The validity and reliability of focus groups as a research method in adult education

N. E. CHIONCEL
University of Nijmegen, The Netherlands

R.G.W. VAN DER VEEN
Columbia University, USA

D. WILDEMEERSCH
University of Leuven, Belgium

P. JARVIS
University of Surrey, UK

The article examines the theory and practice of focus groups in adult education research. Three theoretical positions are described: radical hermeneutic, moderate interpretative and pragmatic realistic position trying to bridge the gap with positivist research. This last position has been chosen as the departure point for a further analysis of the focus group elements, which relate to validity and reliability. The research examines four European research projects in the field of adult education. The most important and surprising data are presented with reference to both socio-psychological and technical problems in the use of this method. The results highlight the importance and dynamics of all the elements discussed.

Introduction

The focus groups method seems increasingly popular in adult education research. For instance, three out of 11 research projects within the UK research programme 'The Learning Society: knowledge and skills for employment' (Coffield 2000) used focus groups along with other methods. Our empirical data are based on four European projects that also used this approach.

Why are focus groups so popular as a research method these days? And why are they particularly significant in adult education? Vaughn *et al.* (1996) suggest some general utilities of focus groups; people are valuable sources of information because they are both capable of reporting factual data and interpreting these data, and they are articulate enough to discuss opinions, feelings and percep-

N. E. Chioncel is Junior Researcher in Adult Education, University of Nijmegen, the Netherlands. *R. G. W. van der Veen* is Adjunct Professor of Adult Learning and Leadership, Teachers College, Columbia University, New York. *D. Wildemeersch* is Professor of Social and Intercultural Pedagogy, University of Leuven, Belgium. *P. Jarvis* is Professor of Continuing Education, University of Surrey.

tions. Assuming these to be correct, structured group discussion is a relevant procedure to obtain people's opinions, feelings and perceptions, although the interactions between group members have to be borne in mind. The information obtained can be about what each person feels and thinks, but it can also be influenced by a phenomenon such as 'groupthink', through which people conform to what others believe. In judging the popularity and the effectiveness of focus groups, it is important to assess the advantages of the method as well as its limitations.

In adult education, focus groups perhaps offer a more concrete opportunity for participants and researchers to learn from the process (Field 2000), being a powerful way of engaging with professionals, policy makers and end-users. The benefits of using such an iterative process in which 'people's views and understandings are shared, debated, challenged and changed' (Field 2000: 324) are more important than the potential disadvantages. Morgan (1996) notes that the majority of the published articles about focus groups refers to the use of this method along with other methods, such as individual in-depth interviews or with surveys. In the research process, focus groups can be used at different points: at the preliminary or exploratory stage of a study; a preparation for decisions about the precise objectives and design of the study; for the collection of the main data; in the closing period of a study to interpret findings, or to generate further perspectives for research.

Despite the fact that there are clear instructions and precise requirements on how to do focus groups research, there is a lack of studies on the methodological aspects, more precisely on the problems of validity and reliability. Morgan (1996) argued that there is a need to do more research on the focus group method, so we set out to undertake such a methodological research on four recent European research projects:

- Education and Training for Governance and Active Citizenship in Europe (Etgace);
- Enhancing Young Adults Participation in Social and Economic Processes (Balancing Competencies)
- The Future Goals and Policies of Adult Education in Europe (Eurodelphi);
- Strategies for Adult Education Improvement in Bicultural Regional Settings (Saeda).

We start with a sketch of the theoretical framework we used. Subsequently, we present our data, based on a survey among researchers, in-depth interviews with co-ordinators and a focus group of researchers on the results of the survey and the in-depth interviews.

Definition and typology of focus groups

Literature presents various definitions of focus groups. The first important and influential article of Merton and Kendall (1946) on the focus group interview set the parameters which relate to the specific experiences and opinions of the interviewees about the subject in relation to predetermined research questions. In order to complete these, an explicit interview guide was recommended. Nowadays,

a focus group is defined 'as a research technique that collects data through group interaction on a topic determined by the researcher' Morgan (1996:130). In some cases, the analyses of these interactions become an extra parameter. These interactions between participants allow observing, understanding and analysis of the degree and significance of agreement or disagreement between participants concerning the specific topics. It is obvious that, in this case, the researchers must consider the group and not the individuals as the unit of analysis. Following an inclusive approach, Morgan (1996:132) treats focus groups as a 'set of central tendencies, with many useful variations that can be matched to a diversity of research purposes'. The present framework and the general definition still allow differentiation of two types of focus groups: lay groups and expert panels. Below, we will make further differentiation.

Focus groups as lay groups

Focus groups as lay groups can be used as a data gathering method. Recently, focus groups, as a 'management tool', have been 'largely associated with market research' (Field 2000: 323). But the marketing domain is not the only application of this type of focus group research, despite the fact that in this area it is heavily represented (Greenbaum 2000). There are many applications of focus group method in the field of applied sociology (Gamson 1992, Cable 1992, McKinlay 1993, Bobo *et al.* 1995—all referred to in Morgan 1996 and Johnson 1996). There are also applications in communication studies (Staley 1990, Albrecht *et al.* 1993—both referred to in Morgan 1996), in education (Vaughn *et al.* 1996, Coffield 2000, Field 2000, Walker and Tedick 2001), in psychology (Vaughn *et al.* 1996) and public health (Kidd and Parshall 2000).

The second possible alternative in this category of focus groups as lay groups, is to employ them for 'social marketing', as a tool in the development and evaluation of social programmes (Morgan 1996). In this respect, focus group research can be designed as an instrument for empowering disadvantaged people or, as in emancipatory research, to 'give a voice' to marginalized groups. For instance, many feminist researchers prefer this approach (Wilkinson 1998). The primary purpose of these action-oriented approaches is 'not to develop academic theories; nor is it to produce theoretical or empirical knowledge that can be applied in action, it is to liberate the human body, mind and spirit in the research for better, free world' (Reason and Bradbury 2001: 2).

Focus groups as expert panels

Participants of an expert panel are not just clients or 'well-informed' citizens but professionals in a particular field of study (Bloor *et al.* 2001). Through expert panels, researchers involve practitioners in scientific research (Emerson 1981, Kruger 1998) in a much more 'co-operative inquiry' (Heron 1996), based on participants' professional experience. The disciplinary knowledge is explanatory knowledge and focus groups tend to orient towards explaining why a particular problem exists or how it develops (Ford 1977).

The second type of focus group as an expert panel is to employ it as an element in engaging experts in a policy project. The purpose of such a research design is to involve them not only in finding explanations for a particular problem, but also to motivate them either to find solutions in their own practice or to collaborate in collective projects to solve a problem. Within such a design focus groups are at the same time a cooperative inquiry in a particular problem and a start for a collaborative action to formulate and if possible to implement solutions.

The third type of focus group as an expert panel is the use of it as a method to reach consensus among experts through debate in order to produce social, scientific knowledge and for developing policy action knowledge. A typical example is the Group Delphi method, which aims at reaching a consensus between experts on expected future developments by a carefully designed programme of sequential individual interrogations, usually conducted by questionnaires and interspersed with information and opinion feedback (Helmer 1965, Linstone and Turoff 1975). Assuming an inclusive approach that treats most forms of interactive group interviews as variants of focus groups, expert panels employed within Delphi technique can be considered part of our typology. In this kind of Group Delphi approach, the questionnaires are replaced or accompanied by a series of expert focus group discussions while the researchers might still intervene between sessions to summarize points of agreement and highlight unresolved issues (van Zolingen and Klaassen 2003).

Criteria for judging the quality of focus groups

A critical inquiry into focus group method might usefully start with a discussion of the different perspectives on qualitative research in general. At one extreme, there is a radical hermeneutic position (Dilthey 1976, Ricoeur 1979, Habermas 1976, 1984); in the middle is a moderate interpretative position (Smaling 1992, 1994, 2000, Lincoln and Guba 1985, Guba and Lincoln 1989); and at the other extreme, there is a more realistic position that in a pragmatic way tries to compare, differentiate or integrate the traditions of quantitative and qualitative research (Kirk and Miller 1986, Maxwell 1992, Silverman 2001).

A radical hermeneutic position

Hermeneutics is the philosophical tradition that develops insights about how people make sense of or give meaning to their world with the help of dialogue and discourse, and focus groups provide a hermeneutical perspective. Such theories and applications share the idea of the hermeneutic circle, or the notion that understanding or defining of something employs attributes which already presuppose an understanding or a definition of that thing (Dilthey 1976). Habermas (1976) and Ricoeur (1979), strengthen Dilthey's distinction by noting that in the human sciences the subject of investigation and the investigator are interlinked in a communicative way. The object is constituted by the researcher's web of meaning, which only makes sense if there are objects to be interpreted. It follows from this that cultural and social research has a role in clarifying the beliefs, plans, motivations and social roles that lead cognitive agents to produce the texts and actions.

Habermas (1981) has, in connection with his theory of communicative action, also developed the foundations of an 'objective' or 'radical' hermeneutics. According to him, this necessitates a two-fold renewal: on the one hand, it should be clear that the social sciences inevitably have to take an interpretative stand, which implies that they can only clarify reality while participating in it; on the other, a reformulation of the objectivity principle is needed. Therefore, the ambition to reach a transcendental truth and objectivity should be replaced by the principle of 'hypothetical objectivity', which is based on rationality. This confirms, complements or rejects the hypothetical claim with the help of the scientific dialogue of a research community (discourse). The consequence of this is that we need criteria that help us give direction to such a scientific dialogue or discourse. These can be derived from the way human communicative action is basically organized. The basic structure of communicative action can be synthesized as follows:

(1) actors involved in human communication refer to three distinct worlds: the objective world, the normative (social) world and the subjective world;
(2) in their communication, these actors take into consideration three different kinds of validity claims:
    (a) the claim of truth, relating to the objective world
    (b) the claim of rightness, relating to the social world
    (c) the claim of authenticity, relating to the subjective world;
(3) communicative action is motivated by rationality, meaning that it is directed towards a inter-subjective recognition of validity claims; and
(4) Communicative action is characterized by consensual understanding, which should be interpreted as the cooperative exchange of definitions of the situation and validity claiming.

Many authors in the hermeneutic tradition have adopted these basic principles and have nuanced them, or have elaborated them further. For instance, Kvale (1996) adopts a moderate postmodernist perspective on the issue of reliability and validity into which he integrates a classical conception of truth. Communicative validity can be discussed in this context if we consider his argument that knowledge is a social construction of reality, not a mirror of reality, and the truth is constituted through a dialogue. He also argues that the discussion in the community of researchers is an important aspect of scientific progress moving from knowledge-as-observation to knowledge-as-conversation. This supposes a strong debate about the relation between the method, the findings and the nature of phenomena investigated. Communication of knowledge becomes significant and 'justification of knowledge is replaced by application, knowledge becomes the ability to perform effective actions' (Kvale 1996: 241). Addressing the question of quality and validity of action research, Reason and Bradbury (2001: 450–454) propose a debate and reflection about five quality issues: as relational praxis; as reflexive-practical outcome: as plurality of knowing; as engaging in significant work; as emerging inquiry towards enduring consequences. It is, in fact, an attempt to shift validity 'from a discourse about quality as normative to a discourse of relational practices' (Lather, in press, referred to in Reason and Bradbury 2001: 447).

The next step in our investigation is the question of how these theories about the hermeneutic circle and about communicative action can be applied to the reality of focus groups and their scientific ambitions. How can these insights contribute to

the interpretations and re-interpretations of the data? What is the relation between the explanation and understanding while interpreting the data of focus groups? Considering the above statements, the hermeneutic perspective seems to be a relevant framework to give direction to the scientific operation of questioning the reliability and the validity of focus group research.

A moderate interpretative position

Beyond hermeneutics, we find different perspectives on qualitative research that still departs from the idea that the purpose of research is to find the truth about an objective reality and argues that the standards for such a truth-finding process in quantitative research do not fit with qualitative research. This perspective, a moderate interpretative one, starts from the assumption that the process in which meanings are assimilated to the scientific data is in fact a process of re-construction, re-interpretation of reality. Thus, the reality is not objective but subjective and, moreover, the (subjective) reality researchers perceive is based on the constructions of those interviewed, for example in focus groups. In this moderate interpretative tradition, researchers do not have a detached, objective role and, moreover, they have to use normal language terms to discuss the true value of their findings.

Lincoln and Guba (1985) proposed the concepts of trustworthiness, credibility, transferability, dependability and confirmability to discuss the truth value of their findings. Credibility means that the results of qualitative research are credible or believable from the perspective of the participant in the research. Transferability refers to the degree to which the results of qualitative research can be generalized or transferred to other contexts or settings. Dependability is based on the assumption that for replicability control of changes that occur in the research context is needed. Confirmability refers to the degree to which the results could be confirmed by other research findings.

For Smaling (1992: 178), objectivity represents the higher goal in the methodological context and it means, 'to do justice to the object of study'. In line with this, he developed a bootstrap-conception, a new conception of objectivity as methodological norm, based on three sub-domains of phenomena that might or might not be objective. The first sub-domain includes the physical attitudes and mental activities of the researcher. The second refers to the ways of operating more specific research methods and techniques and ways of cooperation between researchers. The third includes the results or products, such as data, reports, observations, interpretations and theories. In summary, his model entails that 'objective' could be an attitude, a way of acting, a product that gives voice and does not distort a physical phenomenon, behaviour and a human world.

Even when qualitative researchers agree on objectivity as an ultimate norm, many stress that in fact we can only hope for intersubjectivity. But again, there are different perspectives on the concept of methodological intersubjectivity (Smaling 1992). The perspective of consensus is the most evident one and it can be recognized in different methodological ideas: intersubjective verifiability, con-formability, testability, repeatability, reliability or reproducibility, inter-observer agreement, reliability and univocally. Two other traditional perspectives are intersubjectivity by regimentation and by explicitness. Those two are more focused

on the process of doing research. Regimentation refers to regulation needed in order to assure the virtual repeatability, controllability, correctability and criticizability. This leads almost inevitably to the second traditional requirement of being as explicit as possible about the design, specific procedures and interpretation strategy. All of these approaches are relevant not only for designing and accomplishing the research but also, as we will see in the next section, in assessing focus group method.

A realistic position: bridging the gap with positivist research

A third perspective on judging the objectivity of qualitative research posits that there is not such a strict division between quantitative and qualitative approaches. Within this perspective, qualitative researchers apply the central concepts of quantitative research methodology adapting definitions and more concrete standards to the specific characteristics of qualitative research.

At the outset, it is important to outline the positivist tradition, which underlies quantitative research. It starts from a premise that what is knowledgeable is directly observable, measurable and quantifiable and that data are objective and independent of researchers' perceptions. Thus, the emphasis here is on objectivity and replicability because reality is made up of facts and events that persevere long enough to be generalizable. We consider that a more realistic position integrates some elements from positivist research with qualitative approaches, since the former also constructs explanation and is always seeking evidence to evaluate the method. The data, like reality, are at the same time both subjective and objective. Nevertheless, it is important to search for objectivity; and, in the case of focus groups, this supposes constructing accurate, valid and insightful explanations of what was discussed, so that the interpretations might be reached on further occasions.

Kirk and Miller (1986: 19) state that qualitative research can be evaluated in terms of objectivity and, moreover, 'the partitioning of objectivity into two components: validity and reliability', which is adequate for qualitative research. For Hammersley (1987), research results are valid or true if they represent accurately those features of the phenomena that it is intended to describe, explain or theorize. Reliability is the extent to which a measurement procedure produces the same answer however and whenever it is carried out. Silverman (2001) adopts this perspective in his plea for the claims to validity or truth status of qualitative research. A meta-analysis of the key words related to validity and reliability, in both quantitative and qualitative research, has been undertaken by Winter (2000), who argued that the aggregated definition of validity could be accuracy and that of reliability might be replicability. What is important in searching for understanding of the nature of validity and truth is to enquire whether the research is measuring what it was intended to measure.

One realistic approach to validity is Maxwell's (1992) attempt to develop a typology of validity categories, although he avoided applying or adapting the typologies developed for quantitative research. In the same way that there are differences between quantitative and qualitative approaches, there are also important similarities, so that these approaches are not incompatible. He suggested that understanding is a more fundamental concept for qualitative research than

validity. In his opinion, validity refers primarily to accounts and not to data or method. This is consistent with the position expressed by Hammersley and Atkinson: 'data in themselves cannot be valid or invalid; what is at issue are the inferences drawn from them' (Hammersley and Atkinson 1983, referred to in Maxwell 1992: 283). Three of the five types of validity Maxwell mentions in his typology of the kinds of understanding are directly connected with the research process itself and will be used below to assess the validity of focus group method:

- descriptive validity refers to the factual accuracy of the account. An important question here is 'Was there a problem of mis-transcription or mis-remembrance?' This type of validity can refer to issues of omission as well as of commission;
- interpretative validity refers also to aspects of accounts. It is a fact that these accounts must be grounded in the language of the participants and rely as much as possible on the respondents' own words and concepts. The jump from the words and actions of participants to the accounts constructed by researchers is always a matter of inference; and
- theoretical validity refers to the function of accounts as explanation and more exactly to the inference from data to report. This also applies to the postulated relationships between concepts, so that it is also called internal or concept validity.

While validity refers primarily to the accounts, reliability refers primarily to the data and the method. How can we guarantee that a replication of the method leads to the same results? In quantitative research, the solution is sought in a strict regimentation (representative sample, closed questions). In qualitative in-depth research, the problem is much more to suppose that all relevant data will again appear in a replication. The approach should be designed in such a way that it approximates as closely as possible to completeness, rather than limiting data, as in quantitative research.

Elements of focus group design

The authors of this paper take different positions in the methodological debate as sketched above. On the level of epistemology, some support the radical hermeneutic position, others tend to a moderate interpretative position. Nevertheless, all agree that a pragmatic consensus is possible, which includes all three positions, when it comes to data collection and analysis. In order to judge whether different elements of focus groups research are valid and reliable, we will develop a rather pragmatic approach with respect to these epistemological questions. In developing standards for good focus group research, we will apply mainly concepts of validity and reliability as defined in the realistic tradition. Nevertheless, we also found it helpful to apply methodological standards that were developed in the second tradition. But when it comes to moderating and reporting discussions in the focus group, we acknowledge that, at the end, focus group research is, as the hermeneutic tradition underlines, a discourse on the meaning of action.

(1) While preparing the focus groups, it is important to develop a manual (Kidd and Parshall 2000). If there is no manual, the research does not conform to a basic requirement for replicability. How can something be replicated if it is not clearly specified in the first place? A manual guarantees intersubjectivity by regimentation and explicitness.

(2) Research questions have an important role in focus groups. They induce key points for the discussion and guide the whole process of the debate. A lack of precise and consistent research questions damages, in particular, the theoretical internal validity of focus groups. Clarity of research question can provide both relevant answers to the research and ensure that it is repeatable.

(3) As Crabtree *et al.* (1993, in Morgan 1996) have pointed out, a number of logistical factors are important for the validity and reliability of the research. In particular, the lack of a clear time schedule can lead to a situation in which some of the research questions get hardly any attention.

(4) Identifying and recruiting participants is one of the most important aspects of focus group research (Emerson 1981, Vaughn *et al.* 1996, Greenbaum 2000). Validity requires that participants are competent to answer the research questions, but reliability requires variety, i.e. participants must be able to provide a whole range of responses to the research questions. In addition, there is a problem of the optimal size of focus groups: Morgan (1997) suggests six to ten people. However, a rather small group is better for reliable results because it gives each participant more time to raise facts and arguments. Finally, the number and type of participants who are invited into focus groups are clearly under researchers' control, but whether these people really appear is beyond it. Absence of a particular type of participant can lead to serious shortcomings in the data.

(5) Group dynamics have a good deal of significance for the methodology of focus groups. Specific phenomena such as 'groupthink'—group polarization or dominance and passivity of some participants—must be recognized. 'Groupthink' (Johnson and Johnson 2000), for instance, can endanger the validity and reliability of the results since it is a collective striving for unanimity while ignoring information inconsistent with these views (Flowers 1977, Tetlock 1979, McCauley 1989, Aldag and Fuller, 1993). Closely related to this phenomenon of 'groupthink' is the problem of whether it is really possible or desirable to reach consensus in focus groups. Morgan (1998) sees it, in general, as a source of bias to consider focus groups as a medium for producing conformity. Nevertheless, it is important to recognize the focus groups' capacity to lead the discussion in new directions (Field 2000), and in the meantime to prevent, recognize and correct the symptoms of 'groupthink'.

(6) During the discussion, the role of the moderator is to ensure that participants cover each of the research questions, to ask for clarification or further discussion and, eventually, to offer a brief summary (Field 2000). The moderator must be sufficiently involved in the group to fulfil the role of facilitator, but not so dominant as to bias or inhibit discussion (Morgan 1997, Agar and MacDonald 1995, Kidd and Parshall 2000). Overly enthusiastic or aggressive participants who seek to dominate,

monopolize or control the discussions (Greenbaum 2000) must be restrained with care. Marginalized and apathetic participants also need specific attention. Signs of inattention or boredom and those of defensiveness or enthusiasm are non-verbal signals of which the moderator must be aware.

(7) The gender dimension in relation to both the composition of the group and the moderator represents another interesting aspect that can be related to focus groups. In summary, reliability in the sense of facilitating the delivery of a broad range of facts and arguments requires the facilitator to stimulate the active participation of all without controlling the discussion.

(8) Descriptive validity (factual accuracy) and interpretative validity (grounded in the language of participants) requires a recording technique that is as accurate as possible. The best way is to use audio-taping that is subsequently transcribed literally, but even then a lot of the non-verbal communication is missed. So, in some cases, it is worthwhile to videotape the group session from a viewing room behind a one-way mirror (Morgan 1997). Despite the fact that the discussion is taped, it could be helpful to take minutes, as an instrument for a first analysis of the relations between discussed issues (theoretical validity).

(9) Coding and interpreting data are inevitably time-consuming and complex. Coding data itself is not so difficult, but interpreting the data is trickier. Theoretical validity (referring to the research questions stated in advance) can conflict with interpretative validity (referring to what the participants find interesting). The conflict might lead to an iterative process that adapts the original research question so that it is more in line with the focus of the participants. Moreover this also creates problems with respect to descriptive validity. Are the wordings of participants identical, similar, related or unrelated? What was the context of the comments? Were the two respondents really talking about the same thing when they answered? What was the emphasis or intensity of responses? Did the respondents change their position later in the discussion? (Krueger 1998). It is important to mention that the analysis is contextual and not statistical, which means that the quality of the responses and their association are more relevant than their frequency.

(10) Writing a report requires a balance between the direct connotations of participants (descriptive validity) and the scientific interpretations (theoretical validity) of those connotations. Focus groups are productive in at least two ways (Field 2000: 332): they identify issues that are difficult to obtain and they generate material with substantive value. This is a complex process of producing valid knowledge through dialogue, by understanding and interpreting data and by searching for meanings in explanation of reality represents a clear argument for communicative validity.

(11) Finally, the feedback from participants on the draft of the report is important because focus groups are fundamentally a way of listening to people and learning from them. Focus groups create a line of communication (Morgan 1998). Participants should be invited to react on the accuracy (descriptive validity) and the completeness (reliability) of the draft of the report and it is important to know their response.

The research design of this study

Thus far this brief summary of the literature on focus group method has concentrated on aspects of validity and reliability but, as we have seen, there is an urgent need to do more research on how focus group research is conducted. Our research concentrates on four recent European projects in the field of adult education, sketched briefly below. Apart from collecting documents about each project (Leirman 1995, Aguettaz 2000, Weil 2000, Veen and Raak 2001), we used three different methods of data collection, which will be described later in the paper.

Expert panels in the context of four projects in adult education

The most recent project is Etgace (2000–2002) with a central research question on how education and training for governance and active citizenship in Europe can be improved. Six European countries participated in this project (Belgium, Finland, Slovenia, Spain, The Netherlands and the UK). In fact, it was our own involvement in this project that triggered our curiosity in the validity and reliability of focus group methodology. The second project is Balancing Competencies (1998–2000), having an overall objective of generating a more comprehensive understanding of the potentialities and limitations of existing approaches to post-school education and training for unemployed youth and young adults. The countries involved in this project were Belgium, Denmark, Germany, Portugal, The Netherlands and the UK.

Both projects used focus groups with other research methods. The focus groups in the Etgace project were typically expert panels, bringing together professionals in the education and training of citizens. The Balancing Competencies project used two types of focus group in each country: one was an expert panel of professionals and the other was emancipatory research with young adults. We report only the data on the expert panels in this research project in order to present comparable data with Etgace.

The two other projects were based on the Delphi research methodology combining questionnaires with group sessions designed as colloquies. In the Eurodelphi project (1993–1995), the main objective was to compare different models for adult education in Europe. Sixteen countries were involved, but for very practical reasons we concentrated on the experience with focus groups in two countries. The fourth and last project is Saeda (1998–2000), the objective of which was to identify the strategies for the improvement of adult education in bicultural regional settings in Europe. Four regions participated in this project, from France, Greece, Italy and Spain.

Data collection

In the first stage of our research, we sent questionnaires to researchers involved in both Etgace and Balancing Competencies. We used two forms of questionnaire: the first for all national coordinators and the other for all moderators of the focus groups. The questionnaire for the coordinators followed the elements of focus

group research design as described earlier but concluded with items of a more general nature, such as strengths and weaknesses of the method and unexpected problems. All 12 coordinators returned these questionnaires. The other concentrated on two elements of the group: its dynamics and the role of the moderator. The response rate in this case was lower, 62.1%. In the Etgace project (60 moderators in total), we had responses that cover only 40.7%. For the Balancing Competencies project, we got a response from 83.3% of the moderators of expert panels. We used similar questionnaires and have also conducted two in-depth interviews for the two other projects.

Another important element in our inquiry about the validity and reliability of focus groups was a 'focus group on focus groups methodology'—in fact an expert panel—with researchers of all four projects, for an in-depth discussion about the results of our questionnaires and in-depth interviews. There were invited, as experts, 12 coordinators of one or more projects included in our study. The number of experts that participated was eight. The research questions were related to the following aspects: the definition and design of the focus groups, the strength and weaknesses of the method, and the requirements of validity and reliability of the focus group methodology.

The results of the questionnaires and interviews

In presenting the results, we follow the elements of focus group research method discussed previously and limit ourselves to the most important or most surprising results.

Were the focus groups carefully planned or just in-depth discussions?

In the Etgace project, the coordinating staff developed a manual of 17 pages covering all the elements of the focus group research process. It stated explicitly five research questions that were firmly built on the general aim of the research, guaranteeing its theoretical validity. A precise time schedule was developed for two day-long meetings of focus groups in each country. Each of the research questions was allocated to one or more specific sessions during these two days, but there was some freedom for the national teams to reschedule the focus group sessions. Three out of six countries made some changes, most of which reflected the unwillingness of focus groups in these three countries to split up in gender homogenous subgroups in order to discuss gender questions on active citizenship. This might have damaged the theoretical validity of the research. One of these three countries radically redesigned the whole time schedule, and although applying all original research questions, they rewrote the questions in a way that they hoped were more understandable for the experts and therefore made it easier for them to provide the relevant information. Again this could have damaged the theoretical validity.

Within the Etgace project, only two countries perceived the allocated time as sufficient to complete the tasks. The coordinators of the four other countries mentioned real problems, such as insufficient time for experts to give details about examples they brought forward and the lack of time to reflect on intervention

strategies. This has probably damaged the reliability by limiting the diversity of facts and arguments.

The Balancing Competencies project was based on a guide, but with a flexible and open frame. Some common arrangements were made for the group sessions, such as the number of participants, criteria for selection, time schedule and research questions. Despite the flexibility of this frame, comparable organizational arrangements were established and comparable results were reached. Nevertheless, some respondents of the questionnaires mentioned that the organization could have been done better. In this project, the coordinators reported that by and large there had been enough time for discussion. The research questions of the project were introduced and explained but more specific questions were raised in the debate. During the discussion, in some cases, the topic was re-oriented and made more pragmatic and operational.

The Eurodelphi and Saeda projects were based on a 'like Delphi' methodology, as one of our respondents remarked that they employed 'reflective group discussions' (Leirman 1995) in the national/regional and European/international colloquies designed as focus groups. In Eurodelphi, there was a relative freedom in organizing the national colloquies, rather than a fixed procedure in a manual. However, a guide was available; it indicated details as number of subjects and sessions, the three basic questions, time schedule and some recommendation for moderators. There were no major changes in the planning of the colloquies. The schedule and time for discussions were perceived as satisfactory. One of the interviewed coordinators pleaded for more freedom for the researchers to add more critical questions since he did not see the necessity of exactly the same questionnaire for all countries.

The Saeda project was based on the same written protocol for Delphi-type discussions as in Eurodelphi, but there were two rounds of national colloquies and one international colloquy. Compared with the huge Eurodelphi project, Saeda had a limited number of participants, only four regions, which contributed to a good understanding of the specific topics being investigated.

Major problem: non-participation rate

The coordinators of the Etgace national teams mentioned nine different strategies to identify potential participants: the most preferred ways were brainstorming within the national team to identify participants and asking others (advisory panel, colleagues) for recommendations. Two projects adopted just one approach while the other four combined strategies. The data (see table 1) demonstrate two crucial questions. Firstly, many experts cannot or will not accept an invitation. Secondly, experts that accepted the invitation in the first instance did not attend focus groups meetings.

Although these overall figures seem rather satisfactory, the Etgace project team had a specific problem because its manual and research questions required splitting up focus groups during some of the sessions in subgroups (experts in the field of politics, civil society, workplace). The major problem in all countries was that experts on education and training in the workplace were hard to get; they refused invitations or accepted them but did not show up. Ultimately, in some

Table 1.   The number of invited and presented experts in different sessions in the Etgace project

Countries	Invitation	Accord for participation	Show up			
			1st & 2nd	*1st*	*2nd*	*Total*
A	40 experts	40 experts	9	15	15	21
B	45 organizations	30 experts	No data	No data	No data	24
C	103 experts—mail 15 experts—phone	30 experts	No data	20	17	20
D	20–26 experts	No data	No data	No data	No data	19
E	30 experts	No data	10	18	13	21
F	100 experts 1st day 30 experts 2nd day	40 experts No data	15	29	20	34

countries, the subgroups of experts on education and training in the workplace had only four participants, which might have damaged the reliability.

The Balancing Competencies researchers also mention nine different strategies to identify participants, but generally it seems they used more formal contacts with institutions instead of informal knowledge about individual experts (see table 2).

In Eurodelphi, the total number of respondents in the first part (questionnaire) was 1745, between 71 and 196 per country, with a mean of 109 experts. This was not a representative sample, but a large group of 'well-informed actors' (Leirman 2000). For the national colloquies we have only some examples from one country, but according to the international coordinator the tendency was the same in almost all countries. From the respondents in the questionnaire about 65% accepted the invitation for colloquia, but only 35% invited people actually attended. Around 45% of the experts who accepted an invitation did not attend. But, as one of our respondents maintained, this percentage (50%) is not the most relevant indicator but it is 'more important is to get enough people from all domains and each type

Table 2.   The number of invited and presented participants in the Balancing Competencies project

Countries	Invitation	Professionals		Young people	
		Accord for participation	Show up	Invitation	Show up
G	No data	No data	7	No data	6/9
H	21 institutions	No data	15	No data	11
I	7	No data	6	7	4/5
J	6	No data	5/4	9	6/2
K	8	8	7	8	6
L	20	10	10	12–14	5

of experts'. It was also important to have between seven and ten people in each colloquy and this requirement was met on all occasions.

In the Saeda project, there were 213 respondents in the first round (between 45–59 per country) and 153 in second round (between 32–45 respondents per country) in the questionnaires survey. For regional colloquies there are no data available, but, in the case of International Colloquy, the number of participants (as experts) was 17.

In all four projects, there have been discussions on what day of the week is best to facilitate participation of experts. Most coordinators prefer a working day instead of the weekend, with a slight preference for Friday.

Group dynamics

There has been a discussion in all projects how much information should be presented to participants to focus the discussion on the research questions, which can enhance its theoretical validity and reliability. In the Etgace project, participants were sent between 15 pages and 50 pages of material in advance— which included a national report of life histories of active citizens produced in an early phase of the research. In addition, a report on the discussion in the first meeting was sent before the second meeting. The coordinators indicated that they have the impression that many participants did not read the documents thoroughly and some did not read them at all. During the discussions, the experts hardly referred to the documents. Etgace also took care of introductions on the subject at the beginning of each meeting and/or short introductions at the beginning of each separate session. The Balancing Competencies coordinators mentioned that they needed a lot of time during the meeting to inform professionals in order to have a focused discussion. Eurodelphi and Saeda used a mixture of questionnaires and focus groups. One respondent remarked that 'it is not effective to give participants more than four pages [of] information in advance (with summary of questionnaire results)'. So, the general conclusion of our respondents seems to be that, in order to focus the discussion, some written information may be sent in advance but it is more effective to rely on introductions during the focus group meeting itself.

With respect to the question of whether the focus groups were reliable in the sense that discussions were open and critical, we needed information from moderators of the groups (not the national coordinators). A separate questionnaire for moderators was developed and sent to the moderators in both the former projects. Different forms of question were employed to get an impression of the group climate, but we have selected here the answers on two questions, that combined are representative for the total picture. Most respondents (94.4%) indicate that there was an animated and positive atmosphere (Figure 1), but moderators are more cautious in their answer on the question of whether participants were critical about each other's ideas (only 63% agree) (Figure 2).

Facilitating the discussion

Other items in the questionnaire for the moderators referred particularly to their own role (see Figures 3 and 4). In general, moderators intervened to complete
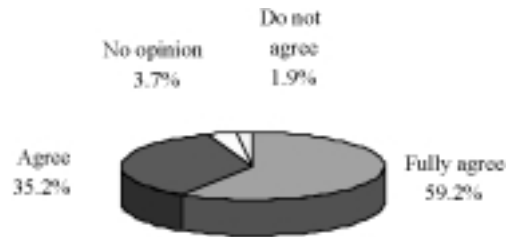
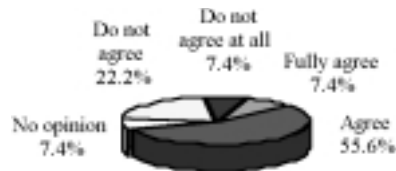**Figure 1.   'It was an animated, positive atmosphere'.**

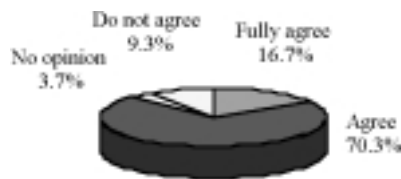**Figure 2.   'Participants were critical about others' ideas, projects or examples'.**

Figure 3.   'As a moderator I have interfered to complete specific tasks during the meeting'.
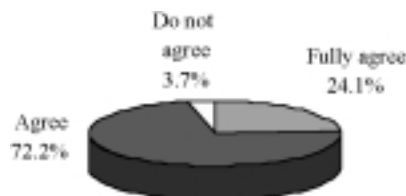

Figure 4.   'I tried to get everyone to participate equally in the discussions'.

specific tasks during the meeting (87.1%), to get everyone to participate equally in the discussions (96.3%), to encourage those experts who might otherwise say little (70%) and to try to limit those experts who dominate the discussions (77.1%). Looking closer, these answers do not say much about what really happened. How often did moderators intervene? How successful were they? Only a precise observation can disclose what really happened.

Recording strategy satisfactory

Earlier in the paper, we tried to formulate a hypothesis about securing descriptive and interpretative validity and suggested that anything less than an audiotape that is subsequently transcribed literally is in fact unacceptable. Our data indicate that, in practice, researchers hold differing views.

In the Etgace project (see table 3), only three out of six countries taped the discussion and each used it in different ways: one country made literal transcripts; another used the tape to make detailed minutes; the third made a preliminary report on the basis of the minutes and subsequently checked it scrupulously by listening to the tapes. Two of the three countries that made an audiotape is reported technical problems leading to a loss of part of the discussion, which is an extra argument to always make detailed minutes.

In the Balancing Competencies project (see table 4), five out of six taped the discussion and five made transcripts (although it is confusing that one country said that it made a transcript without mentioning that it taped the discussion). Three of the countries that made transcripts said the minutes were in fact good enough (which is again confusing because only two of the countries that made transcripts mention that they also have minutes). In general, in the Eurodelphi and Saeda projects, both minutes and transcriptions of tapes were made for significant colloquies.

Table 3.   Recording practice in the Etgace project

Minutes	Tapes	Transcripts	Country
X	X	–	A
X	–	–	B
X	X	X	C
X	–	–	D
X	X	–	E
X	–	–	F

Table 4.   Recording practices in the Balancing Competencies project

Minutes	Tapes	Transcripts	Country
X	X	–	G
X	X	partially	H
–	X	X	I
–	X	X	J
–	X	X	K
X	–	X	L

Coding, reporting and feedback

The Etgace manual suggested a specific procedure for coding based on the central research questions. Countries were asked to divide the records of the discussion into small sections, to attribute each section to one of the central research questions and to make a file of all sections attributed to the same research questions. Within that file, researchers were expected to develop codes that reflect all types of answers to that research question. Four countries followed this procedure. One country mentions that it developed a more systematic, detailed procedure and, quite interestingly, one country followed a more inductive method, not starting with the research questions but with the data themselves, which reflects the central problem in qualitative research, how theoretical and interpretative validity should be reconciled. Nevertheless, all the Etgace national reports had the same structure, reporting the data according to the central research questions. Finally, the draft of each national report was sent to the participants of the focus group, but only a few of them proposed amendments. The international report was structured according to each research question based on the amended national reports.

In the Balancing Competencies project, the coding of the materials from expert panels was reduced to the identification of important data, following the research topics. Relevant data were used for national reports. In this project, the second round of the focus groups was designed to get feedback from participants and, almost incidentally, other feedback concerning this project was obtained.

The Eurodelphi and Saeda projects did not use a sophisticated procedure for coding and interpreting data. In fact, there was only a literal report from discussions or just a summary of the most relevant topics, because the results of questionnaires were seen as the really important part of the research, and colloquies were only designed to check the validity of these data and to determine the potential policy implications. For the participants in the European (Eurodelphi) and International (Saeda) colloquies, it was an opportunity to give their feedback. After the European colloquy, the feedback was not really organized, but, as our interviewees mentioned, there was an informal, positive appreciation.

Strong and weak elements of the focus groups method

We asked our respondents whether there were any unexpected and unanticipated problems in the preparation and the implementation of the focus groups research. The Etgace coordinators mentioned in particular three surprises: the high percentage of invited experts that would or could not accept the invitation or did in fact not show up were, in particular, experts in education and training in the work domain; the lack of time to get all research questions fully answered; the theoretical expectation that there would be a significant difference between opinions from experts from different educational domains was not discovered but rather there was a high level of consensus.

The Balancing Competencies project coordinators did compare in particular the participation of professionals in the expert panels and the participation of youngsters in their emancipatory focus group research. They concluded that

experts were much more motivated and engaged than the young people (although it was difficult to find a day for the meeting of the focus group that was convenient for all).

Researchers in the Eurodelphi project had not expected that it would be so difficult to get all national researchers to agree on the design of the research and on the central theoretical concepts. We also asked all respondents to mention three strong and three weak points of the focus groups method. We combined the answers in some broad categories that are reported in the tables 5 and 6. For the Eurodelphi project, the strong and weak points are presented in table 7. We have no information on strong and weak points from Saeda researchers.

Table 5. The strong points of the focus groups

Etgace	*Balancing Competencies*
Benefits for research – a lot of information in a short time – various perspectives – connections empirical data–theory – potential partners for dissemination Benefits for participants – bring people together – exchange and dialogue – establish connections among experts – start networking	Benefits for research – provide rich insights – multiple perspectives – possibility of communicative validation

Table 6. The weak points of the focus groups

Etgace	*Balancing Competencies*
Outcomes – not all research questions answered effectively – less grip on some issues addressed Organization – limited number of attendees – low attendance and biased composition – time too compressed – changing group composition during two meetings Group process – discussions hardly go in-depth – moderation difficult	Outcomes – rather superficial data – differences between objectives of the researchers and the those of participants Organization – takes a lot of time and effort – methodology without a sound basis Group process – pseudo-democratic – representation of the more talkative, active people

Table 7.   The strong and weak points of the focus groups (Eurodelphi)

Strong points	*Weak points*

Benefits for research
  – possibility for international
    comparison
  – good idea, results
  – well-informed participants
  – good mix of experts from various
    fields
 	
Benefits for participants
  – opportunity to interact with experts
    from the field and policy makers (at
    European Colloquy)
  – good, relaxed atmosphere for
    discussion
  – keynote speakers give inputs to
    stimulate discussions between
    sessions
	Outcomes
	  – difficult to compare results between
	    countries
	  – differences in conceptual level
	
	Organization
	  – rate of attendees at national colloquy
	  – difficult to get people for a whole
	    day
	  – fixed groups not a good strategy
	  – difficult to organize a huge colloquy
	    (European)
	
	Group process
	  – sometimes discussions on a general
	    and abstract level
	  – discussions on same groups not
	    productive
	  – repetitive effects between group and
	    plenary sessions

Discussion

As described earlier in the paper, after collecting and analysing the data from the questionnaires, we organized a 'focus group on focus groups methodology', i.e. an expert panel with researchers of all four projects for an in-depth discussion of a draft of this paper. We invited 12 coordinators; eight of them showed up, including the central co-ordinators of the four investigated European projects. Below we report some important comments of this 'focus group on focus groups research', following the sections of this paper.

The original draft of the 'Definition and Typology of Focus Groups' section mentioned, following the literature, only two types of focus groups: cooperative inquiry and a method to reach consensus. Due to the discussions in the focus group of researchers, we added a third type in the final text of this paper, i.e. the focus group as an element in engaging experts in a policy project. In at least three of the investigated focus groups projects (Etgace, Eurodelphi and Saeda), this was a manifest function of the focus groups. To a certain extent, focus groups were designed as participatory research, involving experts in a project to stimulate their further collaboration on the central topic after the research itself had been finished. At least in one example (Etgace), it was even an explicit and central goal of the financing institution (European Union) to involve in the research the end-users of the results.

With respect to the section about the tension between descriptive, interpretative and theoretical validity, the focus group of researchers stressed the enormous impact of theoretical validity for the quality of the final report. According to these researchers, 'interpretative imagination', 'contextualizing the group' and 'locating

the focus groups in the wider context of research' are crucial. Or, to put it in our own words, the practice of just summarizing the results of focus groups does not make much sense. It is the competence of the researcher to understand the results in a broader context that makes the difference between a correct and an excellent report.

With respect to the technical aspects of focus group research, the focus group of researchers underlined our own point of departure that the huge technical problems should be discussed more in depth. As one of them stated it:

> I think that everybody who has done any real research knows that it doesn't function like that (as the manuals describe it). We do research just because it is difficult and the results are often entirely contrary to what we expect, not even the whole procedure does function as you expect. And it is not good for the research community (to conceal this). I don't see why focus groups as a rather suspect method could not be a beginning place to admit this.

One cluster of problems is around the factual expertise of the participating experts. To begin with, there are no 'total experts in all dimensions (theory, policy and practice)' but experts are just 'people with some reputation in the field'. Next, the drop-out rate is a problem because the researchers 'are depending on experts' and 'it is really difficult to motivate these people to participate'. There is also a danger of a serious bias, because some types of experts have a greater tendency to be absent—for instance, experts working in the profit sector and experts working as high officials for the government. The problem becomes even more serious because a second meeting of the expert panel is often 'more productive, informative and inventive', but it is more difficult to get them there a second time.

Another cluster of problems has to do with group dynamics and the role of the facilitator. The problems range from 'participants that do not read the papers in advance' to 'participants that were so much on their own agenda' to 'I never got what was expected in terms of reflection and open discussion'. The focus group of researchers seems to be even more critical than the literature and our interviewees that returned the questionnaires.

In conclusion, with respect to practical problems, our focus group of researchers stressed the social-psychological problems, such as motivating experts to participate and moderating an open discussion, more than technical problems with respect to recording and coding. They do so probably because these social-psychological problems are much harder to solve, if at all resolvable.

Conclusion

This paper has explored the theory and practice of focus group research in adult education. The methodological literature on focus groups reveals different orientations; some stress the importance of discourse analysis in judging the quality of an account whilst others suggest an interpretative and constructivist perspective, striving for objectivity through intersubjectivity. A more realistic position tries to bridge the gap with quantitative research, in particular to integrate the concepts of validity and reliability in assessing the quality of focus groups' research. Taking this

position as our point of departure, we have tried to illustrate the issues by collecting data on four European projects in adult education that used the focus group method.

Further research into focus group methodology is needed to unravel the importance and dynamics of all presented elements. A survey of researchers involved in focus group projects, as we did, is much too global and rough; it is just scratching at the surface. There is a need for in-depth interviews with invited lay people or experts about their motives to participate or not. There is also an urgent need for a closer observation of the group dynamics in focus groups. Finally, it would be helpful to make a close analysis of how the theoretical validity can be improved without harming descriptive and interpretative validity.

References

AGAR, M. and MACDONALD, J. (1995) Focus groups and ethnography, *Human Organization*, 54, 78–86.

AGUETTAZ, P. (2000) Stratégies d'Amélioration de l'Education des adultes dans des contexts régionaux (Aoste).

ALDAG, R. J. and FULLER, S. R. (1993) Beyond fiasco: a reappraisal of the groupthink phenomenon and a new model of group decision process. *Psychological Bulletin*, 113, 533–552.

BLOOR, M., FRANKLAND, J., THOMAS, M. and ROBSON, K. (2001) *Focus Group in Social Research* (London: Sage).

COFFIELD, F. (ed.) (2000) *The Learning Society: knowledge and skills for employment*, two vols (Bristol: The Policy Press).

DILTHEY, W. (1976) *Selected Writings* (Cambridge: Cambridge University Press).

EMERSON, R. (1981) Observational fieldwork. *Annual Review of Sociology*, 7, 315–378.

FIELD, J. (2000) Researching Lifelong Learning Through Focus Groups. *Journal of Further and Higher Education*, 24, 324–335.

FLOWERS, M. L. (1977) A laboratory test of some implications of Janis's groupthink hypothesis, *Journal of Personality and Social Psychology*, 35, 888–896.

FORD, T. R. (1977) The production of social knowledge for public use. *Social Forces*, 56, 504–518.

GREENBAUM, L. T. (2000) *Moderating focus group. A practical guide for group facilitation* (London: Sage).

GUBA, E. G. and LINCOLN, Y. S. (1989) *Fourth Generation Evaluation* (Newbury Park, CA: Sage).

HABERMAS, J. (1976) Some distinctions in universal pragmatics, *Theory and Society*, 3, 155–167.

HABERMAS, J. (1981) Interpretatieve sociale wetenschap versus radicale hermeneutiek. *Kennis en Methode*, 5, 4–23.

HABERMAS, J. (1984) *The Theory of Communicative Action* (Boston, MA: Bacon).

HAMMERSLEY, M. (1987) Some notes on the terms 'validity' and 'reliability'. *British Educational Research Journal*, 13, 73–81.

HELMER, O. (1965) *Analysis of the future: the Delphi method* (Santa Monica, CA: Rand Corporation).

HERON, J. (1996) *Co-operative inquiry. Research into the human condition* (London: Sage).

JOHNSON, A. (1996) It's good to talk': the focus group and the sociological imagination, *Sociological Review*, 44, 517–538.

JOHNSON D. W. and JOHNSON, F. P. (2000) *Joining together—Group theory and group skills* (London: Allyn & Bacon).

KIDD, P. S. and PARSHALL, M. B. (2000) Getting the focus and the group: enhancing analytical rigour in focus group research. *Qualitative Health Research*, 10, 293–308.

KIRK, J. and MILLER, M. L. (1986) *Reliability and validity in qualitative research* (Newbury Park, CA: Sage).

KRUEGER, R. A. (1998) *Analysing and reporting focus group results* (London: Sage).

KVALE, S. 1996 *InterViews. An introduction to qualitative research interviewing* (Thousand Oaks, CA: Sage).

LEIRMAN, W. (1995) Euro-Delphi a comparative study on the future of adult education in 14 countries between 1993 and 1995. *International Journal of Lifelong Education*, 15, 125–138.

LEIRMAN, W. (2000) *Gearing adult education offerings to the needs of a region and its citizens: Delphi projects leading to the Edufora in Flanders*, Conference 'Education for Social Democracies: changing forms and sites' (London, 3–5 July 2000). Available online: *http://www.ioe.ac.uk/ccs/conference2000/papers/pc/papers/leirman.html*

LINCOLN, Y. S. and GUBA, E. G. (1985) *Naturalistic Inquiry* (Newbury Park, CA: Sage).

LINSTONE, H. and TUROFF, M. (1975) *The Delphi Method, Techniques and Applications* (London: Addison-Wesley).

MAXWELL, J. A., 1992, Understanding and validity in qualitative research. *Harvard Educational Review*, 62, 279–300.

McCAULEY, C. (1989) The nature of social influence in groupthink: compliance and internalisation. *Journal of Personality and Social Psychology*, 57, 250–260.

MERTON R. K. and KENDALL P. L. (1946) The focused interview. *American Journal of Sociology*, 51, 541–557.

MORGAN, D. L. (1996) Focus groups. *Annual Review of Sociology*, 22, 129–152.

MORGAN, D. L. (1997) *Focus groups as qualitative research* (London: Sage).

MORGAN, D. L. (1998) *The focus groups guidebook* (London: Sage).

REASON, P. and BRADBURY, H. (2001) Inquary and participation in search of a world worthy of human aspiration. In P. REASON and H. BRADBURY (eds), *Handbook of Action Research: Participative Inquiry and Practice* (London: Sage), pp. 1–14.

RICOEUR, P. (1979) *Main Trends in Philosophy* (New York: Holmes and Meier).

SILVERMAN, D. (2001) *Doing Qualitative research—A Practical Handbook* (London: Sage).

SMALING, A. (1992) Varieties of methodological intersubjectivity—the relations with qualitative and quantitative research, and with objectivity. *Quality and Quantity*, 26, 169–180.

SMALING, A. (1994) The pragmatic dimension. Paradigmatic and pragmatic aspects of choosing a qualitative or quantitative method. *Quality and Quantity*, 28, 233–249.

SMALING, A. (2000) What kind of dialogue should paradigm-dialogues be? *Quality and Quantity*, 34, 51–63.

TETLOCK, P. E. (1979) Identifying victims of groupthink from public statements of decision makers. *Journal of Personality and Social Psychology*, 37, 1314–1324.

VAUGHN, S., SCHUMM, J. S. and SINAGUB, J. (1996) *Focus group interviews in education and psychology* (London, New Delhi: Sage).

VEEN, R. van and RAAK, J., van (eds) (2001) *Focus Groups: Intervention Strategies for Citizenship and Governance—Analysis of Focus groups* (KUN: Nijmegen). Available online: *http://www.surrey.ac.uk/Education/ETGACE/Focus%20group%20report.pdf*

WALKER, C. L. and TEDICK, D. J. (2000) The complexity of immersion education: teachers address the issues. *Modern Language Journal*, 84, 5–27.

WEIL, S. (2000) *Twinning theoretical and methodological concerns: an emerging paradigm from (youth) research* (UP: Northampton).

WILKINSON, S. (1998) Focus groups in feminist research: power, interaction, and the construction of meaning. *Women Studies International Forum*, 21, 111–125.

WINTER, G. (2000) A comparative discussion of the notion of validity. *Qualitative and Quantitative Research*, 4. Avaliable online: *http://www.nova.edu/ssss/QR/QR4-3/winter.html*

ZOLINGEN, S., van and KLAASSEN, C. A. (2003) Selection process in a Delphi study about key qualifications in senior secondary vocational education. *Technological Forecasting and Social Change*, 70, 317–340.