Chapter 10
The Use of Focus Groups in Design Science Research

The only possible conclusion the ***social*** *sciences can draw is: some do, some don't.*

– Ernest Rutherford

Focus groups to investigate new ideas are widely used in many research fields. The use of focus groups in design science research poses interesting opportunities and challenges. Traditional focus group methods must be adapted to meet two specific goals of design research. For the evaluation of an artifact design, exploratory focus groups (EFGs) study the artifact to propose improvements in the design. The results of the evaluation are used to refine the design and the cycle of build and evaluate using EFGs continues until the artifact is released for field test in the application environment. Then, the field test of the design artifact may employ confirmatory focus groups (CFGs) to establish the utility of the artifact in field use. Rigorous investigation of the artifact requires multiple CFGs to be run with opportunities for quantitative and qualitative data collection and analyses across the multiple CFGs. In this chapter, we discuss the adaptation of focus groups to design science research projects. We demonstrate the use of both EFGs and CFGs in a design research doctoral thesis in the health-care field.

10.1 Introduction

The field of information systems has recognized the importance of design science as an opportunity to increase relevance (Venable, 2006). Hevner et al.'s (2004) information system research framework illustrates how both the behavioral and design science research paradigms in information systems follow similar cycles. Behavioral science research identifies a business need and develops and justifies theories that explain or predict phenomena related to this need. Design science research builds and evaluates artifacts that address particular business needs. Behavioral science researchers search for the truth, while design science researchers seek utility (Hevner et al., 2004).

Monica Chiarini Tremblay, Alan R. Hevner, and Donald J. Berndt

A. Hevner, S. Chatterjee, *Design Research in Information Systems*, Integrated Series in Information Systems 22, DOI 10.1007/978-1-4419-5653-8_10,

Design science research can be described as having two phases: the development of the artifact and its evaluation (which cycles for refinement of the design). A design researcher not only designs an artifact that provides utility but also provides evidence that this artifact solves a real problem. In fact, evidence-based artifact evaluation is crucial in design science research (Hevner et al., 2004). This requires that the artifact be evaluated within the technical infrastructure of the business environment. Several artifact evaluation methods have been outlined by researchers, including observation, analytics, experiments, testing or descriptive analysis, and more recently action research (Baskerville and Myers, 2004, Cole et al., 2005, Hevner et al., 2004, Iversen et al., 2004, Lindgren et al., 2004).

In this chapter we propose focus groups as an effective technique to be used for the improvement of an artifact design and for the confirmatory proof of its utility in the application field. We begin with a brief description and history of the focus group technique. Next, we outline the focus group methodology and propose adaptations for the evaluation of design artifacts. Finally, as an example, we describe a recently completed research study in which focus groups were used to both refine and evaluate the design science artifact.

10.2 Research Focus Groups

The focus group technique has long been utilized in social research to study ideas in a group setting (Morgan, 1988). A focus group is defined as a moderated discussion among 6–12 people who discuss a topic under the direction of a moderator, whose role is to promote interaction and keep the discussion on the topic of interest (Stewart et al., 2007). The term *focus* in the title refers to the fact that the interview is limited to a small number of issues. The questions in a focus group are open ended but are carefully predetermined. The set of questions or "questioning route" is meant to feel spontaneous but is carefully planned. Usually, the moderator encourages the sharing of ideas and careful attention is paid to understanding the feelings, comments, and thought processes of the participants as they discuss issues (Krueger and Casey, 2000). A typical focus group lasts about 2 h and covers a predetermined range of topics. Multiple focus groups allow for understanding the range of opinions of people across several groups and provide a much more natural environment than personal interviews because people are allowed to interact, which allows them to both influence and be influenced by others (Krueger and Casey, 2000). This is valuable to gain shared understandings but yet allows for individual differences of opinion to be voiced.

Focus groups have been effective both as a self-contained means of collecting data (as a primary research tool) or as a supplement to other methods of research (as a secondary research tool) (Krueger et al. 2000; Morgan 1988). The focus group technique is particularly not only useful as an *exploratory method* when little is known about the phenomenon but also can be used as a *confirmatory method* to test hypotheses (Stewart et al., 2007).

Originally coined “focused” interviews, focus groups were used during World War II by social scientists to explore morale in the US military for the War Department (Krueger and Casey, 2000, Merton and Kendall, 1946, Stewart et al., 2007). Though invented by academics, the focus group technique was mostly ignored by researchers because of the difficulties in demonstrating rigor in analysis and the fear of possible contamination of the interview process. Focus groups were, however, widely embraced by market researchers in the early 1950s. In fact, the use of focus groups continues to grow in the for-profit sector, accounting for 80% of industry-related qualitative research, and firms have been created to solely support all aspects of focus groups (Krueger and Casey, 2000, Stewart et al., 2007, Wellner, 2003).

In the 1980s, academics re-discovered focus groups as an alternative to other qualitative research, such as interviews and participant observation. Focus groups are now one of the most widely used research tools in the social sciences (Stewart et al., 2007). Researchers in both basic and applied behavioral science disciplines have utilized focus groups as a source of primary data. Education, management, sociology, communications, health sciences (particularly by clinicians), organizational behavior, social psychology, political science, policy research, and marketing are some of the disciplines utilizing focus groups. The diversity of the aforementioned fields suggests that focus groups can be effectively designed, fielded, and analyzed from varying perspectives and priorities.

Information systems’ researchers have called for a broader variety of available empirical methods to improve relevance of research (Benbasat and Weber, 1996, Galliers, 1991) and we have seen increased attention on the use of focus groups in IS research (Baker and Collier, 2005, Debreceny et al., 2003, Jarvenpaa and Lang, 2005, Manning, 1996, Mantei and Teorey, 1989, Smith et al., 1996, Torkzadeh et al., 2006, Xia and Lee, 2005). Similarly, the software engineering community has suggested a need for a wider availability of empirical methods to improve validity and generalizability of their designs (Basili, 1996, Kontio et al., 2004). Several software engineers have also suggested their use as an evaluation and knowledge elicitation technique (Kontio et al., 2004, LeRouge and Niederman, 2006, Massey and Wallace, 1991, Nielsen, 1997). In the IT industry, focus groups are widely used in human–computer interface usability studies.[1]

We contend that there are several key reasons focus groups are an appropriate evaluation technique for design science research projects (based on Stewart et al. (2007), p. 42):

Flexibility: Focus groups allow for an open format and are flexible enough to handle a wide range of design topics and domains.

[1]For example, usability.gov is a U.S. government web site managed by the U.S. Department of Health & Human Services that outlines the use of focus groups in the design of web pages (see http://www.usability.gov/methods/focusgroup.html).

Direct Interaction with Respondents: This allows for the researcher to clarify any questions about the design artifact as well as probing the respondents on certain key design issues.

Large Amounts of Rich Data: The rich data allow deeper understandings, not only on the respondents' reaction and use of the artifact but also on other issues that may be present in a business environment that would impact the design.

Building on Other Respondent's Comments: The group setting allows for the emergence of ideas or opinions that are not usually uncovered in individual interviews. Additionally, causes of disagreement can point to possible problem areas with the proposed artifact.

10.3 Adapting Focus Groups to Design Research

The traditional literature outlines several steps for the conduct and analysis of focus groups. Obviously, given the breadth of usage and contexts for focus groups, each of these steps can be very different depending on the intent of the research. Figure 10.1 summarizes the basic steps that would be applicable for any research-oriented use of focus groups as found in Krueger et al. (2000), Bloor et al. (2001), Stewart et al. (2007), and Morgan (1988). We analyze each step taking into consideration the focal point of this chapter, the use of focus groups for refinement, and evaluation of a design science artifact.

10.3.1 Formulate Research Question or Problem

In order to effectively define and design the focus groups, the research goals must be clearly identified. In design science research if we seek to design an artifact, incrementally improve the design, and evaluate its utility, we are addressing two complementary, yet different research goals. We propose the use of two types of focus groups to achieve these different research goals: (1) *exploratory focus groups* (EFGs) to achieve incremental improvements in artifact design and (2) c*onfirmatory focus groups* (CFGs) to demonstrate the utility of the design in a field setting. In Fig. 10.2, we illustrate the positioning of the two types of focus groups in the design science research process. As discussed more fully in Hevner (2007), two forms of artifact evaluation are performed in a design research project – the evaluation of the artifact to refine its design in the design science build/evaluate cycle and the field testing of the released artifact in the application environment. We discuss the similarities and differences between EFGs and CFGs in the following focus group steps.

Exploratory focus groups have two roles: (1) the provision of feedback to be utilized for design changes to both the artifact and the focus group script and (2) the refinement of scripts and the identification of the constructs to be utilized in future focus groups. Feedback for improvement of the design of the artifact (Hevner

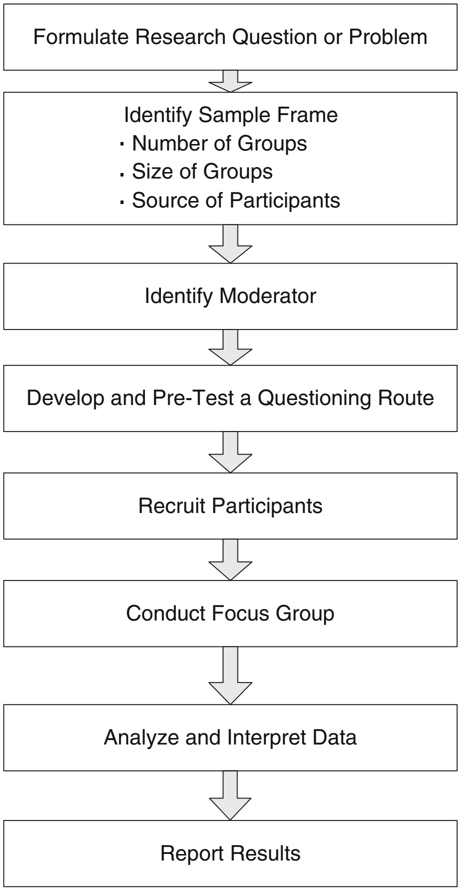

Fig. 10.1 Focus group steps

et al., 2004, Hevner, 2007, Markus et al., 2002) is an essential component of design research. Additionally, the questioning scripts can be refined to improve the quality of feedback received in subsequent EFGs. Finally, EFGs can be used to define and consequently refine the coding scheme that will be used for the analysis and interpretation of field testing in CFGs. The number of EFGs run depends on the number of build/evaluate cycles that use focus groups for evaluation. It is important to note that other evaluation methods (e.g., analytic optimization) may be used for early design cycles while focus groups may be used for later cycles of design refinement.

The CFGs are used to demonstrate the utility of the artifact design in the application field. When using focus groups for rigorous research, the unit of analysis will be the focus group and not the individual participants. Thus, it is crucial not to introduce any changes to the interview script and the artifact when multiple CFGs are conducted. This allows for the comparison of the results across CFGs to demonstrate and corroborate proof of utility of the artifact. The number of CFGs run in the

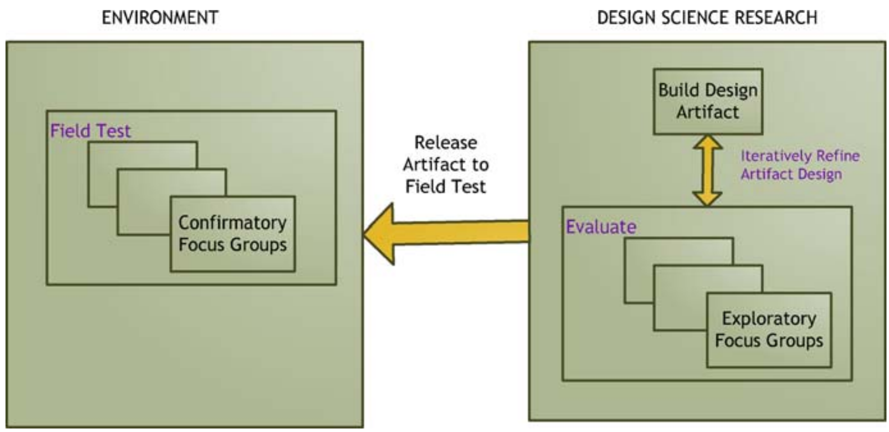

Fig. 10.2 Focus groups in design research

field test depends on the consistency of results across the focus groups and the level of rigor required in the design research project.

10.3.2 Identify Sample Frame

Three decisions are made in this step: (1) number of each type of focus group to run, (2) the desired number of participants in each group, and (3) where to recruit the participants.

10.3.3 Number of Focus Groups

Deciding how many focus groups to run can prove to be quite challenging. Unlike experimentation, there is no power test for the correct sample size. The literature states that focus groups should continue until nothing new is learned (Krueger and Casey, 2000), yet deciding "nothing new" is being learned is a difficult and somewhat arbitrary task. This is especially challenging in design science research. There is always room for improvement of an artifact and certainly a fair amount of subjectivity in interpreting when the design of an artifact is indeed complete. There is certainly a point where we may choose to satisfice in order to move forward. Additionally, there is a need to balance available people and resources, since focus groups can be expensive (most participants receive some sort of compensation) and expert participants may be difficult to find.

In our experience, at the minimum, one pilot focus group, two EFGs, and least two CFGs should be run. This allows for at least two design cycles and enough contrast for field test analysis. Since the unit of analysis is the focus group, it would be difficult to make a compelling argument for the utility of the designed artifact with a smaller number of CFGs.

10.3.4 Number of Participants

Selecting group size has several considerations. It may seem simpler (and less expensive) to run fewer, larger focus groups, since it takes less focus groups to hear from the same number of participants. Yet this could lower "sample size," since there are less groups to compare. Additionally, the dynamics of smaller versus larger groups are different; smaller groups require greater participation from each member and larger groups can lead to "social loafing" (Morgan, 1988). Morgan (1998) suggests a lower boundary of 4 participants and an upper boundary of 12 participants. Depending on the approach taken to demonstrate the artifact to the group, large focus groups (more than six) could be tricky in design research since the subject matter is more complex than traditional focus group topics, for example, a marketing campaign.

10.3.5 Participant Recruitment

The identification of focus group participants is not as statistically rigorous as it would be for survey research. Focus group participants are not randomly selected, but rather are selected based on their characteristics in relation to the topic that is being discussed. In fact, research shows that bringing together groups which are too diverse in relationship to the topic of interests could result in data of insufficient depth (Bloor et al., 2001). For design research the participants should be from a population familiar with the application environment for which the artifact is designed so they can adequately inform the refinement and evaluation of the artifact.

Research is mixed on whether to use pre-existing groups, though for design topics this may be advantageous since the participants have problem solved together and the focus group may approximate a realistic environment (Kitzinger, 1994). Interaction among participants is one of the most important aspects of focus groups. For example, a group of all technical experts may be very different than an expert/non-expert group (Stewart et al., 2007). A design science researcher must consider membership of the focus groups and how it aligns with the research objective early in the participant selection process.

Design researchers should strive to recruit participants that are familiar with the application environment and would be potential users of the proposed artifact. Unfortunately, in many cases such individuals are not easy to find, so plenty of time and effort should be allotted for this task. For instance, it might be possible to conduct the focus group in the evening (most participants will likely work) and offer dinner. Another good approach is to conduct the focus group at a place where the potential participants work, again enticing them with lunch or breakfast. Phone calls and e-mails should be placed at least a month before the focus groups are planned. A few days before the focus groups the participants should be reminded. Researchers should plan for a few participants to not show up, so if the goal is six people, invite eight.

10.3.6 Identify Moderator

Due to the open-ended nature of focus groups, moderation can be complex, especially in social research. Several skills are important when moderating a focus group. Krueger et al. (2000) found the following skills to be highly important: (1) respect for participants, (2) the ability to communicate clearly, both orally and in writing, (3) the ability to listen and the self-discipline to control personal views, and (4) a friendly manner and a sense of humor. For design research, the moderator not only needs to have these skills but also a clear understanding of the technical aspects of the design artifact. In many cases the moderator may be one of the artifact designers. In this case, the moderator has to be very careful not to introduce any personal bias in the presentation of the artifact (we tend to be proud of our work), particularly when conducting an EFG. It may be possible to enlist a second observer to guard against the encroachment of personal views (at least during the initial groups). This is an excellent time to receive good suggestions for improvement of the design and the designer has to be receptive to criticism and suggestions given by the participants, being careful to not justify or defend his work.

10.3.7 Develop and Pre-test a Questioning Route

The questioning route is the agenda for the focus group. In the questioning route you are setting the direction for a group discussion (Stewart et al., 2007) and it should closely align with your research objectives. There should be no more than 12 questions for a 2 h session (Krueger and Casey, 2000, Stewart et al., 2007). Two general principles outlined by Stewart et al. (2007, pg. 61) are to order the questions from the most general to the more specific and to order the topics by the relative importance to the research agenda. Thus, the topics to be discussed are ordered by importance, and within those topics, the questions are ordered from general to specific.

For a designed artifact, this means beginning with a broad explanation of scenarios where the artifact could be utilized, followed by a description of the artifact and how it is to be utilized and finishing with a scenario where focus group participants have the ability to utilize and evaluate the artifact.

For an EFG, the “rolling interview guide” (Stewart et al., 2007) is an excellent approach. With a rolling interview guide, a script is created for the first EFG but is changed for the next EFG, based on the outcome of the previous EFG. One of the advantages of this approach is that it allows for information to unfold over time as you discover more about how people would understand and use the artifact. However, it is imperative that no revisions are made to the interview guide in the CFGs, since continuous change would make comparisons across the focus groups difficult, compromising rigorous interpretation of the results (Stewart et al., 2007).

A promising evaluation approach in design research focus groups (both EFGs and CFGs) is to create a manipulation within the focus group. Participants can be asked to collectively complete a task without the artifact and then again with the artifact. The ensuing discussion should revolve around how the artifact was used and how the completion of the task was altered by its use.

10.3.8 Conduct the Focus Group

Focus group sessions should be fun and stimulating for the participants and moderator (Stewart et al., 2007). The moderator usually greets the participants as they enter and may ask them to fill out demographic information and informed consent forms (e.g., IRB forms). The participants are generally seated in a U-shape arrangement to encourage collaboration (Krueger and Casey, 2000) and allow space for the moderator to demonstrate the artifact. Seating arrangements are also very important. A good approach is to get to know the participants before the questioning route begins, as you greet them when they arrive. The most assertive and expert participant should be seated next to the moderator and the least talkative directly across from the moderator (Krueger and Casey, 2000, Stewart et al., 2007).

Depending on your research protocols, focus groups may be video and/or audio taped. Generally, the participants are told they are being recorded and most institutional review boards require written consent. It is also a good idea to have an observer. The observer will not participate in the focus group, rather will take careful notes, noting in particular any strong reactions, the participants' facial expression, and general tone of any exchange between participants or between the participant and the moderator (Stewart et al., 2007).

Time management is also important when conducting a focus group. A moderator should be able to recognize when all possible issues for a topic have been covered and move on to the next topic. Pilot focus groups can help anticipate and manage the length of focus groups.

Additional guidelines for running focus groups can be found in many excellent texts, such as Krueger and Casey (2000), Stewart et al. (2007), Bloor et al. (2001), and Morgan (1988).

10.3.9 Analyze and Interpret Data

The two design research goals for using focus groups are the incremental improvement of the design of the artifact and the demonstration of the utility of the design. For this reason, we have suggested the different focus group types of EFG and

CFG. While the objectives of the two group types are very different, the methods of analyzing the focus group data from both EFG and CFG can be similar. The interpretation of the focus group discussions has many of the same challenges in demonstrating rigor that all qualitative research encounters share. Several techniques that are used for qualitative data analysis can be considered, carefully selecting those techniques that emphasize the reliability and replicability of the observations and results (Stewart et al., 2007).

One possible approach is template analysis. Unlike a grounded theory approach (Desanctis and Gallupe, 1987), template analysis normally starts with at least a few pre-defined codes which help guide analysis. The first step in template analysis is to create an initial template by exploring the focus group transcripts, academic literature, the researchers' own experiences, anecdotal and informal evidence, and other exploratory research (King, 1998). The contents of the discussions are also examined for the meanings and implications for the research questions. Individual constructs should also be investigated, looking for common themes and variations within the constructs that would provide rich descriptions of the participants' reactions to design features.

In template analysis, the initial template is applied in order to analyze the text but is revised between each EFG session. Once the final template is created after the final EFG, it is used to code the CFG sessions.

10.3.10 Report Results

King (1998) suggests that qualitative results can be reported by creating an account structured around the main themes identified, drawing illustrative examples from each transcript as required. A similar approach can be taken when reporting focus group results. Short quotes are used to aid in the specific points of interpretation and longer passages of quotation are used to give a flavor of the original discussions. Summary tables can be very helpful, displaying both evidence and counter-evidence of the utility of the artifact by focus group. Rich descriptions can further corroborate results by using quotes from the focus group participants.

10.4 A Design Research Example

To illustrate the use of focus groups in a design research project, we discuss a recently completed research project in which an artifact was designed and evaluated in the health-care context. The research investigated issues of data quality in the context of public policy health planning. Three data quality problems are identified and a set of quality metrics are designed to support improved decision making. These metrics aid human decision makers to better understand the quality of the data they have and how to overcome inherent decision-making biases in the presence of potentially incomplete and unreliable information from multiple sources.

10.4.1 Research Context

Like other business organizations, the health-care sector is increasingly becoming an information-driven service (Al-Shorbaji, 2001, Derose and Petitti, 2003, Derose et al., 2002, Friede et al., 1995), particularly for public policy and health planning. In fact, information systems are becoming an integral part of public health decision making. Information acquisition can now be transacted rapidly (Chapman and Elstein, 2000, Maibach and Holtgrave, 1995, U.S., 1995) and from several sources. To improve public health's efficacy and profile, both practitioners and researchers need reliable and timely information to make information-driven or evidence-based decisions (Friede et al., 1995). This study focused on this rich health planning domain, in particular on a set of specific decision-making activities related to community needs assessment.

Figure 10.3 illustrates the research process used to identify potential data quality measures and biases. A field study was conducted (Tremblay et al., 2007) and combined with a review of the literature. Several data quality issues and biases were selected as the focus of the research.

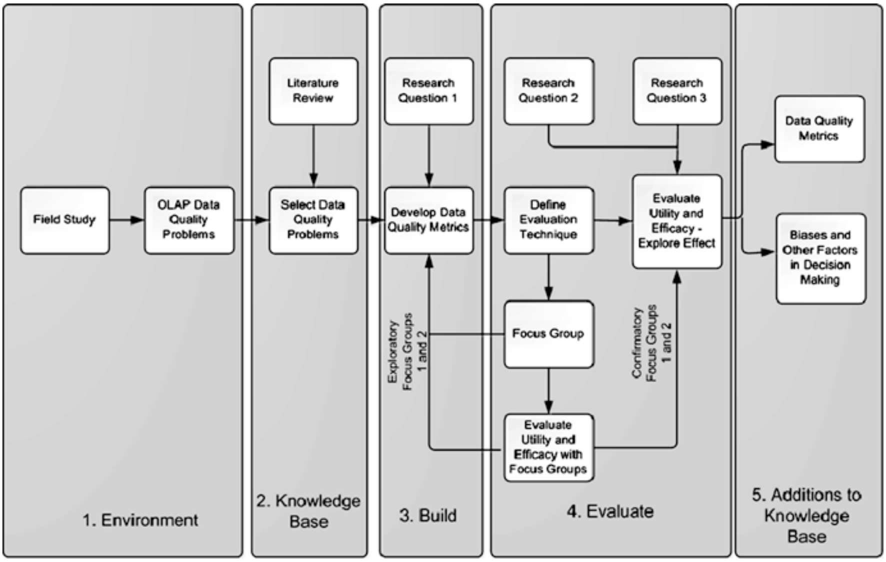

Fig. 10.3 Design research process

10.4.2 Data Quality Metrics Description

Information supply chains (ISC) (Ballou et al., 1998, Ballou and Pazer, 1985, Shankaranarayan et al., 2003) can be complex, multi-step processes that include the collection of raw data from many sources, comprised of intermediate transformations, compositions, and standardizations that ultimately supply the raw data for insightful analysis.

As shown in Fig. 10.4, data quality can be assessed as part of the original data collection process and propagated through transformations and compositions made by the ISC as part of lineage-driven data quality measurement. In contrast, result-driven data quality proceeds from the information product endpoint, with knowledge of the context, and works backward to provide measures that assist decision makers in understanding the uncertainties that account for possible poor decision making due to well-known judgment biases. Result-driven data quality is especially important in an environment where managers and decision makers utilize aggregated data (summary information) retrieved from several data sources in the information supply chain to make tactical decisions.

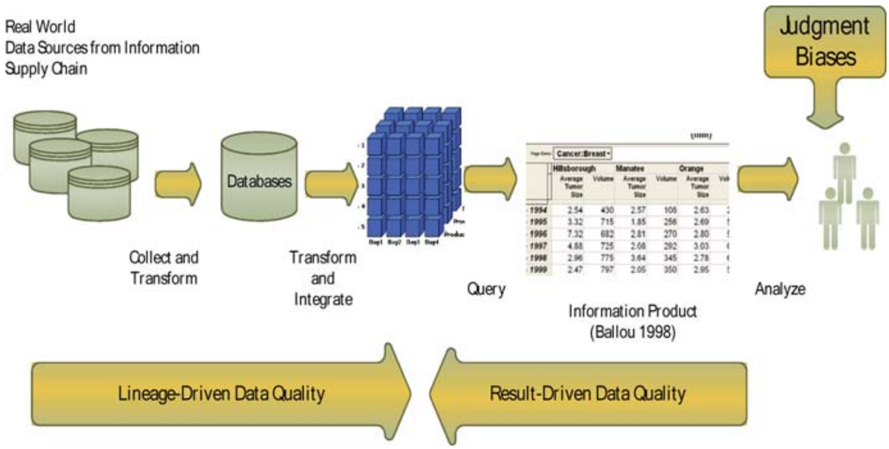

Fig. 10.4 Research landscape

This is true in health care, and in particular in health planning, where health-care resource allocation is often based on summarized data from a myriad of sources such as hospital admissions, vital statistic records, and specific disease registries. These data are utilized to justify investments in services, reduce inequities in treatment, and rank health-care problems to support policy formulation (Berndt et al., 2003).

This project presented methodologies that communicated result-driven data quality (RDQ) information at decision time with simple and comprehensible metrics that can be calculated when the final information product (IP) is created. The decision maker is not involved in the calculation of the metric but considers the metrics as they formulate a context-specific decision. We consider how to present information on the three data quality dimensions for any unique information product in an online analytical processing (OLAP) environment. This project proposed three data quality measures and associated data quality metrics (DQMs) which are summarized in Table 10.1.

Table 10.1 Data Quality Metrics

Data quality problem (Wang and Strong, 1996)	Metric
Completeness: A problem is encountered when combining or aggregating data from multiple sources in the ISC that is missing codes or has codes that do not match other sources of data. This results in data that are not assigned to any of the possible cells in a data cube	*Unallocated data metric* which considers the effects of null values in any of the grouping or filtering variables for *counts* and for *averages*. It proposes a case-based approach for presenting unallocated data to a decision maker, which gives flexibility for the decision maker to consider different "what if" scenarios
Representational consistency: When considering aggregated data or when observing trends, decision makers rely on point estimates, such as an average, which may be biased by noisy data	*Information volatility metric* is a measure of reliability proposed as an addition to OLAP tools when considering aggregated data or when observing trends. Two types of information volatility are defined: intra-cell and inter-cell
Appropriate amount of data: Insensitivity to sample size by decision makers when considering/comparing groupings	*Sample size indicator* is a simple method of drawing the attention of the decision maker in order to mitigate a well-known bias

The data quality metrics are designed and implemented in order to present these metrics in an effective way to decision makers. We considered several alternative evaluation methods and selected the focus group technique as the most appropriate for the research context. The research process in Fig. 10.3 shows the use of both EFGs and CFGs.

10.4.3 Design Research Questions

In order to correctly design the focus group scripts and identify qualified participants the research questions are clearly identified. The research issues for the EFGs are how to improve the design artifact and how to develop a rigorous and comprehensive focus group script and coding schema.

For the CFGs two research questions are formulated:

What are the utilities of the data quality metrics in a realistic field context?

What are the efficacies of the data quality metrics to alter a decision maker's data analytic strategies by eliminating inherent human bias via better understanding of the data?

10.4.4 Identify Sample Frame

A total of five focus groups of 6–12 participants are executed. The first focus group was a pilot to help identify timing issues, refine the questioning route, evaluate the moderator's style, and surface any potential logistical issues. The pilot data were not used further for data analysis.

The following requirements are outlined for the participants of EFGs and CFGs: previous experience with decision making in the health-care field, an advanced college degree, and some training in statistics and decision-making software systems.

10.4.5 Identify Moderator

The moderator is the primary researcher who had some experience in moderating focus groups in both educational and industrial settings. Another researcher serves as an observer to take careful notes and to support the moderator in time keeping.

10.4.6 Develop a Questioning Route

The planning process includes creating a carefully planned script in which all three of the designed metrics are presented to the participants (see Appendix A for a partial script). The research utilizes the "rolling interview guide" (Stewart et al., 2007) for the EFGs. A script is created for the first focus group. Then, based on the outcomes of the first EFG, the guide is revised for use in the second EFG. Based on the outcome of the second EFG, the script, the coding template, and the metrics are revised again. No revisions are made during the execution of the CFGs.

"Vignettes" or story lines are used to create fictitious decision scenarios based on current health-care situations (in recent news reports) and sample health-care data. These data include data from a statewide cancer registry, which has been collecting incidence data since 1981, county data from the US Census Bureau, demographic data from commercial sources, and an internally generated time dimension. The strategy is to present the data with and without the metrics information in order to detect differences in the collective decision-making processes. Thus, we develop an experimental manipulation within the context of a focus group. A PowerPoint presentation is used to describe the vignettes and the metrics. The moderator presents the health-care decision-making context.

Table 10.2 shows examples of the vignettes used in the focus groups

10.4.7 Recruit Participants

Potential participants are identified via personal contacts and phone calls to county public health departments. The potential focus group members are given a brief description of the study and their participation is requested. They are offered dinner after the focus group session. Many of the participants had taken university courses in data warehousing and/or data mining. Several other participants had jobs that

Table 10.2 Example of vignettes

Metric evaluated	Vignette	Decision
Unallocated data metric	Studies have shown that smoking is responsible for most cancers of the larynx, oral cavity and pharynx, esophagus, and bladder. In addition, it is a cause of kidney, pancreatic, cervical, and stomach cancers, as well as acute myeloid leukemia	Is there correlation between smoking and certain types of cancer?
Unallocated data metric	When Hispanics are diagnosed with a certain cancer (fictitious example), they are less likely to receive chemotherapy than non-Hispanics	Is there disparity in care between ethnic groups?
Information volatility metric	Counties neighboring the target county are better at early detection/prevention of breast cancer based on volumes of cases	Examine trend – is this a true claim?
Sample size indicator	Tumor size has been shown to be a good predictor of survival for certain cancers, including breast, lung, and endocrine. Compare average tumor size in the target county to that of neighboring counties	How does the target county compare to other counties?

required use of data analytics (e.g., spreadsheets, business intelligence tools, statistics packages). To illustrate the qualifications of participants in one of the focus groups, Table 10.3 shows the demographic characteristics from one of the focus groups.

10.4.8 Conduct Focus Groups

The focus groups are held in state-of-the-art conference rooms. The participants are seated in a U-shape arrangement to encourage collaboration (Krueger and Casey, 2000) and to allow space for the moderator to demonstrate the design artifacts and PowerPoint presentation. The moderator presents the experimental vignettes and encourages the participants to play the role of a health-care decision maker. In order to analyze the data, the focus group guides the moderator in exploring the health-care data. For example, participants are encouraged to ask the moderator to drill down or roll up data in order to thoroughly understand and compare data for different counties as part of their decision-making process.

The participants are asked to come to consensus on a particular task without the data quality metric. They are then asked to reconsider the decision utilizing the data quality metric. The ensuing discussion revolves around how the data quality metric is used and how the metric affected their decision making. The sessions are recorded and professionally transcribed.

Table 10.3 Sample focus group participants

Gender	Age	Last degree	Current position	Course in statistics?	Years of work experience	Years of health-care experience	Self-reported comfort with data analysis (7-point scale)
M	34	PhD	Health economist	Y	7	6	7
M	51	PhD	Assistant director of measurement and evaluation	Y	28	28	7
F	49	PhD	Researcher	Y	28	28	5
F	35	PhD	Project manager/data analyst/health science specialist	Y	9	9	5
M	56	PhD	Health services researcher	Y	25	20	6
F	31	MA/MPH	Program specialist	Y	8	7	5
F	31	MSPH	Project manager	Y	8	6	7
F	36	PhD	Health economist	Y	NR	3	7

After conducting each of the EFGs, significant changes are made to both the design artifacts (the data quality metrics) and to the focus group scripts and coding templates. The observer helps refine the focus group script used in the EFG. He carefully observes people's understanding of the scenarios, their reaction to the metrics and the flow of the conversation and takes notes. The notes are carefully analyzed and changes are made to the focus group script for the next EFG. For example, the observer noted that the moderator needed to better clarify the goal of research, in particular he needed to give a clearer description of who normally would utilize these types of tools and for what sorts of tasks. Once the CFG begins, no changes are made to the questioning route.

10.4.9 Analyze and Interpret the Data

10.4.9.1 Template Analysis

Template analysis is selected for the interpretation of the focus group discussions. The initial template has a few pre-defined codes which focus on aspects of the data quality metrics. The contents of the focus group discussions are examined for their meanings and their particular implications for the research questions, in our case, changes in data analytic strategies and evidence or counter-evidence of the metrics' usefulness. Individual constructs are investigated, looking for common themes and variations within the constructs that would provide rich descriptions of the participants' reactions to design features and attitudes to decision making with varying levels of data quality as defined by the designed metrics. In addition, several other coding categories are created during coding to explore the entire range of participants' reactions (see Table 10.4 for a partial coding template for the information volatility metric).

Table 10.4 Partial final coding scheme

Construct	Definition
Volatility before	Strategies to deal with volatility prior to receiving metric Interpretation before
Volatility after	Strategies to deal with volatility after receiving metric Interpretation after
Design feature volatility	Mention of the information volatility feature, design improvement suggestion
Speculation	Speculation on data quality problems
Other factors in decision making	Including stakeholder issues

Once the template is completed and agreed upon by the researchers, the transcripts for the first EFG are coded by identifying sections that are relevant and annotating the appropriate codes from the initial template. Cohen's kappa is used to measure inter-rater reliability (Cohen, 1960). The results are then reconciled between coders. The two independent coders discuss the areas of disagreement, stopping when agreement is reached on all higher ordered codes and most of the lower order codes (King 1998). The transcripts are then recoded based on the reconciliation between the two coders.

10.4.10 Report Results

The identified constructs of utility and efficacy are investigated. Utility is defined as "usefulness of the metric" and efficacy as "having the ability to change data analytic strategies." To analyze utility of the metric all passages that are coded as "design feature" are analyzed. Changes in data analytic strategies are evaluated contrasting the passages coded as "before" and "after" for each metric.

For each of the metrics proposed in the study both evidence and counter-evidence of the utility and efficacy of the metrics are presented. The qualitative data are summarized for both utility and efficacy, and then rich descriptions are given using quotes from the focus group participants to corroborate the results.

Table 10.5 is an example of the summary for one of the data quality metrics, information volatility. The results are summarized by focus groups. For this particular metric, the design is improved by adding benchmarking information based on the results from one of the exploratory focus groups. The benchmarking idea came from a participant in the second EFG:

> You (need to) draw a line in the sand and say, this is a problem, this is not. And maybe if it goes over that line, it pops up and says, 'Hey, check this out.'

Table 10.5 Utility of information volatility metric

Focus group	Evidence of utility	Counter-evidence of utility
EFG1	Yes	Difficulty interpreting
EFG2	Yes	Difficulty interpreting
CFG1	Yes – saw several instances where this would be useful in their daily data analysis	None
CFG2	Yes	None

This was corroborated by a remark from a participant on one of the CFGs:

> ...benchmarking is a necessary component of it.

Similarly, Table 10.6 shows how the efficacy of the same metric is evaluated. This particular example points out one of the limitations of the use of focus groups. We ran two CFGs to field test the set of design metrics. For the volatility metric, no efficacy data were collected from one of the CFGs. This group was dominated by an individual who rejected the validity of the vignette presented to evaluate the metric. The individual convinced the group to refuse to make a decision. Thus to study this metric, at least one more focus group needs to be run to show stronger evidence.

Table 10.6 Efficacy of information volatility metric

Focus group	Change in data analytic strategies?	Comments/observed changes
EFG1	Yes	
EFG2	Yes	
CFG1	N/A	Rejected task, group disliked low realism of the vignettes, refused to make decision
CFG2	Yes	

10.5 Limitations on the Use of Focus Groups for Design Research

We observed several limitations in the use of focus groups in the evaluation of artifacts. Generalization to a larger population can be difficult for several reasons. The first is due to the convenience nature of focus group recruiting practices. It is particularly difficult to find adequate participants when evaluating artifacts due to the technical nature of the subject which limits the pool of possible participants. Additionally, individual responses cannot be considered because of the interaction between respondents and between respondents and the moderator. A strongly opinionated member may bias the results and discourage other participants from speaking, as we saw in the above example.

Another limitation lies in the difficulty of deciding how many focus groups to run. Unlike experimentation there is no power test for the correct sample size. The focus groups literature states that focus groups should continue until nothing new is learned (Krueger and Casey, 2000), yet deciding whether “nothing new” is being learned is a difficult and somewhat arbitrary task. When considering the design of the artifact, the EFG continuously produced new ideas and suggestions, making the decision to stop and move on to CFG somewhat subjective. Deciding how many CFGs to run is also difficult. In our study the two CFGs found some differing results, thereby highlighting the need for additional CFGs. The choice will most likely be driven by the costs of running focus groups and difficulty in finding additional expert participants.

A very important aspect of conducting focus groups is an effective moderator who is skilled in drawing information from the participants, encourages interaction between participants, and is non-authoritarian and non-judgmental (Stewart et al., 2007). The moderator has to be careful to not bias the results during the focus groups. In our example, the moderator had control of the interface in which the data and metrics were presented to the groups, which certainly led to different results than if the focus group participants had been able to access them directly at their own workstations. However, the goal was to focus the attention of all the participants at the same point to enable common discussion.

10.6 Closing Remarks

The goal of this chapter is to propose focus groups as a useful method for two of the fundamental goals of design science research: refinement of a proposed artifact and demonstration of its utility. We outline how traditional focus group methods can be adapted for these purposes. For the evaluation of an artifact design, *exploratory focus groups* (EFGs) study the artifact to propose improvements in the design, continuing the cycle of build and evaluate until the artifact is released for field test in the application environment. Then, the field test of the design artifact may employ *confirmatory focus groups* (CFGs) to establish the utility of the artifact in field use.

Rigorously designing, planning, selecting participants, conducting, analyzing, and reporting the results of the focus groups have unique concerns as adapted to design science research and we outline several potential approaches for each step in the focus group process. Additionally, the data generated by this methodology are qualitative and we describe a process to capture, code, analyze, and report these data in a rigorous manner.

As we conducted the focus groups, we were encouraged by the emergence of rich ideas and concepts that emerged when using this qualitative technique. The intent of this example research study was the evaluation of the proposed data quality metrics, but several other "user views" of data quality emerged that merited serious consideration and will stimulate further research. The open-ended nature of focus groups allowed for the identification of new artifact ideas, several of which we are currently pursuing.

The focus group technique allowed the researchers to observe data quality in action in actual decision making. One interesting finding, for example, was that though participants were skeptical of the data in the examples (which for the most part was from a real ISC!), they were not skeptical about their own data (data that they utilized in their jobs), perhaps because they have very high ownership of those data and believe their data to be of high quality, even though this is rarely true. We observed that several "irrational" approaches were taken to analyze the data. These included speculating on the reasons for poor data quality without any real evidence (e.g., Hispanics do not go to the doctor as much as other ethnic groups.).

Certainly there are many other avenues to explore in the use of this technique. In every step we outline, there may be diverse approaches that are contingent on the artifact and the application domain. For example, the design of the focus group script will be very different for varying application domains. In our case, we dealt with a decision-making environment in the health-care industry, but another approach may be needed if the context is significantly different (for example, a supply chain bid recommendation agent). In fact, this technique is most appropriate where observational methods can be used for evaluation. Also, template analysis worked well for our study but we undoubtedly can draw from the work of other qualitative researchers for guidance on other ways to analyze transcribed focus groups.

To conclude, we believe that focus groups are a highly relevant and rigorous approach for improving and evaluating design artifacts. However, it is critical that researchers adapt traditional focus group methods to the goals of design science research projects in the forms of exploratory focus groups and confirmatory focus groups. The contributions of this chapter are the explication of how design science focus groups can be performed in order to achieve these research goals and the presentation of an exemplar design science research project that effectively used focus groups for both exploratory and confirmatory evaluations.

References

Al-Shorbaji, N. (2001) Health and Medical Informatics: Technical Chapter, in *Health Information Support, Regional Office for the Eastern Mediterranean, World Health Organization,* Cairo, Egypt.

Baker, T. and D. A. Collier (2005) The Economic Payout Model for Service Guarantees, *Decision Sciences* 36 (2), p. 197.

Ballou, D., R. Wang, H. Pazer, and G. K. Tayi (1998) Modeling information manufacturing systems to determine information product quality, *Management Science* 44 (4), pp. 462–484.

Ballou, D. P. and H. L. Pazer (1985) Modeling data and process quality in multi-input, multi-output information systems, *Management Science* 31 (2), pp. 150–162.

Basili, V. R. (1996) The role of experimentation in software engineering: past, current, and future, in *Proceedings of the 18th International Conference on Software Engineering,* Berlin, Germany, IEEE Computer Society.

Baskerville, R. and M. D. Myers (2004) Special issue on action research in information systems: making is research relevant to practice-foreword, *MIS Quarterly* 28 (3), p. 329.

Benbasat, I. and R. Weber (1996) Research commentary: rethinking "diversity" in information systems research, *Information Systems Research* 7 (4), p. 389.

Berndt, D. J., A. R. Hevner, and J. Studnicki (2003) The CATCH data warehouse: support for community health care decision-making, *Decision Support Systems* 35 (3), p. 367.

Bloor, M., J. Frankland, M. Thomas, and K. Robson (2001) *Focus Groups in Social Research*, Sage, London.

Chapman, G. B. and A. S. Elstein (2000) Cognitive Processes and Biases in Medical Decision Making, in G. B. Chapman and F. A. Sonnenberg (eds.) *Decision Making in Healthcare: Theory, Psychology and Applications*, Cambridge University Press, Cambridge.

Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20 (1), pp. 37–46.

Cole, R., S. Purao, M. Rossi, and M. k. Sein. (2005) Being proactive: where action research meets design research, in *Twenty-Sixth International Conference on Information Systems,* Las Vegas, 2005.

Debreceny, R., M. Putterill, L.-L. Tung, and A. L. Gilbert (2003) New tools for the determination of e-commerce inhibitors, *Decision Support Systems* 34 (2), p. 177.

Derose, S. F. and D. B. Petitti (2003) Measuring quality of care and performance from a population health care perspective, *Annual Review of Public Health* 24 (1), pp. 363–384.

Derose, S. F., M. A. Schuster, J. E. Fielding, and S. M. Asch (2002) Public health quality measurement: concepts and challenges, *Annual Review of Public Health* 23 (1), pp. 1–21.

Desanctis, G. and R. B. Gallupe (1987) A foundation for the study of group decision support systems, *Management Science* 33 (4), pp. 589–609.

Friede, A., H. L. Blum, and M. McDonald (1995) Public health informatics: how information-age technology can strengthen public health, *Annual Review of Public Health* 16 (1), pp. 239–252.

Galliers, R. D. (1991) Choosing appropriate information systems research approaches: a revised taxonomy, in H. E. Nissen, H. K. Klein, and R. Hirschheim (eds.) *Information Systems Research: Contemporary Approaches and Emergent Traditions*, Elsevier Science Pub., Amsterdam, The Netherlands.

Hevner, A., S. March, J. Park, and S. Ram (2004) Design science research in information systems, *Management Information Systems Quarterly* 28 (1), pp. 75–105.
Hevner, A. R. (2007) A three cycle view of design science research, *Scandinavian Journal of Information Systems* 19 (2), pp. 87–92.
Iversen, J. H., L. Mathiassen, and P. A. Nielsen (2004) Managing risk in software process improvement: an action research approach, *MIS Quarterly* 28 (3), p. 395.
Jarvenpaa, S. L. and K. R. Lang (2005) Managing the paradoxes of mobile technology, *Information Systems Management* 22 (4), p. 7.
King, N. (1998) Template Analysis, in G. Symon and C. Cassell (eds.) *Qualitative Methods and Analysis in Organizational Research*, Sage Publications, London.
Kitzinger, J. (1994) The methodology of Focus groups: the importance of interaction between research participants. *Sociology of Health* 16 (1), pp. 103–121.
Kontio, J., L. Lehtola, and J. Bragge (2004) Using the focus group method in software engineering: obtaining practitioner and user experiences, *International Symposium on Empirical Software Engineering Proceedings of the 2004 International Symposium on Empirical Software Engineering*, IEEE Computer Society, Washington, DC, pp. 271–280.
Krueger, R. A. and M. A. Casey (2000) *Focus Groups: A Practical Guide for Applied Research*, 3rd edn, Sage Publications, Thousand Oaks, CA.
LeRouge, C. and F. Niederman (2006) Information systems and health care Xi: public health knowledge management architecture design: a case study, *Communications of the Association for Information Systems* 18, p. 15.
Lindgren, R., O. Henfridsson, and U. Schultze (2004) Design principles for competence management systems: a synthesis of an action research study, *MIS Quarterly* 28 (3), p. 435.
Maibach, E. and D. R. Holtgrave (1995) Advances in public health communication, *Annual Review of Public Health* 16 (1), pp. 219–238.
Manning, P. K. (1996) Information technology in the police context: the "sailor" phone, *Information Systems Research* 7 (1), p. 52.
Mantei, M. M. and T. J. Teorey (1989) Incorporating behavioral techniques into the systems development, *MIS Quarterly* 13 (3), p. 257.
Markus, M. L., A. Majchrzak, and L. Gasser (2002) A design theory for systems that support emergent knowledge processes, *MIS Quarterly* 26 (3), pp. 179–212.
Massey, A. P. and W. A. Wallace (1991) Focus groups as a knowledge elicitation technique: an exploratory study, *IEEE Transactions on Knowledge and Data Engineering* 3 (2), pp. 193–200.
Merton, R. K. and P. L. Kendall (1946) The focused interview, *The American Journal of Sociology* 51 (6), pp. 541–557.
Morgan, D. L. (1988) *Focus Groups as Qualitative Research*, Sage Publications, Newbury Park, CA.
Nielsen, J. (1997) The use and misuse of focus groups, *IEEE Software* 14 (1), pp. 94–95.
Shankaranarayan, G., M. Ziad, and R. Y. Wang (2003) Managing data quality in dynamic decision environments: an information product approach, *Journal of Database Management* 14 (4), pp. 14–32.
Smith, H. J., S. J. Milberg, and S. J. Burke (1996) Information privacy: measuring individuals' concerns about organizational practices, *MIS Quarterly* 20 (2), p. 167.
Stewart, D. W., P. N. Shamdasani, and D. W. Rook (2007) *Focus Groups: Theory and Practice*, 2nd edn., vol. 20, Sage Publications, Newbury Park, CA.
Torkzadeh, G., J. C.-J. Chang, and G. W. Hansen (2006) Identifying issues in customer relationship management at Merck-Medco, *Decision Support Systems* 42 (2), p. 1116.
Tremblay, M. C., R. Fuller, D. Berndt, and J. Studnicki (2007) Doing more with more information: changing healthcare planning with OLAP tools, *Decision Support Systems* 43 (4), pp. 1305–1320.
U.S. Congress (1995) Bringing Health Care Online: The Role of Information Technologies, Office of Technology Assessment, U.S. Government Printing Office, Washington, DC.
Venable, J. R. (2006) The role of theory and theorizing in design science research, in *DESRIST Conference, Claremont, CA, 2006*.

Wang, R. Y. and D. M. Strong (1996) Beyond accuracy: what data quality means to data consumers, *Journal of Management Information Systems* 12 (4), pp. 5–34.
Wellner, A. S. (2003) The new science of focus groups, *American Demographics* 25 (2), p. 29.
Xia, W. and G. Lee (2005) Complexity of information systems development projects: conceptualization and measurement development, *Journal of Management Information Systems* 22 (1), p. 45.