Field Survey Experience in End-User Behavioral Research

Authors: Liping Gu

Date: 2016-06-07T00:00:00+00:00

Abstract

With an emphasis on research ethics and research attitudes running through the entire text, this paper briefly introduces six survey experiences conducted around 2006-2009 during the completion of the PhD dissertation "Network Information Behavior Differentiated User Models—Creation and Use of Personas," including research ideas and strategies, research perspectives and positions, outlines of research design, the special value of research subjects, sampling and survey verification processes, principles of reasonable inference, and operational steps for organizing SPSS tools, among other aspects.

Full Text

Fieldwork Experience in End-User Behavior Research

FlySheet System Technology Co., Ltd., Taipei 100140

Abstract This paper emphasizes research ethics and attitude throughout, briefly introducing six survey experiences conducted between 2006-2009 before and after completing the doctoral dissertation *Differentiated User Models for Online Information Behavior: Creating and Using Personas*. The discussion covers research approaches and strategies, perspectives and positions, design frameworks, the special value of research subjects, sampling and survey review processes, principles of reasonable inference, and step-by-step procedures for using SPSS tools.

[**Keywords**] user behavior; sampling survey; hypothesis testing; reasonable alternatives; methodology; research ethics

1 Research Approach and Strategy

Contemporary library science primarily faces the impact of the internet. New web technologies pose tremendous challenges to digital libraries and informa-

tion services, decisively influencing technical research in digital libraries. Alongside technological innovation, the processing, analysis, and computation of massive information have become new topics in informetrics, bibliometrics, and scientometrics. Changes in information organization, retrieval, and needs have prompted shifts in user behavior, moving beyond traditional library science assumptions. Ultimately, these cumulative changes and impacts have fundamentally transformed the nature of libraries, giving library science new directions and research questions. Thus, studying library and information science today is indeed timely.

However, modern information analysis is increasingly complex, library services continue to expand, and digital library construction involves participants with diverse skill sets. This places new demands and challenges on the disciplinary backgrounds and research work of knowledge service providers. Professionals must possess a broad interdisciplinary knowledge base to develop unique insights and innovations in their library and information work from a wide perspective. They must also start from broad, novel, and even unconventional ideas, gradually focusing on two or three specific frontier issues through literature review and discussion, integration with practical work, and rigorous scientific research processes. Only after delving deeply, tackling challenges, and overcoming various difficulties can they propose innovations and contributions.

For end-user behavior research, there are two basic approaches: (1) starting from library science and integrating other disciplines; or (2) using the practical foundations of other disciplines—such as management, sociology, psychology, and economics—as guiding frameworks or theoretical sources to address difficulties in library and information development, establishing hypotheses using other disciplines' content, conducting empirical investigations with statistical or other methods, solving specific problems currently facing libraries, and obtaining reliable conclusions.

New research projects always encounter numerous difficulties. Before execution, there are imaginable challenges; during execution, unforeseen difficulties arise. Even after completion, the truly important questions may only become apparent. The research experience introduced here, from *Differentiated User Models for Online Information Behavior: Creating and Using Personas* (hereinafter referred to as *Personas*), began by establishing hypotheses from information-seeking behavior, using falsification to obtain basic behavioral descriptions. During interpretation, reasonable alternative hypotheses were constructed from several psychological and sociological theories for research inference, thereby developing a hypothesis. This paper does not focus on *Personas* itself but outlines the methodology of the fieldwork conducted at that time, hoping to serve as a modest spur to induce more valuable contributions.

2 Research Perspective and Position

The writing norms and elements of contemporary research papers include identifying problems, posing questions, solving problems, and specifying the methods, techniques, and tools used to address these issues. However, since serving on the editorial boards of *International Journal of Library and Information Science* and *International Journal of Science and Technology Education Research* from 2009 to the present, I have frequently reviewed numerous papers from South and Southeast Asia. Most rejected papers suffer not from lacking these elements but from insufficiently clear positions. Among many review comments, the most common is "Please clarify your research value and significance." In other words, we hope researchers challenge their own aspirations, demand more from themselves, and rigorously examine their work to align with field realities and clarify their research purposes and directions. Simply put, simplicity far surpasses cleverness.

Furthermore, for end-user behavior research, we also pay attention to details revealed in research reports, examining whether researchers have stepped into ethical and political pitfalls. The most basic requirements include:

- **Voluntary participation principle**: Not forcing respondents to participate; all participants join voluntarily.
- **No harm to respondents principle**: Research content does not involve personal evaluation, privacy, or the attitudes of research units.
- **Anonymity and confidentiality principle**: Using Solomon encryption for all respondent data and storing it securely.
- **Transparent research process principle**: Researchers disclose their names and affiliations; research questions do not involve special care or welfare for individual respondents by their institutions.
- **Neutral research report principle**: Researchers specifically point out the limitations of their research in reports rather than concealing them.

However, some important aspects are invisible to reviewers. For example, another practical issue beyond the research itself is: How do researchers face their own and their institutions' vested interests? Since scientific research itself is a form of social and cultural practice, it inevitably involves questions of: From what perspective? What issues are of concern? What conclusions are intended? Everyone has vested interests in the outcomes of a particular research endeavor. If researchers are clearly aware of their own vested interests and the complex vested interests of all parties in the research project, they should prevent these interests from affecting the research design and guard against deliberately, intentionally, or unintentionally misleading results.

In fact, once these points are clearly thought through, it becomes easy to find entry points and breakthroughs from one' s own research questions and others' research findings. Research ethics is not moral preaching but the most critical aspect of all research. It is no exaggeration to say that a researcher' s research ethics determines the height and contribution of their research. Other research

difficulties can be solved—such as seeking guidance for SPSS, redesigning or remedying flawed research designs, or improving weak inferences. Only research ethics exists in one' s heart, permeating the entire research process, and must be handled with utmost caution.

Researchers should maintain the conscience of an intellectual and be vigilant against "policy-driven conclusions."Years ago, before *Personas* took shape, when it was only known that certain social science surveys needed to be conducted, both the advisor and the researcher faced the question: "Whom am I researching for?"This simple self-questioning had a decisive impact on the eventual formation of *Personas*.

3 Outline of Research Design

Contemporary researchers no longer engage solely in theoretical reviews and writings. Due to the development of library and information work, almost all disciplines (at least the science and engineering disciplines at the Chinese Academy of Sciences) treat "literature review" as something that "must be done" rather than the "only thing to be done," and regard "innovation" as the decisive condition that "must be achievable."

Years ago, after careful study, we concluded that regardless of which subfield or specific topic one starts with, the innovation process still revolves around three aspects: (1) Literature review—identifying unsolved problems—using old methods to solve new problems; (2) Research methods—identifying limitations of old methods—developing new methods; and (3) Research results—summarizing results of new methods and new problems—generating new fields. These three aspects are interlinked, forming innovative research. Special attention must be paid: the proposal of new problems, new methods, and new fields must be based on empirical evidence that can be replicated or modified by other researchers.

Using the research design of *Personas* as an example, the initial plan was to conduct six identical questionnaire surveys as a longitudinal study. Later, many practical execution issues were discovered, so it was changed to six different questionnaire surveys as a cross-sectional study. Halfway through the implementation, we found it executable but unable to deeply describe end-users' actual work content and online behavior, so interview and natural observation methods were added, shifting to a mixed-methods study.

During overall design, adjustments can be made according to circumstances, but during actual execution, the plan cannot be changed arbitrarily. Therefore, to ensure the research plan is completed "on time as agreed," we consulted many senior researchers' reports and textbooks during the design phase to establish some rigid rules. For example, for the rigid processes of questionnaire and interview surveys, the key points are:

Questionnaire survey specific steps: (1) Determine questionnaire framework and main content; (2) Design specific questions; (3) Pilot test questionnaire

in small scope; (4) Revise questionnaire based on pilot issues; (5) Design distribution and collection methods; (6) Distribute questionnaire; (7) Respondents answer questionnaire; (8) Collect questionnaires; (9) Track non-respondents; (10) Organize questionnaires and eliminate invalid ones; (11) Code answers for processing and analysis.

Interview survey specific steps: (1) Determine main interview content; (2) Determine basic interview format; (3) Design interview outline; (4) Arrange interview process; (5) Conduct interview; (6) Organize interview records; (7) Recall interview situations and reorganize records; (8) Summarize interview results for processing and analysis.

By fully considering every detail of research ethics and maintaining a firm and clear research attitude, we could accurately assess the feasibility of the research design and the efficiency of actual execution under "justifiable" circumstances.

4 Special Value of Research Subjects

When conducting user research, it is crucial to be very clear about who is and who is not a research subject; otherwise, subsequent research inferences cannot be made. Using *Personas* as an example, we initially chose master' s and doctoral students at the Graduate University of the Chinese Academy of Sciences as research subjects because of their special characteristics. Unlike the previous generation of researchers, this generation' s basic demographic, social, economic, and network communication parameters at the Chinese Academy of Sciences all had new values compared to the past 15 years. The total number of scientific researchers grew exponentially, with their proportion within the Academy increasing significantly. Moreover, under the unified enrollment management of the Graduate University, they were assigned to various research institutes directly under the Academy. Their situation in these institutes was as follows: together with regular staff, they transmitted technology, methods, theories, research norms, and values through laboratories or individual guidance. These senior colleagues not only imparted their technology, methods, information, and theories to newcomers working with them but also passed on the research institute' s norms and values to their colleagues.

Studying these subjects does not easily yield universal user behavior, but their sample representativeness is special and their characteristics concentrated, making it easy to find interesting or novel discoveries. However, on the other hand, this is also very challenging and prone to poorly designed or poorly executed problems. The way to overcome this difficulty is to repeatedly and carefully consider research ethics and attitude. During the research process, there was almost daily internal struggle: direct impact from users continuously revised the researcher' s existing knowledge structure and "common sense." This training was very helpful for future research projects and even for working in different libraries and companies. On the other hand, although facing different audiences, the purpose of *Personas* was to improve library systems and services rather

than increase the researcher' s publication count and citations. The research position was from the readers' standpoint rather than the library director' s and librarians' standpoint, and the research interest was in readers' benefits rather than the director' s and researcher' s benefits. Therefore, when the researcher later faced various voices claiming to "represent users," they could directly refer to the fieldwork experience with end-users years before and intuitively detect the hidden intentions and unapparent crises in matters. To others, this seemed like a special ability, but it was simply the result of hard work years ago and periodic review.

5 Sampling and Survey Review Process

Sampling seems to be a very complex and difficult task. However, if one masters the principles of research ethics and attitude and keeps them in mind, considering how to sample is not too difficult. Although there are many sampling methods, techniques, and scales, what use are they if they do not conform to research ethics and attitude? Conversely, even without understanding sampling, as long as one grasps research ethics and attitude, one can derive relatively simple sampling methods. If one only emphasizes sampling techniques while neglecting research ethics, the final result will inevitably be dramatic research failure, which must be handled with caution. Below is the thought process before the *Personas* research.

5.1 Is the Sampling Frame Available?

The biggest operational difference between random sampling and convenience sampling is that random sampling selects respondents (samples) from all survey objects (population). Therefore, without a sampling frame containing information on all population members, random sampling is impossible. At this stage, besides presenting a clear research proposal, commitments and specific measures for information confidentiality were also made.

5.2 Is the Sampling Frame Complete?

Having a frame does not equal mastering relevant population information. For questionnaire surveys, the frame must at least record: name, department, whether studying in Beijing area, and whether still enrolled. This information allows researchers to judge whether random sampling is feasible. In fact, the frame included personal information such as gender, degree level, grade or enrollment year, and age, but did not include contact information. Researchers at this stage began to judge which type of random sampling to adopt.

5.3 How to Conduct Sampling?

After careful consideration, *Personas* adopted stratified random sampling, sampling according to the proportions of students in various Beijing-area institutes, gender ratios, doctoral-to-master's ratios, and grade ratios. From each sampling

unit, selections were made proportionally to the unit' s size. Its advantages are: (1) ensuring representativeness of characteristics forming the classification basis, thus producing less variability than simple random or multi-stage random sampling; (2) reducing the chance of population members being excluded through the classification process; and (3) enabling estimation of characteristics for each stratum, thus allowing comparisons. However, its disadvantage is the possibility of misclassification, which increases variability. To overcome this defect, a brief study of variation coefficients was conducted. Although the possibility could not be verified, structural equation modeling was accidentally learned. Consequently, purposive sampling was adopted for the interview survey. Based on questionnaire survey analysis results, a sub-group was selected from the population as representative of the whole through judgment, with a portion of this group taken as the sample.

5.4 How to Effectively Distribute Questionnaires?

This relates to: (1) time, location, and manpower—the longer the distribution period, the more dispersed the locations, the more effective, but the higher the manpower requirements; (2) manpower, gifts, and funding—more manpower and gifts increase the number of distributed questionnaires, thereby increasing research validity, which in turn allows sacrificing validity for reliability to solve reliability deficiencies caused by sampling; and (3) questionnaire quantity, number of questions, gift value, and funding—more questionnaires should have fewer questions; there is a statistical formula to calculate reference values, which was only learned after completing *Personas*, though *Personas* did meet standards within the "safe zone." However, from a practical perspective, rather than considering questionnaire quantity and questions, it is better to consider more the administrative factors and respondents' expected psychology that cause response bias and significantly reduce reliability.

5.5 Will Samples Be Willing to Respond?

Willingness to respond includes two issues: (1) How is the survey conducted? This is the questionnaire distribution issue mentioned above. (2) Does the questionnaire itself cause respondent anxiety? There are many causes of respondent anxiety. Besides the distribution and collection process, the most important is whether the questionnaire can be easily answered. For example, it is inappropriate to ask about too new topics, such as recently emerged internet terminology. If asking about the latest online behavior is necessary, then questions about understanding the term must be included, or the term must be explained in a few sentences. However, this creates pressure on respondents and affects their answers. To reduce this pressure, questionnaires can be designed structurally, such as classifying questions first and using five-point scales for each question, so respondents spend time reading questions but not options. However, this approach creates four situations that reduce reliability: random answering, lazy answering, inertial answering, and response fatigue (less concentration toward

the end).

Additionally, interview question response methods should be as diversified as possible. Structured questions ensure comparability of data across individuals, time, and researchers, making them particularly useful for studying variable questions that examine differences between things. Conversely, unstructured methods allow researchers to focus on specific phenomena that differ from others and require targeted approaches. They abandon universality and comparability in pursuit of internal validity and contextual understanding, making them very useful for revealing processes leading to specific results.

5.6 How to Effectively Recover Questionnaires?

Time, location, manpower, gifts, funding, and external factors must all be considered for questionnaire recovery. But convenience is also crucial. Effective recovery depends on the steps and time required for questionnaires to go from respondents to researchers. For example: (1) Mail survey recovery includes putting it in a return envelope, sealing it, finding the mailbox, and delivering—at least 6 steps. (2) Electronic surveys require remembering the task, finding the URL, and clicking complete after answering—at least 3 steps. (3) Face-to-face interviews, whether collecting at appointed times or filling during interviews, require remembering collection or interview times and finding the researcher—at least 2 steps. (4) Telephone surveys require none of these steps. Interestingly, more recovery steps relatively allow more questionnaire items. The disadvantage of mail surveys is that the sampling frame lacks correct current addresses, and the recovery steps are too troublesome for respondents. However, compared to telephone interviews that cannot ask many questions and electronic questionnaires with uncertain recipients (and no email addresses in the sampling frame), mail surveys are suitable for questionnaire distribution. For interview surveys, considering the smaller number of people, face-to-face interviews were adopted.

5.7 How to Track Non-Responding Samples?

Ideally, tracking is unnecessary because even if tracked, at least one more hypothesis test is required: assuming the most recently recovered questionnaire answers are no different from previous ones. In random sampling research, even without tracking, "missing questionnaires" and "invalid questionnaires" must be analyzed to discuss whether missing sample characteristics affect sampling—that is, whether there is systematic error.

Tracking methods relate to initial distribution and recovery methods, but generally include telephone calls, follow-up mail, email, revisiting interview locations, and indirect communication through primary and secondary contacts. All involve the accuracy of the sampling frame. Theoretically, everyone believes that through diverse and different contact methods, non-responding selected respondents can always be "tracked" or "found." But in practice, due to sampling frame inaccuracy, final research data becomes highly "dangerous." Before the

Personas research, the following issues were fully considered: (1) **Possibility of minor errors**: The sampling frame was full of "false" information; many people did not provide real data at the time, so halfway through the sampling survey, we found the sampling invalid and needed to temporarily modify the plan. (2) **Possibility of moderate errors**: Later, through many methods, we contacted respondents indirectly. Respondents knew the researcher was "serious" and thus replied, but the questionnaires were full of corrections from "serious" answering, which in fact increased the invalid questionnaire rate for researchers, and sometimes the respondent answering was not the selected person at all. (3) **Possibility of major errors**: Research team members, not considering the above statistical errors, compiled the data with other data into the database. After descriptive and even inferential statistics were completed, due to theoretical model instability, data had to be refitted, revealing a large number of possible erroneous values in the database. At this point, data cleaning had to be redone and analysis repeated from scratch, and error rates had to be analyzed (comparing results from both analyses). Worst of all, if initial data had not been archived, there would be no chance to redo the work. Researchers could not perform "debugging" and "error correction," causing complete research failure.

5.8 How to Effectively Reduce Sampling Error?

In the *Personas* research, two methods were adopted: (1) Maintaining equal proportions between sampling ratios in each stratum and the departmental structure ratios of the Chinese Academy of Sciences Graduate School Beijing area (equal proportion), so that the composition of sample survey objects could be selected proportionally to the population, thus reducing error compared to simple random sampling. (2) Flexibly changing sampling ratios across strata (unequal proportion) because the sizes of various departments in the Chinese Academy of Sciences Graduate School Beijing area differed. Larger departments had higher sampling ratios, also achieving the purpose of reducing sampling error.

5.9 Which Questionnaire Distribution and Recovery Method to Adopt?

In terms of questionnaire item measurement, *Personas* used categorical scaling to measure psychological responses such as attitudes, opinions, evaluations, and impressions. The mail survey method was adopted based on the special circumstances of surveying the Chinese Academy of Sciences Graduate School Beijing area: (1) Because survey objects were mostly busy with work and inconvenient or less available for face-to-face interviews, mail surveys saved their time without requiring them to take extra time to contact and wait for researchers. (2) Because the various departments of the Chinese Academy of Sciences Graduate School Beijing area were not centrally distributed, making sample distribution overly dispersed, only mail surveys were relatively suitable for such distributed academic institutions. (3) Chinese Academy of Sciences graduate students stud-

ied, researched, or worked in their relatively independent institutes. Because the Academy itself 汇集了多所国家级重点实验室 (houses many national key laboratories), access control, security inspection standards, and information confidentiality regulations were quite detailed and strict, making procedures very cumbersome if visiting a particular institute. For a survey covering many institutes, mail surveys were more appropriate. (4) With a lack of experienced interviewers, mail surveys were one of the effective ways to reduce manpower while obtaining reliable information.

5.10 How to Avoid Non-Response Bias?

Mail surveys have low recovery rates and are prone to non-response bias—in other words, respondents' interest in the survey topic more easily influences the entire survey results. To prevent this, several effective methods were adopted: (1) Avoiding vaguely themed questionnaires by defining the topic as a graduate student-related subject, "Chinese Academy of Sciences Graduate Student Online Service Usage Survey," to increase respondent participation. (2) Randomly sampling from "non-returned" respondents for face-to-face interviews. (3) Using high-recovery-rate fixed samples (monitor samples) to compensate for errors in the actual survey phase. Because the researcher had participated in or conducted three questionnaire surveys on different times and topics before this survey, there were graduate students who had established trust relationships and served as friends, compensating for some potential survey errors caused by non-response bias.

5.11 Which Depth Interview Method to Adopt?

Because there was a need to analyze different user behaviors based on questionnaire survey results and attempt to obtain explanatory reasons for these behaviors from user interviews, the in-depth interview method was adopted. Because most people cannot accept long interviews, but depth interviews require a relatively long time, on the one hand, a larger sample needed to be selected from purposive sampling, and on the other hand, suitable interview objects needed to be identified from these samples' interview willingness feedback, using quota sampling for purposive sampling.

5.12 How to Ensure Close Integration Between Questionnaire Survey Results and Interview Plans?

Qualitative research can adopt fixed question set interviews (nonscheduled-structured interview). In *Personas*, survey questions were determined based on questionnaire survey results. Therefore, interview object selection was determined according to four user groups classified from questionnaire survey analysis results. However, the expression form, order, or response methods of interview questions were not strictly regulated. Thus, interviewers were allowed to express questions in their own language and according to the specific interview environment, decide question order, allow respondents to answer in their own language

and manner, and allow interviewers to ask further about unclear answers or potentially related other questions. But the entire interview had to closely revolve around fixed content, with no substantial changes to the questions asked.

For data summarization, a tabular approach was used for organization. In addition to summary tables, the researcher repeatedly refined each respondent' s answers. After forming an overall impression of their lives and behaviors, analysis was conducted combining data tables and prior assumptions. In fact, the above errors could appear in any research step. The only thing that could be done was to meticulously record work logs during the research process and properly preserve every data document, then carefully consider the details of the research process after completion. Without adhering to research ethics and with a casual research attitude, it would be impossible to truly complete the entire research plan.

6 Principles of Reasonable Inference

Adopting reasonable alternative hypotheses is one of the main approaches in exploratory research. When researchers want to develop new research fields, they often generate new perspectives through exploratory research. However, its disadvantage is that it rarely satisfactorily answers questions. The solution is to integrate two completely different approaches— "hypothesis testing" and "reasonable alternatives." In *Personas*, because exploratory research could provide hints for research questions and observations of research methods to accurately answer research questions but could not provide absolute answers, the latter half of the theoretical construction process (still at the hypothesis stage) attempted to consider methodological, procedural, and population influence factors to judge the correspondence between survey results and previous hypothesis tests.

In the latter half of *Personas*, the basic strategy adopted was "evaluating reasonable alternative hypotheses" —that is, testing and evaluating whether other alternative explanations exist for a particular phenomenon. This strategy' s applicability is regardless of whether the data nature is quantitative or qualitative, regardless of whether the research method type is experimental, cross-sectional design, longitudinal design, or case study, and regardless of whether the data collection method is observation, questionnaire, or other. Therefore, in mindset, there is an expectation of other methods to interpret findings, and any interpretation is viewed as modifiable after further testing. However, reasonable alternative hypotheses still have certain reasoning limitations and conditional requirements.

This reasoning is closely related to its "unit of analysis" and should avoid "ecological fallacy" and "reductionism" related to the unit of analysis. Units of analysis can be divided into individuals, groups, organizations, and artifacts. Therefore, when *Personas* unit of analysis is the individual—each user—rather than organizations like the Chinese Academy of Sciences or the Institute of Engineering Thermophysics, there is no "ecological fallacy" error of drawing

conclusions about individuals based on observations of an institute or group. Conversely, because *Personas* research used mail questionnaires, returned responses only represented a portion of people. Conclusions drawn from selecting respondents from this portion for interviews were not suitable for inference to all Chinese Academy Sciences graduate students or all Chinese people. Therefore, operationally, it was corresponded to the hypothesis testing in the first half of *Personas*, breaking the "common sense" of online user behavior through counterexample demonstration, obtaining possible differentiated user groups, and thus providing theoretical space for creating personas in the latter half.

In the latter half of *Personas*, considering the influence of "reductionism" (the tendency to assume one unit of analysis or variable is more important or relevant than others, which can lead to fallacious conclusions if inappropriate units are used), reasonable alternative hypotheses were adopted to find different answers from different fields, selecting the *likeliest* explanation for research results (evidence) among many answers. Therefore, reasonable alternative hypotheses include not only those from library and information science (such as information-seeking behavior) but also from other fields (such as psychology, sociology, economics, anthropology, education, etc.).

In other words, "reasonable alternatives" are not arbitrary interpretations according to one' s own wishes, because "reasonableness" does not come from whatever is convenient for the researcher to think, write, and publish. Rather, it is the result of "competition" and "alternatives." Academic theory is not truth but relative reasonableness, and this reasonableness comes from dialectics and competition. This repeated demanding process ultimately presents the research value and contribution of *Personas*—completing a verifiable and simply operable hypothesis.

7 Tools: Operations of SPSS v12

There are no shortcuts in research; only more thinking, more attempts, and more practice. However, following guides can save some exploration time. Below are notes compiled from relevant textbooks and actual operations of SPSS v12 during doctoral studies (2006-2009). As mentioned earlier, tool operation is not difficult at all; following steps can produce data and charts. The key is research ethics and attitude: if the intention is good, even simple descriptive statistics can produce respectable research. Similarly, if one emphasizes dazzling mathematical equations while neglecting simple research ethics, the interpretation under beautiful charts will be full of holes.

- **Simple descriptive statistics**: (1) Open file; (2) Statistics → Summarize → Descriptives; (3) Select variables, click "←" for Variable(s); (4) Options → select all Statistics options → Continue. Statistics include: Mean (arithmetic mean), Variance, S.E.mean (standard error of mean), Skewness, Range (difference between max and min), Kurtosis, Sum, Minimum, Maximum, Std.deviation (standard deviation).

- **How to know if categorical variables relate?−Chi-square analysis (without SPSS)**: (1) Null hypothesis opposite to assumption; (2) Degrees of freedom df = (r-1)(c-1), where r = number of rows, c = number of columns; (3) Look up "chi-square distribution table" based on df and significance level (0.01, 0.05, or 0.1); (4) Obtain "test parameter"; (5) Use formula; (6) Obtain "chi-square value"; (7) Compare "test parameter" and "chi-square value." Large difference rejects null hypothesis, thus supporting the assumption. No significant difference supports null hypothesis, thus the assumption is not established.
- **How to know if categorical and interval variables relate?−One-way ANOVA**: (1) Open file; (2) Statistics → Compare Means → One-Way ANOVA; (3) Select dependent variable, click "→" for Dependent List; (4) Select independent variable, click "→" for Factor. Large difference rejects null hypothesis, thus supporting the assumption. No significant difference supports null hypothesis, thus the assumption is not established.
- **How to know if interval variables relate?−Using Pearson correlation coefficient−R coefficient**: (1) Open file; (2) Statistics → Correlate → Bivariate; (3) Select variables to analyze, click "→" for Variables; (4) Correlation Coefficients → Pearson; (5) Test of Significance → Two-tailed; (6) Options → select all Statistics options → Continue. Interpretation: 0 < |R| < 0.3 = no correlation; 0.3 |R| < 0.5 = low correlation; 0.5 |R| < 0.7 = moderate correlation; 0.7 |R| < 1 = high correlation (can conduct simple regression).
- **How to measure trend between two variables?−Simple regression**: (1) Open file; (2) Statistics → Regression → Linear; (3) Select dependent variable, click "→" for Dependent; (4) Select independent variable, click "→" for Independent(s); (5) Method → Stepwise; (6) Standardized Residual Plots → Histogram.
- **How to describe multiple variables for an individual (variables not correlated)−Factor analysis**: (1) Open file; (2) Statistics → Data Reduction → Factor; (3) Descriptives → select Univariate descriptives, Initial solution, Coefficients, Significance; (4) Extraction → Principal components; (5) Extract → Eigenvalues over → ">1"; (6) Display → select Unrotated factor solution, Scree plot; (7) Factor Scores → Save as variables; (8) Method → Regression, Display factor score coefficient matrix.
- **How to describe multiple variables for an individual (variables not correlated)−Cluster analysis**: (1) After factor analysis, obtain "fac1-1" and "fac2-1"; (2) Statistics → Classify → Hierarchical Cluster → select "fac1-1" and "fac2-1"; (3) Plot → Dendrogram, All clusters, Vertical → Continue; (4) Method → Between-groups linkage; (5) Save New Variables → Single solution 3 clusters.

Experience in user behavior research comes from advisors, teachers, and friends during doctoral studies, who had many direct and indirect influences on the

researcher. Relevant guiding books can refer to Zhang Xiaolin' s *Information Management Research Methods* and Earl Babbie' s *The Practice of Social Research*. The experience of conducting six large-scale social science surveys over three years, although limited to the Chinese Academy of Sciences in Beijing, was helpful for later work supporting the Economics and Finance Department at City University of Hong Kong, institutional ranking and scientist behavior research at National Taiwan University, and developing market customer groups for local and multinational companies.

Although end-users have different attributes, characteristics, behaviors, and positions in various regions and institutions, and serve in different job positions with different standpoints, high sensitivity to user behavior enables quick entry into research frontiers or breakthrough of market barriers, contributing abilities in various workplaces and doing one's part for the library and information cause. This paper serves as a modest spur to induce more valuable contributions.

[**Author Introduction**] Gu Liping, male, born in 1978, Ph.D., has published 33 papers.

Note: Figure translations are in progress. See original paper for figures.