Evaluating a Dairy Herd Improvement Project in Uruguay to Test and Explain Q Methodology

Brett Kramer
Graduate Student
Iowa State University
206D Curtiss Hall
Ames, IA 50011-1050 USA
bkramer@iastate.edu

Pedro de Hegedus
Professor
Social Sciences Department
Agriculture College of Uruguay
Garzon Avenue 780
12.900 Montevideo, Uruguay
phegedus@adinet.com.uy

Virginia Gravina
Professor
Statistics Department
Agriculture College of Uruguay
Garzon Avenue 780
12.900 Montevideo, Uruguay
virginia@fagro.edu.uy

Abstract

This paper has two purposes: 1) to describe the evaluation of a dairy herd genetic registry project in Uruguay and 2) to describe and explain Q methodology. The evaluation focused on the complex social, economic, and contextual reasons why some producers in Uruguay had not participated in the registry. Q methodology was used to illuminate non-participating producers' perspectives in a way that kept their viewpoints whole yet provided for a reduction of the data into a few manageable and interpretable factors. These factors (considered aggregated perspectives) were to be analyzed in the Spring of 2003 to identify points of agreement and areas of divergence among them in order to inform program planners on possible future marketing strategies. Therefore, the following paper relies on qualitative data systematically collected from participant interviews and observations.

Four emergent themes explaining producers' non-participation were identified: productive efficiency (nutrition and herd health were more important, thus prioritizing financial resources); unaware (producers were thought to be simply unaware of the project's existence); economic resources (low milk prices and costs related to the project depleted available financial resources to participate); social pressures (lack of time and competition with local producer organizations discouraged involvement). Practical implications point to the suitability of using Q methodology in an agriculturally related evaluation context where tapping diverse and sometimes ignored voices is critical for program improvement.
Introduction and Setting

According to the International Dairy Federation, world milk production was forecasted to exceed 501 million tons in the year 2002 (IDF, 2002). In 2001, Uruguay produced 1.2 million tons of fresh milk with approximately 350,000 cows. Its northern neighbor, Brazil, was responsible for over 22 million tons of fresh milk from over 16 million head of cows in 2001. In light of this level of milk production, Uruguayan levels pale in comparison. Moreover, Argentina - Uruguay’s southern neighbor - produced 9.4 million tons with roughly 2.4 million head for the same year. Staying competitive with its immediate neighbors is an important concern for the Uruguayan dairy industry. The following paper describes an evaluation of a project designed to assist Uruguayan dairy producers in their efforts to remain competitive regionally and provide domestic foodstuffs.

In 1997, the Instituto Nacional Mejormiento Lechero (INML) was formed in Uruguay to help improve the dairy industry. The INML fairly closely resembles the Dairy Herd Improvement program (DHI) in the United States and translates from Spanish to "The National Dairy Herd Improvement Institute". The INML is financed through user fees and the monetary contributions from seven agricultural agencies. This broad support suggests its importance and indicates the many sources of funding necessary to implement agricultural development projects in Uruguay. Central to its mission is the goal of assisting producers to produce more milk through better informed production decisions, especially those based on expected progeny differences (EPD’s) data. In 1998, INML program planners began a project to improve the genetic base of the dairy industry in Uruguay through a genetic registry. The genetic registry entails recording individual-level production data from producers’ herds with the expectation that producers will use the data to make better production decisions, specifically those related to culling unproductive cows.

Purpose, Objectives, and Evaluation Questions

There was one specific programmatic issue on which the evaluation was expected to shed light. In 2002, INML had approximately 200 dairy herds registered out of the roughly 6,000 dairy herds in Uruguay. Although INML had managed to attract the involvement of many large producers and a scattering of medium to small-sized dairy producers, more widespread participation, particularly from operations with fewer than 100 head, had eluded them. This situation made logical sense when viewed in light of two factors: cost and price of milk. The user fees that partly support the registry are computed on the number of production units in the registry, thus making it more cost effective for larger producers than for smaller ones. Moreover, the price of milk in Uruguay is at its lowest price in years – approximately seven cents per liter (compared to roughly 32 cents per liter in the United States). However, despite these two barriers, INML program administrators believe that the program can be beneficial to a large number of producers in the country.

In October of 2002, the authors, in collaboration with program planners, began a focused evaluation of the genetic registry project in Uruguay. However, in addition to assisting INML personnel with an evaluation undertaking, the authors also wished to use the context to test a novel approach to evaluation, termed Q methodology. The evaluation
question posed by the INML would therefore serve as a way to learn about Q methodology in the context of evaluating a practical and important agricultural project.

The process began by identifying two program population groups along with the respective key evaluative questions for each group. The first population group about which program planners had questions was composed of individuals who were familiar with, or “inside” the system, such as program technicians, producers who had cattle registered with the project, and program personnel. In regard to this group, three questions were formulated:

1. What is the current level of knowledge in regard to the project’s purpose, function, and services?
2. What was the principal reason for producers to register their cows with the project?
3. What improvements can be made in the project and what does the current system lack in the way of delivery?

The second population group about which program planners had questions was limited to those individuals who had no knowledge of, or association with, the genetic registry project. The main focus in this population group was producers who had not registered with the project. Three evaluation questions were identified to ask of this group:

1. What is producers’ current level of knowledge with regard to the registry?
2. If producers are aware of the registry, what are the reasons behind their lack of participation or registration with it?
3. What changes or additional services could the project undertake in order to involve producers in the project?

Discussion with program planners developed the process of the evaluation, which focused on the second population group and can be summed up in the following question: Why have producers in Uruguay decided NOT to participate in the genetic registry project? More specifically, what economic and social forces exist and interact to form the basis on which this decision is made?

Due to the complexity of economic and social forces – and how producers perceive and respond to them – the authors and program planners viewed Q methodology as an appropriate alternative to conventional research and evaluation methodologies. The authors hoped that the methodology would function well in uncovering diverse, expected, and unexpected orientations toward the program. Moreover, the authors hoped that the methodology could serve a more practical purpose: that of identifying strategic points of consensus and difference in producers’ holistic perspectives that program administrators could leverage to increase participation in the registry project.

Methodology

Q Methodology has a rich, if little-known, history. It is not surprising then that it does not receive a great deal of attention by teachers of research design courses. In response to reviewers of this paper proposal, the authors will introduce the history, theory, mechanics, and distinguishing characteristics of Q methodology.
In 1934, British psychologist and physicist William Stephenson (a student of Charles Spearman) penned a letter to the editor of Nature magazine (See Stephenson, 1935). In it, he wrote that he had undertaken work on re-conceptualizing correlation analysis such that in place of correlating tests vis-à-vis random variables believed to be expressions of traits, he had developed, instead, a method to correlate whole persons. Specifically, that to which Stephenson referred would later grow into a scientific method of its own right – what is currently known as Q Methodology.

Q Methodology (hereafter simply referred to as Q) involves the study of human subjectivity: the self-referential frame through which human beings define and express their world. Q is more than a technical data analysis tool. It is a way of approaching the study of human behavior with its own epistemology and ontology. Q has been used to explore phenomena in fields such as food and agricultural policy (see Pelletier, et al., 1999), political science (see Lipset, 1963), public policy (see Focht, 2002) communication (see Stephenson, 1967), public health (Dennis, 2001), and psychology (Block, 1961).

Q shows up only sparingly in the evaluation literature. In fact, only one such article was unearthed (Garrard, J., & W. Hausman, 1985), although from personal communication with the principal author, other Q researchers have used the methodology in evaluative undertakings. Q could be useful to evaluation, as one of many types of inquiry.

Central to Q is concourse theory (Stephenson, 1978). A Q concourse can be thought of as a population of statements, thoughts, visual depictions, or many other such human expressions. For example, in any given program, there are different opinions, perceptions, feelings, thoughts, and/or ideas about what it is like to be part of the program – or even outside of it. These can be captured and recorded using either qualitative data gathering techniques (i.e. interviews) or other techniques such as document review (e.g. reviewing the program’s stated goals and objectives) or survey techniques. These “statements”, so gathered and recorded, are kept close to the language in which they are originally expressed.

Q concourses may be virtually limitless (such as people’s views on capitalism) or relatively discrete (such as Scriven’s theory of program evaluation). However, the concourse itself (meaning the entire population of statements) does not provide an efficient nor structured way through which program participants, planners, and administrators can “interact” with it. For this reason, a Q sample needs to be taken from the Q population, or Q concourse. The Q sample, like many samples, is not undertaken haphazardly. The structure of the sample is best driven by theoretical concerns so as to provide a subset of the concourse in relation to the particular issue at hand. In evaluation terms, it makes for good practice to structure the sample according to the evaluation questions or theory deemed to be appropriate and suitable to the context.

The authors and INML program planners worked to identify both program participants and non-participants to interview. It was our assumption that the two groups would have different stories to tell about the genetic registry project. More specifically, we wanted to understand why dairy producers were not participating in the project from the point of view of participants and needed to capture verbatim statements in order to
implement the next step of Q. We interviewed one dairy cooperative administrator, three technicians (two of which were not connected to the project), and four dairy farmers (all of whom who had not participated in the project). All interviews were transcribed in the speakers’ native language (Spanish) and coded according to emergent and theoretical themes.

The theoretical structure develop for use in this study consisted of two main dimensions (also referred to as “main effects”) with two “levels” within each of these, thereby resulting in the 2x2 matrix displayed in Table 1. The two main dimensions deal with perspective (that of the farmer and that of the larger context) and pressures (economic pressures and social pressures), resulting in four cells that serve to structure and organize the Q sample.

Table 1: Theoretical Structure of the Uruguayan Q sample

<table>
<thead>
<tr>
<th>Main Effects Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perspectives and Pressures</td>
</tr>
<tr>
<td>Farmer (a)</td>
</tr>
<tr>
<td>Context (b)</td>
</tr>
</tbody>
</table>

The coded data (i.e., statements) from the interviews were then divided according to the relative “fit” of each within one of the four cells (see Table 1). It is worth noting here that although the Q researcher may choose to identify a particular statement with a specific cell or category, this a-priori “labeling” makes little difference to the subsequent interpretation of the data. No assumption is made that the statements themselves “measure” the identified categories or the theory or structure that undergirds the sample. In other words, what Q concerns itself with more directly is the use of theory, and not an attempt to prove it directly (Brown, 1993). Rather, what we search for is a manifestation of the theory used to build the Q sample by way of the subsequent factor analysis and interpretation. The meaning we strive to find via Q does not reside in the statements; rather meaning is constructed by the study participants as they construct their Q sort. This is what Stephenson (1963) alluded to when he wrote, of meaning being not in the statements themselves, but in terms of what people do with them”.

By holding too strongly to a belief that any one particular statement “means” something definitive (usually to the researcher), the researcher risks losing valuable insight into a person’s pattern of subjectivity. And this is not merely a risk to those who are adept at and familiar with survey research and modern questionnaire design. Interpretive researchers may also fall prey to such a trap by placing too much importance on a-priori interpretation of the rich qualitative data often obtained through interviews. What distinguishes Q from either of these two approaches is that it privileges the respondent’s inference as to what the statements “mean” via the patterning in the Q sort. In this way, Q is more “informed” than either of the two previously mentioned approaches to inquiry because analysis and interpretation of the phenomena is directly aided by the subjects themselves: we allow them to “speak” for themselves via the Q sorting task.
While sampling theory of persons oftentimes uses random sampling, Q most often harnesses the flexibility of purposive sampling (of statements as well as persons, as will be addressed shortly). However, as with most cases in survey research, the goal is to provide a miniature of the population that mirrors the larger one in terms of its comprehensiveness, without sacrificing representation. That is, statements are homogeneous in respect to their kind (meaning relating to the same thing) but heterogeneous in regard to variance inherent in difference. To provide a concrete example, take the Uruguayan Q sample: statements were alike with respect to the broad topic of non-participation in the project yet diverse with respect to the specific mechanisms underlying the non-participation. Furthermore, care was taken to ensure representativeness. To achieve a representative Q sample, a simple yet effective framework is put in to practice: the Fisherian design.

The Fisherian design is useful in that it enables the researcher to “carve up” his/her theory into applicable dimensions. Many times, more than one dimension is at issue and so the Q sample design takes the form of a matrix or table. In the Uruguayan case, the authors, in consultation with the program administrators, decided that the sample had two fundamental dimensions: that of context (farmer and technical) and that of pressures (economic and social). In this way, the design proceeded along a 2x2 matrix, resulting in four cells.

It is important that the Q sample be manageable in terms of size: it is very difficult and time consuming for respondents to distinguish between and among upwards of 100 items in a sample. Therefore, Q samples generally tend to number roughly between 30 and 60, with the exact number being decided by the number of replicates in a given Fisherian theoretical structure. For example, in the Uruguayan design, the authors chose to sample 8 statements from each of the four cells in Table 1 (i.e., ac, bc, ad, bd), thus resulting in a Q sample of 32 statements.

One aspect that needs addressing at this point is the need for balance within a Q sample and specifically, within a given cell of the Fisherian design. In this context, “balance” refers to the respondent having an equal opportunity to react positively and negatively to items in at least one of the main dimensions (such as perspective or pressure). Therefore, within one cell (such as that would be created from combining farmer and economic), four statements would be chosen that generally reflect a positive assertion while four statements would be chosen that generally reflect disagreement with the positive assertion (Stephenson, 1953, 79). However, caution must be employed to avoid selecting statements that are antonyms (e.g., “high” and “low”) for they serve little purpose in illuminating the more fine grained discrimination reflective in most concourses. Additionally, it is fruitful to select statements that exhibit a certain degree on continuum within each sub-cell. This allows the respondent to respond to, interact with, and construct his or her own interpretation of the subtle shades of meaning between and among statements. To write here that the task of selecting “positive” and “negative” statements is an easy one would be misleading; it is anything but and requires a great deal of effort, analysis, and sometimes, careful editing of statements. That is not to say it is impossible, even for the most novice Q researcher.
Once the Q sample has been drawn and statements numbered randomly, they are submitted to respondents for the Q sorting task. Respondent selection ("p-set", meaning “person-set”) advances similarly to the Q sample selection in that the persons are selected purposefully in relation to the probability that the perspective they might offer is fundamentally linked to the evaluation question under consideration. That is, they are selected because they might have something to say in relation to the topic that is the focus of the Q sample. As is the case with the Q sample, a Fisherian design is helpful in structuring this process.

Respondents are asked to place the statements in an array that resembles a quasi-normal distribution. The distribution is oftentimes more platykurtic (i.e., “flatter”) than a normal distribution but nevertheless retains the shape and properties of symmetry (see Figure 1).

![Figure 1: Example of a Q sort array for a 32-statement Q sample](image)

Respondents are directed to begin sorting the Q sample into three piles: statements on the left reflective of those *most unlike* the respondent; statements in the middle having no relevance for the respondent; and finally, statements on the right reflecting those *most like* the respondent. Once all statements have been placed into their respective pile, respondents are instructed to select the two statements (or however many places there are in the extreme tail of the distribution) the respondent feels are most uncharacteristic of his or her position and place them on the far left of the sorting surface. Once complete, the respondent is then instructed to select the appropriate like number of statements that are most reflective of him or her and place them on the far right of the sorting surface. The respondent then proceeds to work alternately from opposite ends of the distribution, finally arriving in the middle – the location of least relevance. After the sorting has been completed and the respondent carefully re-examines the entire array of statements, he or she is asked to record the array on a sheet of paper. This is done by writing the number of the statement in the cell that corresponds to where it was placed in the distribution.

At this time, it is sometimes worthwhile to collect any demographic information about which the researcher may be interested on the same piece of paper. For example, one might collect sex, programmatic role (e.g., participant, administrator, and planner), and other information deemed critical for the subsequent analysis. For the study that is the topic of this paper, the respondent’s diary herd size, programmatic role, current level of awareness about the program, and area of residence (e.g. department) were collected. Finally, the sheet of paper is coded according to preferred conventions.
It is at this stage that the Q sorts are analyzed, generally with the assistance of modern computing technology. The sorts are first correlated, and then submitted for factor analysis. In general, two methods of factor analysis are most widely marshaled for this task: centroid and principle components. Of the two, principle component analysis is by far the most recognized and used method of factor extraction. However, it has its limitations in the application of Q methodology, too involved to go into detail here. Suffice it to add that there exists debate within the Q community over this very issue despite Stephenson’s strong theoretical arguments in support of the centroid method.

Factor analysis on its own is of limited use without rotation. To simplify, rotation consists of changing the reference points of the geometric coordinate system to more closely fit the data and obtain “simple structure”. Simple structure refers to a situation in which individuals’ Q sorts are maximized on one factor with near-zero loadings on all others, thus enhancing interpretability of the results (McKeown and Thomas, 1988, p. 52). Again, there are two methods most widely practiced by modern Q researchers: judgmental rotation and varimax rotation. While judgmental rotation does not have the convenient statistical properties of varimax, what it lacks in this arena it more than makes up for through its flexibility; flexibility made possible through centroid analysis. In summary, the centroid method is most often used in conjunction with judgmental rotation and principle components analysis is most often augmented by varimax rotation.

The result of this statistical manipulation (regardless of what methods are used) is factors that reduce the data into a few perspectives held in common, or individually. The outcome from the data analysis is by no means prescribed and regardless of how many perspectives the researcher feels are present before the analysis, there is no set number of factors (i.e. perspectives) that result from any one particular analysis. Indeed, with judgmental rotation, the researcher is free to pursue several different factor analytic “solutions” that carve the data in different ways so as to illuminate exactly where perspectives merge – and where they diverge. The authors are anxious to share an example of this from the Uruguayan data at the conference.

**Findings**

Q sort data collection was not complete as of the time of this paper submission. INML staff and the Uruguayan authors anticipated collecting the Q sorts in February of 2003. The paper will therefore rely upon qualitative data that were systematically collected through formal interviews and participant observation. These data are all bounded by the concept of motivations for non-participation in the project.

**The Productive Efficiency Perspective**

Interviews and conversations with project personnel and technicians revealed an overriding theme: concern that the project, while laudable, failed to address key aspects critical to the dairy operator’s productive efficiency, such as nutrition, herd health, and management inputs.
This concern was primarily voiced by a high level dairy cooperative administrator as well as by technicians unrelated to the project. Although there is little apprehension over the potential benefits farmers might enjoy from registering their animals with the project, there exists a great deal of anxiety over how well the project can address other factors believed to be more critical to farmers’ productive capacity and economic profitability. As one respondent put it (this quotation, and all others, has been translated from Spanish):

> What good does it do to identify genetic traits that can raise the daily milk production of one cow by two to three liters per day when the cow’s post-partum interval is six to nine months?

The respondent added that gains from improved genetics could not be offset by the negative impacts of poor nutrition. What this illuminates is a belief that farmers are aware of the potential benefits from genetic improvement but are faced with economic conditions that cannot justify the expense to pay for their participation in the registry. Furthermore, technicians added that genetic improvement had to be viewed within the larger context of the entire dairying operation. That is, when considering increasing the genetic productive capacity of cows, care must be exercised to balance the investment with other farm resources, such as nutrition and herd health. Because farmers operate on such a thin margin, decisions of this type are not taken lightly and are decided, in part, by the relative profit each management option can provide.

**The “Unaware” Perspective**

Marketing the project was a central concern to program personnel. Program personnel believed that a lack of knowledge about the project was contributing to farmers’ lack of participation. Two pieces of evidence provide strong support for this claim. Until 2001, program personnel had marketed the project on a one-on-one basis primarily through farm visits to the largest producers. However, as registration began to drop, they changed marketing tactics. In 2002, program personnel began to make contact with, and visits to, small dairy management groups. These groups are self-organized and usually pool their resources to employ a single technician. Because of the increased number of project registrations from this group, program personnel concluded that the tactic was working and furthermore, that the main reason behind lack of participation in the project was simply that producers were not aware of its existence.

**The Economic Resources Perspective**

The most prevalent of all themes was that of economics. It was clear from every interview that depressed milk prices have crippled the range of management options a producer can afford. This theme pervaded all interviews and ranged from subtle to direct, as can be seen from the following quotation from an interview with a producer: “The only thing that is keeping me from participating in the project is the fact that prices are so low - I just don’t have the money to pay for the cost of the service.”

Yet it wasn’t only low prices in combination with the actual registration cost that served to define the extent to which economics played a role in producers’ decisions: there was also the issue of computing resources needed to enter and send the data to INML.
example, one producer expressed it this way: “I do not have a computer so I would have to buy one. With prices as they are, where would I get the money to do that?”

The Social Pressure Perspective

Perhaps the most unanticipated theme that emerged from interviews was the degree to which social forces were acting upon producers’ decisions to forego registration with the project. The social forces identified were as simple as they were complex. For example, as one producer put it: “My son doesn’t want to help with the operation. I want to participate in the program but I don't have enough help to do all of the things that must be done and type in the information.”

Time, then, was a factor for the above respondent, as it was for others. However, other social forces in existence were not that simple. Producers perceived the project being in “competition” with other local organizations. Although program personnel denied that services were duplicated and in fact were correct in their appraisal (the local organizations, while providing technical assistance, do not provide a genetic registry of members’ cows), producers nevertheless had the impression that some services were duplicated. Furthermore, because both the project and the local organizations required a financial commitment, they were in direct competition for the producer’s financial resources as well as their loyalty. This conflict is interpretable in the following statement from one producer: “Because my local association provides the same thing and helps to influence policy, I don't see why I would participate in this project.”

Implications and Practical Importance

Although the data do not support any definitive conclusions with respect to producers’ non-participation there are implications of the research with respect to Q methodology. Three main points come to mind.

The authors were pleased at how receptive program personnel were to Q methodology. Although they had never heard of, nor used Q, they were open to its potential to highlight the complexities of the issues related to the project. This was unexpected not because program planners appeared rigid, but rather because it is not uncommon for Q researchers to meet methodological inflexibility on the part of those first introduced to Q. Secondly, a hands-on training using Q at the beginning of the research endeavor proved useful in educating program personnel on the intricacies of the methodology. It is our suspicion that this training reaped dividends, as program personnel were motivated, intrigued, and excited about the prospect of trying something new. Finally, it is important for researchers and practitioners interested in using Q to realize that there is a learning curve in undertaking such an endeavor. The methodology continues to challenge the authors’ epistemological assumptions and approaches to inquiry. Because Q borrows from two distinct traditions, forging these into a distinct approach to inquiry is at times daunting.
Notes

1 The latest developments in Q Methodology as well as current applications of this method of inquiry appear in the journal *Operant Subjectivity*, via an on-line discussion group, and at the annual meeting of the International Society for the Scientific Study of Subjectivity. Additional information can be obtained from the aforementioned organization’s website at http://www.qmethod.org.

2 To summarize briefly, the debate over the different methods of factor extraction have more to do with statistical rigidities than anything else. Centroid factor extraction uses an average correlation estimate (on average, the correlation between the sort under scrutiny and all others) to place on the diagonal of the inputted correlation matrix. This allows the researcher to pursue theoretical hunches for it does not require a determinant solution. Principal components analysis, on the other hand, uses a perfect inter-sort correlation estimate (1.0) to place on the diagonal of the inputted correlation matrix. This is very clean from a statistical perspective for it allows the researcher to pursue an ordinary least squares solution. However, it also dictates there exists one “best” solution. Although this may have wonderful statistical results (the sum of the squared differences are minimized), it nevertheless places a restriction both on the data and the researcher, thus discouraging any theoretical pursuits deemed to be interesting. For additional information, see Brown (1980).

3 Rotation in $Q$ is undertaken in order to arrive at one single factor solution at a time, such that sorts are purely loaded on one factor and near-zero on others, thus focusing the lens, so-to-speak, through which we can view the factors and their relation to one another. This "focusing", via rotation, does nothing to disturb the fundamental nature of the data; nor does it change the coordinates of any data point (i.e. Q-sort) in geometric space. What it does do, however, is to aid in the interpretation of factors at the other end. This is done through analyzing the similarities and differences in the now-aggregated Q-sort arrays that represent each factor. More specifically, modern computing technology has advanced sufficiently to provide output that identifies specific "consensus" and "differentiating" statements in the factor arrays, thus simplifying at least the identification of the starting point for the interpretive work that follows.
References


