# Publishing house of the Matej Bel University in Banská Bystrica Belianum

Ivan Kusý, Jarmila Androvičová

# Methodology of Social Science Research

Banská Bystrica 2015 This textbook is a result of the Innovative Steps in University Education in the 21 Century project (ITMS: 26110230109) in compliance with the Operational Programme Education. The project is co-funded by the European Union. Matej Bel University in Banská Bystrica is the grant receiver.

Title:	Methodology of Social Science Research		
Authors:	Ivan Kusý, Jarmila Androvičová		
Reviewers:	Prof. PhDr. Peter Kulašik, CSc. Doc. RNDr. Štefan Kováčik, PhD.		
Science Editor:	prof. PhDr. Peter Terem, PhD.		
Approved by:	Textbook approved January 13, 2015 by the Editorial Board of the Faculty of Political Science and International Relations, Matej Bel University in Banská Bystrica		

© Matej Bel University in Banská Bystrica, 2015

ISBN 978-80-557-0971-0

# Contents

Introduction	5
Understanding Social Science Research	9
Science and Scientific Research: Principles of Social Science Research10	0
Defects and Problems of Scientific Research1	7
Social Science Knowledge and Research: Two Basic Groups of Research Methods	:h 9
Understanding Quantitative Methodology in Social Sciences	1
Different Types of Social Science Research	6
Stages of Research and Research Project	2
Methods and Techniques of Data Collection	6
Explorative Methods: Questionnaire and Interview	9
Observation	9
Document analysis7	2
Data Analysis and Interpretation74	4
Methods and Techniques of Data Analysis: Statistics in Social Science	
Research	1
Statistical Theory and Practice8.	2
Statistics as Science and as Methodological Tool8	3
Charm and Problems of Statistics8	7
Basic Population and Representative Sample8	9
Absolute and Relative Numbers9	1
Means	4
Standard Deviation and Variance	8
Statistical Dependencies10	1
Large Scale Phenomena and Law of Large Numbers10	7
Understanding Qualitative Methodology in Social Sciences112	1
Philosophical and Theoretical Sources of Qualitative	7
	/ _
Methods and Techniques of Qualitative Research	С

Case Study	135
Qualitative Interview	140
Ethnography	143
Discourse Analysis	146
Stages of Research and Research Project	
Data Analysis and Interpretation in Qualitative Research	
Bibliography	163

# Introduction

This book can be used as a didactic text for students of political sciences but it can equally serve students of other social sciences (economy, law, pedagogy, etc.) and others that need to familiarize themselves with the basic principles of social science research. Of course, it is not possible to mention all relevant methods and methodologies and fully and thoroughly describe all methodological problems, procedures, methods and techniques of social science research. This is not our aim, because there is a plenty of different methods and their variations and neither a specialist in social sciences can fully understand and work with all of them.

The aim of this textbook is to describe some selected purposeoriented issues. We suppose that our selection will provide you with sufficient information so that you (if there is no other possibility) will be able to plan and carry out the simple sociological research. No less it is important for us that you will understand the importance of social science research in the formation of a healthy and harmonious society, recognize the need of specific reflexivity that can be found only in the social sciences. We hope you will be able to appreciate what social sciences can tell us about our history and our personal biographies (in spite of their methodological restrictions), that you will be able to better understand the world around us and become resistant to attempts to docile interpretations of facts, able to create your own opinion based on relevant data, information and knowledge. In short, we hope you will become a part of the real intelligentsia of our society and you will not unnecessarily reinforce those who only pretend to be a part of it. To value a word with its real content and a fact with its real meaning. This is the way worthy to follow and in this book we would like to offer you a small piece of it.

We have mentioned the importance of social science research in the discovery of social and individual past. We do not want to argue about the role of social sciences in the present and past. There is an opinion platform that accentuates the fact that the main role of a scientist is to predict the future. For example, **K. Popper** (1902-1994) pointed out that it is not true and he refused negative historicism in social sciences. The idea of historicism meant that on the basis of the recognition of natural

relations of social changes it is possible to prognosticate social development. We want to point out the possibilities of social sciences in an attempt to formulate scientific predictions from this point of view. When we think about the future state in the development of society, we should not consider it as a result of necessity (of the development) but like a set of possibilities (of the development).

In the background of this principle there is a requirement of not violating human freedom. In spite of this, the fact is that in social sciences it is impossible to make predictions, manage and generalize, principally arising from fundamental differences between nature and social reality. The difference is also in the relationship between a scientist and the investigated object – in natural sciences the object of research is not individuality with its own consciousness. On the contrary, in social sciences it is possible to say together with Wallerstein that scientists investigate themselves, their mirror or their antipode.

In addition, simultaneously with learning we change the social reality in social sciences. Knowledge as information becomes a part of social reality. Giddens talks about the reflexivity of knowledge. The principle of reflexivity is based on continuous verification and change of social action and it is reshaped in the light of new information about the action itself. This information changes its character. Action as a realization of many theories must be influenced by these theories because they are a part of social reality. If a vision of future is introduced on the basis of some theory of social development it raises pressure on the actions of an individual. As long as the vision is acceptable and accepted it requires that an individual should integrate some values determined by this development. The connection of knowledge and reality is one of the most important methodological characteristics of research in social sciences. It is important to keep in mind that the theories and facts are in a reflexive relation and it is difficult to distinguish between them.

The next attribute of social sciences is their cultural interdependence. Different theories are built up on presuppositions that are basically culture conditioned prejudices. Prejudices are based on different things such as religion, nationality, race, gender, political orientation... In the past, social sciences were particularly oriented to nation, state and national culture. Within them social processes should not only occur, but they should also be investigated (they become a part of

national science). The majority of sociologists sympathize with the national society they belong, similarly most of the political scientists like the liberal democracy and most of the economists prefer market economy. These are certain frames for us, but simultaneously limits of action that are not (naturally) continuous – they have their own time interdependence, time dimension. So we can consider them prejudices.

In conclusion, we can say that we are not able to exist without prejudices because they are our certitude. Initial "prejudice" in the process of learning is language that influences the way we think. **Understanding Social Science Research** 

# Science and Scientific Research: Principles of Social Science Research

As we have mentioned in the introduction, social scientists are always a part of the subject of their research. They are a part of the society that they study so they always have some kind of relationship or attitude to the studied problem. The results and the ways of their research are influenced by their own experience as well as by the experiences of others. Their own (empirical) experience is a natural part of social science research. It is also very valuable and useful, particularly when social scientist knows how to use it, if he/she has developed a **"sociological imagination"**. **Charles Wright Mills** (1916 – 1962) says that sociological imagination means to understand connections – to understand the world of an individual in a context of structural determinants and limitations of the world he lives in. It is the ability that belongs not only to sociologists but also to other social scientists, but as well to the journalists and writers. The works characterized by sociological imagination were often awarded the Nobel Prize. (Mills, 2002)

An important tool of sociological imagination is differentiation between personal problems of an individual and public problems. Many personal problems are conditioned by an individual's personal characteristics, his individual situation and his own biography. Solutions to these problems are in this case personal. An individual has to search for adequate strategies to solve his problems using available tools. On the other hand, public problems cannot be solved within individual strategies and possibilities. Mills (2002) introduces an example. Imagine that in a city with 100 000 citizens only one person is unemployed. It is his personal problem. He should think about himself and ask himself what he is doing wrong and seek for a solution. However, imagine that in a city with 100 000 citizens one third of the people are unemployed. In this case it is a public problem. It is not possible to solve it personally with reference to the personal effort of individuals. The structure of personal possibilities in this case failed itself. This simply means that in that region there are not any available jobs, or the structure of those jobs does not correspond with the needs and qualifications of the people seeking jobs. The problem cannot be solved by an individual's strategies. It has a **structural character**. In this case the problem will persist till structural changes caused by different factors (one of them can be and should be the targeted activity of politicians) will occur. We can see this problem in the example of Slovakia, particularly in large regional differences in living conditions and standards.

The personal experience of a social scientist is thus a part of his cognitive process like unreflected or reflected experience. Social sciences are in general **theoretical and empirical sciences**. Empirical experience is a crucial component of the cognitive process. Theory as a set of knowledge tries to generalize empirical experiences. To build up, widen and enrich social theory, it is necessary to confront theoretical knowledge with empirical experiences on one hand, and on the other hand to generalize and integrate knowledge gained from empirical experience into a certain theoretical system. In this case we do not speak only about subjective living experience of a researcher entering into this process but we speak about a **systematic, planned and meaningful collection of experiences – about the research**.

We can describe the relationship between theory (T) and research (R) as the following:

 $T \to R \to T'$ 

Research is carried out on the basis of a certain theory consisting of existing knowledge, whereby interpreting research findings contributes to further development and enrichment of the theory. T' is thus theory enriched by empirical findings. So the gathering of scientific knowledge and building of scientific theory is marked by a kind of cumulativeness.

We have used the expression "a kind of" on purpose. In his book, The Structure of Scientific Revolutions, Thomas Kuhn (1997) points out, that sciences do not always develop in a cumulative way but in jumps that he calls **scientific revolutions**. Scientific revolution means refusing of a big part of theoretical knowledge accumulated in the previous period. In natural sciences, which were Kuhn's priority interest, scientific revolution can be caused by a discovery, invention or ground-breaking theory. Newtonian Physics transformed into Einsteinian Physics could be an example. The change in this case does not mean that all principles of Newtonian Physics lost their validity but that their validity is limited and it is not universal. Kuhn calls this process that occurs in natural sciences the **substitution of paradigms.** Old paradigms that have showed to be insufficient in explaining new knowledge and topics are substituted by new ones that bring principally different views and logical and acceptable explanation of new findings.

However, this process is much more complicated in social sciences. Kuhn calls them **multiparadigmatic**. It means that in one period more paradigms can co-exist. Scientists working within one paradigm are in mutual relationship because the subject of their interest is the same society that is historically determined and unique in a given time. It is different from any other society that passed through the different historical development and is on a different level of cultural development. Simply, it is not possible that the same paradigms would be valid in different societies. Dynamics of development of these subjects of social science research lead to another specific – they must repeatedly (continuously) re-describe what has been already described, but with relatively valid conclusions. In comparison to natural and technical sciences, time is significantly important in social sciences.

Scientists working within one paradigm share the same basic epistemological and philosophical presuppositions, a certain base of theoretical knowledge. And what is important for us, even a certain set of methods and methodological procedures. Theory and methodology influence and complement each other. Different philosophical and theoretical presuppositions of scientists influence their approach to study of social reality.

For example, if I suppose that the political behaviour of individuals is mainly influenced by the character of political system, I will probably focus my research on an analysis of that system – the institutions and laws that create it. If, on the other hand, I suppose that political behaviour is conditioned especially by their social relationships and typical interests or qualities, my research will be focused on them and it can have a form of secondary data analysis (for example, statistics or another already existing research) or field data collection in the form of a standardized questionnaire. On the contrary, I can put my concern to the way in which politicians perceive their own political behaviour and their place in politics and the political system. It is possible that in this case, I will make a series of in-depth interviews.

In simple terms - certain presuppositions that a scientist bases his research on are conditioned by his choice of a certain method or methodology. The choice of method is to a large extent a function of the research intention. It is evolved from what the scientist wants to find out and which kind of questions he is asking himself. Even the kind of questions he is asking, partly evolves from his theoretical orientation. On the other hand, we should not absolutize theoretical and paradigmatic affiliation. One scientist can effectively combine and use knowledge from different theories but also different methodological approaches.

Talking about equilibrium of theory and empirical experience we can find two extremes. On one hand, **absolutization of theory** is based on the presupposition that a society and its structures exist only in consciousness and because of this it is not possible to approach it by research tools. Absolutization of theory neglects confrontation of theoretical knowledge with empirical experience and main tool of scientific work is deduction. It produces deductive theory based on the reasoning that is derived from axiomatic statements. On the other hand, another extreme is so-called **empiricism**, as a one sided orientation on the facts. Empiricists approach rectilinearly to measuring without previous theoretical rooting of measuring tools.<sup>1</sup>

Based on what we already know, let's try to answer the question: what is science? This question could be answered in two ways. On one hand, we can search for certain objective criteria. Most often the basic criterion is that science must have its own subject and method of

<sup>&</sup>lt;sup>1</sup> Both of these problems were identified by C.W. Mills in American sociology in the '50s. One of these extremes Mills calls the "grand theory", and assigned it mainly to structural functionalists, that dominated in American sociology of that time, with leading person Talcott Parsons. To make it clear - Mills does not call into question the contribution of structural functionalism to the development of sociological theory, however, he points out too much complicated nature of the grand theory and the underestimation of empirical verification of findings in this approach. On the other hand, Mills analysed so-called abstract empiricism in the American sociology of those times. Methodology is a prison of sociological imagination in it. Application of method in the new context is without sufficient theoretical reflection. Researches are thus produced like in a factory and methodological procedures are taken over like recipes. Thus, according to Mills abstract empiricism is actually bureaucratizing sociology. (Mills, 2002)

research. In this case, several problems would appear. Firstly, different sciences, especially social sciences, often share the subject of their research. For example religion is the subject of interest in sociology, religious studies, anthropology and from certain point of view even in political science. And on the other hand, different scientific disciplines often share methodology, methods and techniques of research, or even a whole paradigm. For example, ethnography, as a specific method, can be used in ethnology, sociology, but also in political science. Paradigm of structural functionalism has influenced the development of ethnology and anthropology, as well as the development of sociology and political science. Searching for a specific subject and method of certain scientific discipline was the focal point, mainly in the period of inception of particular social science disciplines. In those times it was necessary to present "tangible" evidence in front of the scientific community about the rightfulness of the designation of institutionalizing cognitive discipline (for example, sociology) into a science.

The fact that that evidence was required by the scientific community guides us to the second possible answer to our main question: **the science is what the scientists** (those who create scientific institutions) **consider to be a science** in the given field, in the given period and, in the case of social sciences, in the given territory. Since this agreement is in constant development, it seems to be useless to try to find a universal definition of the given science.

The main activity of scientists is refinement of a paradigm, its application and elimination of inner conflicts and ambiguities of the paradigm. In the light of new facts, building blocks of the paradigm or paradigmatic system of the given science can be changed or even rejected. The science, in which a scientific community accepts common paradigm, is called by Kuhn "normal science". **Paradigms** are accepted examples of contemporary scientific practice including laws, theory, application and instrumentation. This all provides a model that creates a certain, coherent tradition of scientific research.

The existence of a paradigm economizes science - it is not necessary to again and again excuse, give reasons for and promote new conceptions if they are a part of a paradigm. On the other hand, the paradigm is a kind of restriction - it determines a set of problems that can be solved within normal science and it defines a set of allowed solutions and procedures. It is like a jigsaw puzzle that has only one solution, when accepting these rules or restrictions.

Kuhn does not consider as criterion of a good jigsaw puzzle the fact that the result of scientific research is interesting or important. It must principally and exclusively correspond with the paradigm. To work out of the paradigm is very risky. He asks an open question whether social sciences have any paradigm at all. He states the thesis that sociology is not a paradigmatic science. Among other things, it is a consequence of the fact that for every combination of place and time that determine an observed problem, it is possible to set another paradigm.

We cannot compare the reliability of findings in social sciences with exact sciences. Their conclusions are of **probable character**. Conclusions in exact sciences are deterministic and they are universally valid. Natural sciences are able to use experiment to describe causal relationship among variables. This is very difficult and often impossible in social sciences. When analysing social phenomena it seems that they are mutually connected and interrelated. This is one of the reasons why it is more difficult for social scientists to formulate paradigm.

It is necessary to find some help in this case. It can be the conception of a "natural system" mentioned in work of the Czech sociologist Miroslav Disman (2002). A natural system is a set of interrelated variables, since its connections with other natural systems are weaker. Natural system is informatively closed. This is an important presupposition we either really or sincerely believe in. If this was not true, it would not be possible to explain the behaviour of that system. If a natural system is appropriately created (defined), it is enough to control only a few known inputs that come from other natural systems.

Take an example from physics – determination of a boiling. It is necessary to observe the temperature of water, air pressure or water purity. To describe this natural system (in physics) we need just a few elements and we do not need any faith because we can be sure that elements mentioned above are really all the ones that can influence the whole process.

Social scientists do not have at their disposal such a simple situation. We can compare it to sociological intention - to state the factors

influencing family's annual income. It is necessary to observe jobs, work position, working place, education, experience, age, gender, medical condition of all family members, take into account a size of household and family relationships, membership in different organizations, economic situation - profits and losses, relatives living abroad, access to cheap goods, etc. We could continue to calculate other possible factors. There are a lot of items and some of them need a more detailed description and measuring of several sub-items. It depends on experience and possibilities of a researcher which items will be put in "his/her" natural system. His/her definition then definitely influences the results of research, the method of hypothesis verification and influences perceiving and treating of the given social problem in a society, thus attributes to the definition of social reality itself.

# **Defects and Problems of Scientific Research**

Complicated mutual relationships in social sciences can lead to several misrepresentations. Some of them are going to be presented here in simple examples. To simplify the explanation we will use only three variables instead of the whole set of variables in natural system. One of these variables is unknown or unmeasurable. Possible misrepresentations in perception of relationship among variables and in their interpretation can be, according to Disman (2002) classified into four types.

# 1. False Correlation:

As an example we can use the relationship between a higher number of storks and a higher birth-rate in a certain area. On the basis of our experience, of course, we will not support this, on the first sight apparent, relationship and after a deeper study we will discover a variable in the background that influences both inquired variables. It is an improved state of environment.

### 2. Developmental Sequence:

An example is a higher tendency to read pornographic literature in those students that behave at school in a violent way. There was one cause above the supposed one of the observed aggression. The cause of reading pornographic literature was in an authoritarian approach of the parents' upbringing and the need for compensation and escape. It is possible to go further and seek for a cause of a cause. Can we in this case speak about parents' dissatisfaction with their own lives? Can it be influenced by uncertain job or unemployment? Where to seek the cause – and what finally influences the atmosphere in the classrooms (and even maybe in the whole society)?

# 3. Middle Element Missing:

The research has shown that gender notably influences the results of the intelligence test. Women had worse results in the given example. The important thing is that men were authors of the tests and they acquired different skills during the socialization than women. When intelligence is studied as an ability to cope with chosen skills and abilities, it is evident that it is more difficult for women to deal with those chosen (and naturally acquired) by men.

#### 4. Dual Cause:

We can use an example of the relation between the level of education and income. Education is a very important factor for income. However, education itself is not sufficient enough to explain the level of income. Even people without education or people with an education that does not correspond with their job (or occupation) have high incomes. Income has several independent variables and it is necessary to choose the most relevant into a natural system. Multi - conditioning is typical for social phenomena. Thus, not taking into account all important factors is the most common misinterpretation.

Since we never work with a natural system that is completely described, we will always have to deal with above mentioned misrepresentations. They will not always be as obvious as those mentioned above because we will usually work with more than three variables. Another problem is the **transformation of obtained information** - almost every time there is a need to record and transform reality into language. In sociology we often ask questions and get answers, even though we could obtain information by direct observation. To ask a question is simply quicker and cheaper (in case we need the information such as number of children, ownership of a car etc.). Every indirect measurement (based on the statement of a respondent) is thus increasing the risk of misrepresentation because a statement of the respondent can be to some extent distant from reality. We have to work with these results because we do not have others.

It is important to ensure validity of observation in social science research - achieve that we measure what we really planned to measure<sup>2</sup>. The choice of indicators should ensure the reliability of measurement - it is the measurement that gives the same results when repeated. We will mention the problem of validity and reliability of the research later.

<sup>&</sup>lt;sup>2</sup> If an indicator of attendance of exhibits in a museum would be the shabbiness of the carpet in front of a particular exhibit, we could misleadingly conclude that the most popular is the one placed near the only toilet in a museum.

# Social Science Knowledge and Research: Two Basic Groups of Research Methods

We have already mentioned basic differences in the nature of cognitive processes in natural and social sciences. Natural sciences are described as allegedly exact because they are able to produce exact and reliable conclusions such as "if *x* then *y*, under certain permanent conditions". Scientific character is in a large extent determined by the ability to present generally valid conclusions. We know that any generalizing statements have only probable character. They are valid with some restrictions – restrictions of time, place and context, and they are often ambiguous. For example, if I claim (and empirical research confirms my claim) that older people are more conservative than younger ones, this statement is not 100% valid. In other words, I can always find exceptions that do not confirm the rule – there are older people who are not conservative at all and on the other hand, there are younger people who are conservative.

At the same time we need to say that formulation of generally valid findings (although only of probable character) is not the only aim of social science research. In other words - we are not interested only in to what extent the statement is true. Often we want to know what conservatism means, how it is demonstrated in different contexts, for example in particular countries. We want to know in what way these conservative opinions are formed and how the participants justify them. Many questions that we ask are aimed at inner particularities of the phenomena and their aim is to understand these particularities.

These two cognitive aims – **to generalize vs. to understand inner particularities** - are those that are the basic source of differences between so-called qualitative and quantitative research. Both kinds of research have a strong tradition in sociology. However, the paradigm of quantitative research has dominated for a long time. It is based on the positivist conception of science that proclaims the necessity to build up social sciences following the model of natural sciences. On the other hand, from the very beginning of the existence of social sciences, opinions that refused that tendency were quite strong. Scientists following these antipositivist opinions claimed that social reality is so specific that it is not possible to apply methods and procedures valid in natural sciences to understand social phenomena. The aim of social sciences should be to understand the activity of participants. This is what natural sciences do not have to do and, at the same time, are not able to do.

We will mention both traditions of social science research in this book; moreover, they are the criterion structuring the content of the book.

Favouring quantitative or qualitative research is sometimes paradigmatic, i.e. we can find exclusive supporters of one or the other methodological approach which, who at the same time underestimate the meaning of the other. On the other hand, a strong opinion group has been formed among sociologists, according to which such a restriction is useless and unproductive. Quantitative as well as qualitative approaches to social research are both meaningful and significant in social sciences. We should not perceive both traditions as mutually exclusive but as two methodological approaches that we can choose from, mainly according to the research aim, and at the same time, as methodological approaches that can be effectively combined.

This is the position we adopt too. Since the social reality has its own special meaning and quality on an individual, a group or a social level, it is natural to carry out research on all of these levels. Every level has its own meaning and it is essential if we want fully understand certain social phenomenon. Quantitative and qualitative researches are not opposite each other, they complement each other. One studies collective phenomena, the other typical ones. Individual and collective can be described as well as particular and general. A doctor or a social worker examines the particular case of an alcoholic. The subject of their interest is the personal qualities of the addicted, his medical condition, social conditions - his family life, housing conditions, working conditions etc. According to the individual diagnosis, either the therapy or concrete strategy in social work is chosen in order to help the person to cope with his/her problems with alcoholism. On the other hand, the general phenomenon of "alcoholism" is examined by the representative research using so called representative sample that is large enough (for example, 2500 respondents) to obtained results that can be generalized. Which means, thanks to the information gained on the representative sample (those 2 500 respondents) we are able to say something about all alcoholics (e.g. in Slovakia).

# Understanding Quantitative Methodology in Social Sciences

In the previous part, we have already pointed out that generalized findings have only **probable character** in social sciences. This fundamental difference between social and natural sciences can be demonstrated in the following example: from physics we know that water boils at 100 degrees Celsius (at certain defined conditions – at a certain altitude, temperature and density), it has general validity in comparison to the sociological statement that older people are more conservative. This statement is not equally valid in all societies and cultures, but its validity is restricted and differentiated by time and area.

Is there any other difference between these two findings? Look at the nature of causal explanations. When we would in the first case ask the question "why water boils at 100 degrees Celsius", a physicist would answer it by exact description of what happens with water molecules at this temperature. So he would explain the causality of the phenomena. But look at the sociological finding: what is the cause of the fact that older people are more conservative? How should we explain the causality of this phenomenon? Causal explanation would be much more difficult in this case because of a lot of variables included and because of many possible alternative explanations. This means that causal explanations are more difficult in social sciences and often quite impossible if we treat them in the way natural sciences do. Although, this does not mean that social sciences should give up on causal explanations completely. The search for causes of the studied phenomena is a basic driving force of the human desire for knowledge in the whole spectrum of science research. But they have to use different level of analysis and, on the other hand, they can go much further in the interpretation of obtained facts and findings.

The basic tool for detection of causality is an **experiment**. Experiment is more often used in natural science research and less likely in social sciences. In social sciences, experiment can be done only in some very specific situations. Here, the causal connections are most often concluded without any confirmation by an experiment. There are several reasons for this. One of the most important is the problem of the artificial laboratory environment that should be a simulation of the real social world. People would not behave naturally in such conditions and moreover many phenomena are not possible to be included in them. In natural sciences, to simulate the necessary conditions in laboratory is much easier. Many correlations in social sciences do not have a causal but only a functional character. The measured connection of phenomena does not necessarily mean that the occurrence of one is a direct cause of another one. In case the researcher discovers certain relationship between the two or among several phenomena, then it depends on his/her interpretation whether he/she will present this relationship as causal and what other phenomena and variables will be included in the explanation scheme in order to prove the causality.

The aim of quantitative research in social sciences is to verify the general validity of a statement and to confirm the correlation between two or among several phenomena. Since at the beginning of the cognitive process there is a non-verified statement, it has the nature of presupposition. This kind of non-verified statement is called a hypothesis. Quantitative research is thus focused on testing the validity of those statements, so we speak about the **testing of a hypothesis**.

For example, if I want to verify the validity of the statement that older people are more conservative and to realize whether this claim is valid in the Slovak population, I will design my research in a way that will enable me to find evidences thanks to which I will prove or refuse this presupposition – thus accept or refuse my hypothesis. In social sciences, a situation wherein we are able to examine whole population rarely occurs. More often, it is not possible and so we have to examine a "research sample". Its important characteristic is that it reflects selected essential characteristics of the target population. It means we will choose only certain people/objects from the target population, but at the end we will generalize our findings on a whole population. This is called representativeness of a research. Thus one of the most important qualities of quantitative social science research is its representativeness that can be achieved by appropriate (proven and adequate) selection of the sample following some generally accepted and effective rules of selection.

Representativeness of research results is one of the important criteria for evaluation of their quality. Other important criteria are: **validity and reliability**. Validity – is a quality of a research and its specific measuring tools and procedures that testifies whether the research/tool measures the right thing. In another words, whether it depicts characteristics of the given phenomena and covers its content.

For example, in research about living standards in the fifties and sixties one question, used as one of the indicators, was common. Researchers used to ask whether the given family eats meat at least two times per week. This indicator (together with many others) should help to create an image about the living standards of Slovak households. At those times it was a valid indicator. It is not necessary to emphasize that nowadays it is no more a valid indicator of living standards, as the conditions have changed and phenomenon in alimentation – vegetarianism has occurred.

Another important characteristic of research is its reliability. Reliability of primary data means that we achieve the same results in repeated measurements. Reliability is a characteristic of measuring tools, measurements, operations or whole research procedures. By violation of validity a systematic error occurs in research because an invalid indicator "deviates" all answers. Reliability is related to the occurrence of random error. For example, a question based on a valid indicator can be formulated incomprehensibly so that only a part of respondents understand it. Another part of them can thus skip the entire question. This is a random error. Reliability expresses the degree, size and number of errors and deviations that is to some extent tolerated (most often up to 10%).

A degree of reliability is then a quotient of correct answers from all the answers.

$$R_{tt} = \frac{V_t - V_e}{V_t} = 1 - \frac{V_e}{V_t}$$

 $V_e$  – measurement errors  $V_t$  – total results  $R_{tt}$  – reliability of research results

It is clear that we are not able to calculate or guess the degree of reliability. Because of this, the best thing is to try to decrease the possibility of error occurrence – the random as well as the systematic ones. While we can to some extent guess the reliability, validity can be secured only by logical procedure with clear definition of key terms and concepts.

So, how we can define research? **Research** is a planned and thought out procedure that presents some contact with practice that is not intuitive but it is based on empirical experience and is focused on a certain aim (research intention). A certain more or less standardized (united according to some rules) procedure is applied. It is important that this procedure is planned, systematic and logical. Otherwise, two extremes can occur:

a) <u>instrumental approach</u>: automatic and insufficiently reasoned application of certain standardized methodological procedure on different problem or problem in a different context; without the presence of the subjective contribution of a researcher.

b) <u>methodological intuition</u>: unplanned improvisation, not reasoned and distant from standards. Subjectivity of the researcher is unscientifically dominates.

To avoid extremes like this we have to know well not only the internal logic and procedure of particular methods but also conditions enabling their use, and their essential pros and cons. The application of any standardized method or procedure in sociology must always be confronted with the nature of the problem, with emphasis on its context.

Since we use in this book the terms "method" and "methodology" it is necessary to clarify their meanings. Methodology indicates certain general approach in revealing the social reality. When we speak about the methodology of the cognitive process, we include some epistemological and theoretical bases - basic paradigm. For example, in the methodology of qualitative research those bases and rules that all qualitative methods have (an interview, an observation etc.) in common are included. The second term - method can be defined as a set of universally connected and interrelated rules that create a specific procedure in the discovering of social reality. The term "methodology" is more general and it usually contains a set of methods. In the textbooks and texts about social science research other standard terms can be found - for example, the term technique. **Techniques** are operations of schematic and routine character (an exact border between them and methods cannot be defined). The term "technique" can be used also for complements of used methods - for example, techniques of keeping records.

# **Different Types of Social Science Research**

We come across different types of research in research practice. We introduce the classification inspired by Alexander Hirner (1976) according to the following criteria:

### 1. Criterion of the Use of Research Results

#### **Basic Research**

Its results are determined for direct development of science and only by means of it for practice. Research is planned in a way so that the solution of the problem would contribute to the development of sociology as a science and to the development and broadening of the general sociological knowledge database. Basic research is beneficial to the progress of a science by extending basic knowledge about relevant phenomena and processes, facts and relations among them.

# **Applied Research**

It is planned in a way that its results could be used directly in practice. It uses already existing and verified findings in a new context and to solve a new practical problem. Results of applied research can indirectly become a part of general sociological knowledge - by their renewal that reflects developmental changes in society. Applied research is thematically oriented on contemporary problems in social practice - it helps to detect, solve and predict them.

Example: Within a subdiscipline of the sociology of organization we are interested in informal relationships in a workplace and their impact on work performance. We will carry out the research in several chosen organizations and we will generalize and systemize our results. It is the basic research. On the other hand, as sociologists working in a company that provides complex audits in different organizations we can examine informal relationships at workplace too. In this case our conclusions will be used as a part of possible practical changes in the organizational structure of the company. They will serve to managers of the company as foundations for further practical activity – their work with these informal relationships in order to increase work quality and efficiency, to create better working atmosphere, and so on. This is applied research.

2. Criterion of the Research Sample Selection

#### **Comprehensive Research**

Comprehensive research includes all members of the examined sample or all cases that reflect the examined social reality. It is possible to apply comprehensive research only when we know (if we have enough funds, time and workers) to capture the whole mentioned population. This is possible usually in the case of smaller populations (for example, patients in particular treatment programmes), in bigger ones only exceptionally (for example, census, house and flat census once every 10 years).

#### **Representative Research**

Representative research is based on the selection of a sample from the examined population. We are forced to use representative research when it is not possible to capture all the elements of the research subject. We have to choose an adequate part (research sample) using standardized rules that provide a representative selection.

Example: We examine the activity of organizations aiding refugees – we are interested in the type of activity and cooperation with the state or other participants. Since there are only a few of this kind of organizations, we can visit all of them when making our field research. We can make interviews with selected employees and study different materials from all of these organizations for example. It is a comprehensive research. In another type of research I focus on another type of organization – on senior centres. They can be found in almost every town and in many villages in Slovakia. So it is necessary to select for our research sample only some of them. Of course, it is important to choose relevant criteria for this selection – for example the size of the community where the centre is, number of members, length of its functioning and so on. If we work with the sample selected in this way it is representative research. We will talk about the methods of sampling later.

# 3. Criterion of the Scientific Aim

# **Descriptive Research**

The aim of the research is scientific description of particular social problem (so called "sociography"). We can mark this kind of research as diagnostic as well – it answers the question how certain part of social reality looks like and thus provides the necessary materials for its diagnosis. It is possible that later the research results will be further developed; they will serve for example as a base for the formulation of hypotheses in explanation research.

# **Comparative research**

The aim of this research is the comparison of a certain social reality in two or more situations that represent area, group or time. So there are two equally designed but independently carried out areas of research. The question of comparability must be solved in advance.

# Explanatory research

The aim of this research is not only to find out and describe a certain social phenomenon but also to explain it. It means to reveal the whole context –the relations with other phenomena, and mostly its possible causes and consequences. This type of research is thus most demanding.

Example: As political scientists, we are interested in the healthcare security system in country X. Since we do not have sufficient information in our area, we must, at first, collect and analyse relevant documents and legal regulations for this area and to describe how it works in the given country. It is a descriptive research which gathers the information that was absent, not available or not systematized. In the next stage we are interested in a comparison of this system with the system in another country in order to detect differences and similarities, advantages and disadvantages. It is a comparative research. Later, we will try to explain these elements of the healthcare security system and its particularities resulting from a specific political culture and tradition in the given country but also in connection with the political and economic system. It is an explanatory research.

# 4. <u>Criterion of the Level of Complexity</u>

**Complex research** – research focused on a survey of all fundamental aspects of the phenomenon

**Partial research** – research is focused on some narrowly restricted problem. Main task is to find out what is the current state of a phenomenon – for example, a survey of the market or public opinion polls about hot issues.

**Pilotage** – is a field research operation that precedes the work on the research project (preparation for the main - complex research in order to get as much as possible information about the examined field). Another definition describes pilotage as verification of the research tool, suggested methods, suitability of the contents and of formulation of questions in a questionnaire. It is thus a kind of a trial research.

Example: We are interested in the economic level of Slovak regions. We evaluate it according to several complex indicators, for example, rate of employment, GDP per inhabitant, transport and social infrastructure, available services, etc. It is a complex research. Further we find out that a high unemployment rate of women in some regions may be a consequence of a lack of preschool institutions – kindergartens and nurseries. We carry out a quick survey of their occurrence and availability (survey of statistical documents for example). It is a partial research. Since we have decided to focus next research on one specific marginal region and to examine strategies that the inhabitants choose to "survive" in such unfavourable conditions, we visit some households to get a better image of their situation and to verify our research tool. It can be modified before the main research on the basis of acquired information. It is a pilotage.

5. Criterion of the Time

**One time research** – is a research action that takes place only once when examining the given problem.

**Panel research** – is a multiple research action of the given problem on the same research sample (panel). This research serves for recording

changes in time. At the same time we are able to record changes in opinions and needs of individuals or groups creating our panel.

**Longitudinal** (or continuous) **research**– is a research that is planned for the long run – for example, long-term research monitoring of rural development. It takes place continuously or in certain intervals. Compared to panel research, it is not carried out on the same sample. During repeated and long-term research new problems are created. We reflect them in the contents (and form) of follow-up research actions.

Example: Before the 70th anniversary of the Slovak National Uprising (SNP), we carry out research focused on opinions of the public about the importance of SNP in Slovak history and the importance of commemorating it. The research takes place only once - it is a one-time research. In the pre-election research, research agencies carry out pre-election polls of particular candidates or political parties. They repeat the process but the sample is different. It is longitudinal or continuous research. If I want to understand the real development of election polls and the formation of election choice of potential voters, the most adequate is the method of panel research so that I will repeatedly ask respondents during the pre-election period (like Paul Felix Lazarsfeld in his research in 1940s called "The Peoples' Choice"). I can follow up changes in their political preferences, reactions to the campaigning etc.

In sociological literature we can find an enormous amount of systematizations of kinds or types of sociological researches. For example, J. Wiatr (1964) presents another typology:

#### 1. Descriptive (diagnostic) research

He defines it a little differently from A. Hirner. He puts there mainly monographic research. These are those that answer the question of how some part of social reality looks like. If that part is, for example, a territorial community, research will examine its composition, institutions in its area and they work in it and social processes taking place there. Public opinion polls and analysis of opinions and attitudes of the given community in a relationship to the examined problem are as well part of a descriptive research.

#### 2. Research focused on verifying hypotheses

These researches are focused on looking for connections among phenomena (variables). The subject of the research must be well known for example, due to results from previous descriptive researches.

#### 3. Research elaborating theories

Aim of the research is to find out whether particular theory it is adequate for explanation of the given part of social reality, what are the conditions and limits of its validity in practice. Sometimes the results of several empirical studies focused on one problem are considered as well as elaborating theories in case they generalize particular research results and search for their connections.

This division reflects the initial intention in research planning. Otherwise, it is true that descriptive and verification researches can contribute to the elaboration of the theory too.

# **Stages of Research and Research Project**

Implementation of social-science research can be divided into several phases. In the literature, these stages are classified differently, which is the cause of different level of concretisation – some authors prefer to go more to detail than others, or it depends on different key point that serves for designation and division of one stage from another. In general, implementation of the research can be divided into three basic phases:

- **1. preparatory phase** (all steps that a researcher does before entering the field and before data collection)
- 2. implementation phase (data collection)
- 3. data analysis and interpretation

# Project of the Research

The preparatory phase, in a standard quantitative research, includes elaboration of the research project. The important thing about the research project is that it should be made in a way so that anyone would be able to proceed according to it and carry out the same (or very similar) research following its instructions. At the same time, it is necessary to say that the need to prepare a research project in written form is not only an academic requirement invented by teachers to keep students busy when teaching methods of social-science research, but it is a need that arises from research practice. For example, it is necessary for an agreement between a client and an organizer of the research in a commercial field. In academic field, the finances for research are often obtained from different grants and grant agencies demand the well prepared research project when making the decision whether to support particular research or not. Moreover, the research is often carried out by a team of people that need to bring their ideas into accord and unify the conception, content and form of the research. In all those situations it is required to "put down" the basic features of the research. It means to prepare a more or less formalised and a more or less specific and detailed project (outline) of the research.

Thus, every research action is preceded by the preparatory stage. During it the researcher prepares the project of the planned research. To have a better idea about the needs and tasks of the research project we draw an analogy with the project of the construction activity. If we do not build a sandcastle but a building, that must be placed somewhere, has its purpose and certain technical and aesthetic qualities, it has to be designed and projected. When designing, it is possible to eliminate imperfections and mutually harmonize different stages and changes in them. The project is a material that can be independently evaluated by different specialists. They can suggest beneficial changes and increase the probability of successful realization. Even the project of social science research is an independent material that can be evaluated from different points of view in order to harmonize all stages of the research process. Harmonization of planned stages (they mustn't themselves become problematic or exclusive of one another) is the practical aim of the process of creation of a research project. If needed, the project is used to carry out the research later again. It is as well a prerequisite for the comparability of those researches.

Based on our research and pedagogical experience and knowledge of literature in the field, we suggest that an adequate research project consists of these stages:

- 1. Formulation of the research aim
- 2. Selection of the research subject
- **3.** Hypothesis formulation based on the analysis of the given problem
- **4. Choice of the relevant research technique/method.** (Kusý, Vavrinčíková 2009: 27)

As we have already mentioned, the research project can be planned in various situations (for a client, a grant commission, etc.). Now, we will follow up the situation that is familiar to us and the most probable for the readers of our publication - planning of the research project as a part of an academic paper - for example, the final thesis consisting of theoretical and empirical parts. A theoretical part of the paper can serve as an analysis of the given problem and the subsequent research project can refer to it. In the real world, problem analysis follows after the determination of the research problem. University students choose the subject of their paper from the offer presented by teachers and they can take on the research subject in it.

#### 1. Research Subject

The first stage in the research project is formulation of the research subject. If there is no subject the research does not have any sense. The research subject is any problem (practical or theoretical) that we need to examine or explain. It is evident that we cannot explain such complex social problems like unemployment, poverty or drug addiction in only one research. We can consider formulation of the subject incorrect if the formulation, for example, would be: "Unemployment in Slovakia". It must be a research problem, not generally social - it must be a problem that is specific enough and can be examined in the research. We can reach this by place and time restriction of the given problem. Problems that can be examined in the best way are those that students come across when writing the theoretical part of their papers.

Unemployment can be, for example, place restricted to Banská Bystrica and time restricted from 2010 to 2015. When the research subject is then formulated like "Development of Unemployment in Banská Bystrica from 2010 to 2015" it is clear that students have an exact idea about what he is going to be investigated and are able to imagine how the investigation of this problem will look like. We can in this case suppose that it will be a research of changes in the extent and intensity of unemployment in this specific case based on the analysis of statistical data. However, the results of that research can be only freely related to specific problems of social and economic policy. Even if an absolvent exceeds the limits of a simple description we can suppose that it will only proffer suggestions in some fields. So it is possible to doubt the effectiveness and meaning of the research planned in this way.

However, we want to go further and we encourage the student to look for the research subject in another place. A prerequisite for a higher (even in the final thesis) level of professional approach to the selected topic is formulation of the research subject as the problem that lies in <u>the</u> <u>relation of two variables</u>. Their possible connection can students reveal while studying theoretical background of the problem. If we would continue in our example following the recommendation to formulate the problem in terms of relation of two variables, the subject of our research could be "The Impact of Long-Term Unemployment on Young People and their Political Orientation". The political orientation is a dependent variable in this relationship, and long-term unemployment is an independent variable – the one that makes changes in the dependent one.

Let's continue in planning the research project that will further elaborate this relation. In this case we could use the questionnaire or interview as a main method of data collection. We could determine the representative research sample by at least two features (the age of the respondents and the duration of their unemployment). The necessity of representativeness is more urgent in case of questionnaire; in the case of the interview the demands on the representativeness are lower because its results are of qualitative, not quantitative importance.

It would be sufficient to carry out the research on another sample (for example, on the sample of young employed people or people of a higher age category in long-term unemployment) that would serve as a comparative sample. It is a sample that would allow us to say with certainty how long-term unemployment can influence the taking of drugs. If its influence would not be proved it is necessary to find other influences.

The research coming out from an attentively formulated research subject would give the absolent satisfaction for his efforts in the form of professionally interesting and in practically useful findings. The important thing is the depth of knowledge connected with the simplicity of result interpretation. In this case the absolvent can feel content with his/her thesis. This kind of thesis is easy to defend because it is "about something", it directs the student towards important interpretative possibilities. We will mention them later. It is important to say that in interpretation of research results, it is important to strictly follow up the facts (research results) in order to avoid the temptation to just verify what has been expected from the beginning.

# 2. Research Object

The second stage of the research project is the selection of sample or object. The idea of research object is often implicitly indicated in the "research subject", but has to be elaborated explicitly later on. The answer to the question "who (or what)?" will be the object of the research is very important for the final results of the research. The selection of the object should (usually) focus on such groups of people or other information sources that can submit (or provide) relevant information on the given subject of research.

The object selection directly relates to the subject of research. They should not be in a conflict in any stage of the research project. Particular items of the research project must create a harmonious unit because only in this case can research be carried out successfully. If one of the items is changed than all other parts of the project must be rethought.

For example, in research of youth unemployment and its impact on political orientations, only certain groups of people and only certain sources of information can be useful. In our case that can be the unemployed themselves but also their relatives or household members or neighbours and acquaintances. Even former long-term unemployed can recall this period – currently middle-aged people – and give the information about changes of their political orientation before and now. As we can see, there can be a lot of appropriate research objects, it is necessary just to choose, ideally, in such a way that the selection would help to fulfil the intentions of the research.

Analogical to the term "research object" are terms "representative sample" or "research sample". The important thing is that it is a limited sample – the most often inhabitants, respondents. From this fact a serious consequence for the character of the obtained results arises, as we can see in the work of M. Disman (2002): reduced analysis of the reality leads to the statements of probable character. On the basis of the research results we try to deduce the conclusions valid for complete sample the unit that we call the "basic population". This "basic population" is created by all units for which our conclusions should be, according to our expectations, valid. The aim of the research object selection is that the representative sample should as much as possible represent the basic population.

#### 3. Hypotheses Formulation

Let's say again that we analyse problems of quantitative research. It uses deductive reasoning. It is based on the theory or general problem formulation. The problem is reflected in hypotheses (statements that can be verified) about the relations among variables. Consequently, we collect data that verify the validity of hypotheses. The contrary of this is inductive reasoning - here we begin with data collection and observation and we
seek for certain regularities. If we find them, we convey provisional conclusions. We verify these by further observation and in the case of a success a new theory may be a result of it. This procedure was necessary in the positivistic period in the development of sociology when sociological theory was only slowly created.

In general it is possible to say that we meet three types of hypotheses in sociological research. The first type is called <u>initial hypotheses</u> – they are a part of initial research stage. We can say that it is a part of research problem formulation or they arise in the preliminary theoretical analysis of the examined problem. The second type is the <u>working hypotheses</u> that present formulation of the relation between two or more variables. They direct the orientation of the whole research. In our research we try to verify or disprove these hypotheses. Finally, there are <u>statistical</u> <u>hypotheses</u>, i.e. those that are directly statistically tested and are based on working hypotheses.

Formulation of (working) hypotheses is a necessary stage in the cognitive process even if it may seem boring and unnecessary to a nonexpert. The importance of hypotheses lies in the fact that they test whether the research is possible at all. They test whether problems we are interested in are possible to express as relations among variables that we observe and measure. Another purpose of hypotheses is that they are a tool for optimization of the inevitable information reduction - we examine only that problem which is analysable. If we proceeded reversely and at first we formulated questions that seem interesting to us, we could find out that observed results are not applicable. And thirdly, hypotheses are a logical part of the research project: they connect the object and the aim of the research, they prefigure the research technique selection and technical demands on the data processing. They are primary in the estimation of the research extent - even a seemingly simple problem can grow up into many necessary variables, and hypotheses guard their manageable number.

"Hypotheses" in this part are "working hypotheses"<sup>3</sup>. We paraphrase a text by Miroslav Disman about this topic when we define them as a tool

<sup>&</sup>lt;sup>3</sup> Working hypotheses are based on operative definitions. Working hypothesis must be empirically verifiable. If we theoretically suppose that the level of likeness to a teacher influences the level of discipline in a classroom, we have to put both variables into a form that we are able to measure. In the given example it is the number of pupils that like the

with which we translate theoretical problem into the research language. Quantitative research uses deductive reasoning which means that it is based on the theory or general problem formulation.<sup>4</sup>The problem (different aspects of social reality) consists of (is reflected into) hypotheses (statements that can be verified by specific tools and by specifically projected research). (Disman, 1993: 78) Consequently, we collect data that verify the validity of the hypotheses.

"If we are to explain some fact and if we do not find a verified statement among recognized theses, we take some new thesis into consideration. We do not know whether it is true or false and we examine it in order to verify it. We take this kind of not verified thesis into consideration in the course of our attempts to explain some fact. This thesis that undergoes the verifying procedure is usually called a hypothesis." – Polish philosopher Kazimierz Ajdukiewicz (1953) wrote.

A working hypothesis is under the best conditions formulated like a statement that presupposes relationship between two or among more variables. Variables must be definable in an operative way, i.e. in a way so that it is possible to work with them in our research - they must be measurable and observable. A working hypothesis testifies to how we will measure the expected relation. (Disman, 1993: 79)

Every indicator that we measure must have its own working hypothesis. "In quantitative research, we can examine only those problems that can be translated into the language of the working hypotheses, i. e. the kind of problems that can be expressed as relations among variables that have valid operative definitions." (ibid: 85)

We will describe the role of a working hypothesis in the research project and the research implementation itself together with Miroslav Disman in this way:

teacher and the number of bad behaviour marks. Working hypothesis will then be: The more pupils like the teacher, the smaller number of bad behaviour numbers. It is verifiable because we can detect and measure this kind of expression of likeness and discipline. Even then, one problem occurs - the problem of selection of adequate indicators for operational definitions. In this case, it is questionable to which extent the bad behaviour mark represents violation of discipline.

<sup>&</sup>lt;sup>4</sup> The contrary of this is inductive reasoning - here we begin the research and observation and we seek for certain regularities. If we find them, we convey provisional conclusions. We verify these by further observation and in the case of success a new theory may be a result of it. This procedure was necessary in the positivistic period in the development of sociology.

1. If we are not able to formulate the working hypothesis we are not able to carry out the research. Hypotheses formulation tests whether the research is possible at all and whether problems we are interested in, are possible to be expressed as relations among variables that we can observe and measure.

2. Working hypotheses help us to optimally reduce information (we have to be relatively sure that we have chosen the most important variables and that we do not collect information that is useless) and to estimate the extent of the research. We examine only that which is possible to analyse in the solution of the problem and what is related to it. If we proceeded reversely and at first we formulated questions that seem interesting to us, we can be satisfied with ourselves as great researchers but when analysing we would find out that observed results are not applicable.

3. Working hypotheses contain the basic information for selection of a research technique (according to the expected demands in data collection and processing).

4. Hypotheses formulation is not an opportunity for the student to show off with the right estimation of the result. However, it is a step in which a student expresses why he/she is interested in the problem and why he considers the current level of knowledge about the problem to be insufficient or outdated.

5. Hypotheses must be in accordance with all preceding steps of the research project that emerge from the "theoretical part" of the research. They represent the exact formulation of such aspects that the researcher focuses on in his own research. Those are the aspects of the research subject. It is necessary to verify - confirm or disprove - the validity of their formulation in the research (we "test" the hypothesis). Because of this, these hypotheses must be in a form of the statements and it is not important whether they are formulated in a positive or in a negative way.

6. Formulation of (working) hypotheses is a necessary stage even if it may seem unnecessary.

Let us see the example of the positive and negative hypothesis formulation mentioned in point 5. A hypothesis that further develops the possible influence of unemployment on social behaviour of an individual can, for example more susceptible to become drug addicted than employed ones of the same age."

If the research claims that this correlation exists, we would obtain the same result with the negative formulation of the hypothesis:

"Long-term unemployed men **are not** susceptible (are not equally susceptible) to become drug addicted than employed ones of the same age."

Even in this case, the hypothesis directs the research towards the relation of the unemployment and drug addiction. It would just not confirm this hypothesis (it is not true that they are equally susceptible). The result is the same as when verifying the first hypothesis, the result is that the relationship exists.

It is not possible to state the minimal or maximal number of hypotheses in the research. It can be only one or several tens of hypotheses. It is necessary for the student to check the maximum by the basic demand - he must verify every hypothesis in his research. One hypothesis does not mean one question. One hypothesis can be verified by several questions, and what is more, this link between the number of hypotheses and the number of questions is welcomed because more questions help to verify the hypothesis more precisely.

We would like to point out that hypotheses should not be formulated in a way to be confirmed. Whether the hypothesis is confirmed or not the result is equally valuable. For example, one of the parts of Einsteinian Physics emerges from a non-verified hypothesis about the existence of ether in the space. Initially, it seemed that if ether (a substance with certain properties) existed in space we would be able to explain many recorded phenomena. The research that was aimed at verification of the existence of ether did not confirm it and because of this Einstein searched for another explanation and he found it – a new theory.

We point out that hypothesis are not a part of fortune-telling. In spite of this, many students defended their theses with persuasion that if they guessed the right result of their research, their whole empirical part would be successful. This unfounded attitude considerably helps to accept hypotheses formulated in this way (we will mention an absurd case): E.g.: "56.4% of men and 47.8% of women that are long-term unemployed are drug addicted."

Even if the research confirms this percentage (we can claim that the student formulated this hypothesis from the point of view of logical and time sequence in reverse - on the basis of the obtained results), the result is of small cognitive value, even if the student may have "guessed". The first problem arises from the fact that if the reached value for men is 56.3% or 56.5% it is necessary to say that the hypothesis was not verified. We have thus said that it does not matter, but on the example like this, we can demonstrate that it is necessary to interpret very similar results in reverse when the hypothesis is formulated the wrong way. The second problem emerges in that the right guess of some number does not mean the confirmation of the association we seek to achieve.

#### 4. Selection of research technique (method)

At the beginning of this chapter we would like to explain the reason why the term "method" in the title is in brackets. The reason for this is that we prefer the term "techniques" to designate data collection. In our textbook we incline to those conceptions that use the term "methods" to describe other procedures or science research tools. At first, we would like to point out the variety in which are both terms used in sociological literature.

Let us begin, for example, with the division of methods and techniques in sociological research according to Czech sociologist Václav Lamser (1966: 33-38). Methods are divided into general and specific. General methods are: comparison, analogy, dialectic method (unification of the main logical processes, analyses and syntheses, induction and deduction, quantitative and qualitative procedures), and Mill's methods of inductive thinking, direct method of agreement, method of difference, method of residues and method of united variations. Specific methods, on the other hand, are methods of specific sciences and thus also social science methods. Lamser puts social science methods into the group that he classifies as special sociological methods: Weber's method of ideal types, Le Play's method, sociometric method, Simmel's method and some others. Václav Lamser (ibid), apart from the term "method", defines also the term "**technique**" and he differentiates it from the terms "methods" and "procedures." According to him technique is a special operation within some stage of cognitive process in the research. It is the term that is logically subordinated to the term "procedure" because procedure can include several techniques. Technique is connected not only to observation, but also to data processing. For example:

- questionnaire technique
- technique of standardized (structural) interview
- technique of content analysis and
- technique of analysis of variance

Implementation of the technique has, according to Lamser, a character of routine operatin. It requires professional training, experience, and practical training (e.g. interview, or direct observation). Further, he speaks about the techniques of observation (while this term is used with the meaning of "investigation"). There are three observation techniques: direct observation, questionnaire and interview. The use of different kinds of documents could be added. (ibid: 88)

Miroslav Disman (1993: 123) does not exceed the listing of basic kinds of techniques (standardized or non-standardized interview, group interview, questionnaire, observation, sociometric techniques, semantic differential scale, secondary analysis, participant observation, etc.). After listing many specific techniques he concludes with the statement that fortunately all those are applications of the following basic techniques:

- direct observation
- interview
- questionnaire
- document analysis.

In the research project, the student (the researcher) chooses those methods and techniques, he supposes to be the most sufficient for fulfilling the research aim, and for verifying the hypotheses. In the selection of methods and techniques it is necessary to take into consideration the practical aspect – their feasibility. In order to confirm or disprove the hypothesis it is possible to combine several methods and techniques. For example, I can use already existing statistical data from other researches (method of secondary data analysis) and at the same time I can create a questionnaire that will be used to complete the information needed to verify the hypothesis.

# 5. Operationalization: Defining the Indicators of Social Phenomena

The research problem in the case of standard quantitative sociological research can be formulated as a relationship of two or more **variables**. To formulate the research problem then means to determine the research subject and object, in a way mentioned in the previous part.

The term "variable" is adopted from natural sciences. Students probably know it from mathematics. It designates some reality/entity that can take on different values. The relationship of two variables is then the one in which the values of one variable change according to the changes of values acquired by the second variable. One variable is **independent** (that influences the changes of the other variable) and another one is a **dependent variable** (that changes its values according to the values acquired by the independent variable). For example, in mathematics we can speak about direct correlation in situation when the independent variable (x) is increased by one unit and consequently the dependent variable (y) is equally increased by one unit.

We already know that we do not meet with this kind of relation and such obvious determinism in social sciences. Because of this, the links between dependent and independent variables will not have the form of a direct or inverse correlation, but their relation will only be of probable character. Their connection will be functional, not causal. If the dependent variable is age and the independent one is the level of conservatism (expressed, for example, by total score on some conservatism scale) we can observe that with increasing age the level of conservatism will increase, too, however not directly, one by one.

The term "variable" determines those realities, qualities and phenomena in social science research that are of variable and complex character, at the same time they show the kind of regularities or rules that we want to observe in our research. The basic creative process in quantitative research is formulation of the research problem into a relation of two or more variables. To get from this still too general level to the level of specific, empirically measurable **signs and indicators** is one of the most important tasks in a whole research process. This cognitive operation is sometimes called: **operationalization**<sup>5</sup>.

The key question of operationalization is: how to transfer, to project the problem that is in the general (theoretical) level into the empirical level. It should be done in a way so that the collection of empirical data would be possible and at the same time the quality and structure of the given social phenomena would be preserved. We have mentioned above that in the case of social science research a kind of reduction of the examined phenomena is necessary. The important thing is that this reduction is conscious and meaningful and emerging from what the researcher wants and is able to find out.

Example: A certain research team examines the problem formulated in a following way: Belonging of an individual to the social networks and their influence on his/her carrier success. The research team is going to investigate a well-known and verified finding that social networks (relationships and contacts that an individual creates and builds with the others during his life) are an important part of social capital that can be useful for an individual and they can bring him, besides other things, professional success. In the above formulation, the problem is still laid on the theoretical level and empirical research is not yet possible. To examine the problem empirically it is necessary to put it on the lower level of generality - the level of its signs and indicators, because for example, if the researcher wants to use the technique of questionnaire, he/she cannot ask: "Can you tell me, please, what your belonging to the social networks is?" It is evident that the respondents would not understand the question like this. In the case of the dependent variable - professional success, it is evident that the most of respondents would understand the question, but every one of them would understand it in another way. One could speak about the success because he finally found a permanent job after a long

<sup>&</sup>lt;sup>5</sup> The term invented and used by American sociologist Paul Felix Lazarsfeld (1901 – 1976)

time of part-time jobs and unemployment, the other would be dissatisfied because he is waiting ages to be promoted to managing director and he is still the director of only one section. The answers would not be compatible because the denominator would not be common and answers would not correspond with the needs of the researcher. The researcher must, in his own interest, state what he/she understands under the belonging to the social networks or professional success and how he will examine and measure these phenomena. And thus, in what way he will **analyse** that phenomenon. He must try to project that phenomenon on the level of particular signs and indicators that are clear and unambiguous to operationalize them, so that for example in the case of a questionnaire he/she id able to ask concrete and precise questions that are easily answerable.

It is necessary to point out that operationalization does not mean linguistic analysis of the term, thus not its definition. It is not important to define what the profession or professional success is, but to define the indicators we will use in the examination of the problem in this specific research. Of course, we can use any ordinary non-professional definition of this term or base it on some theory, however not only linguistic and semantic but also ontological projection of the term or variable is important in this case. It is important to mind the correlation of the terms with reality they represent.

Because of this Hirner (In Schenk, 2003), in comparison to Lazardfeld's conception of the operationalization, develops the conception of social reality identification to name the process of discovering relevant signs and variants of the given phenomenon. Hirner says that while the subject of operationalization is the term, the subject of identification is social reality. The relation of the term and social reality (between the term and its determinant) is not evident. Not only in social sciences has it been proven that many (even scientific) terms do not have a real content or do not correspond to any reality. A classic example is the term world "ether" that was used in physics particularly at the end of the 19th century. This link is more difficult in social sciences because the language is often contaminated by terms from ordinary language, by many synonyms, etc. As a result, it is not possible to start scientific research from the term and the attempt to define it. The first presupposition is the examination of the connection between the term and reality that denotes that problem with elimination of all cases when the term relates to an empty set or is a synonym of another term.

In operationalization sometimes we work with so called terms with direct empirical covering – and the operationalization is not necessary in this case. The variable can be, for example, age, gender or so. Most of the variables we work in our research must be further operationalized, i.e. analyse. We can indicate it by the following diagram:



(D1 – 3 are dimensions of the phenomenon, I1-3 are indicators, x1-3 determine variations of the indicator/sign)

Let us return to our example. The research team wants to examine professional success, or professional life. Let us say that we want to take into consideration both essential aspects – objective (with clearly determined criteria what the professional success is) and subjective (emerging from the recognition of what the professional success or failure is, to a great extent it is determined by subjective perception of this problem). So the first dimension (D1) of this phenomenon – professional success – will be created by the objective aspect and the second dimension (D2) by the subjective aspect. It is necessary to find a relevant number of indicators in each dimension that will be used in the examination of that phenomenon. In other words – as a researcher I will ask myself a question: "What indicates professional success (in its objective dimension)?" The indicator can be the level of income, job position etc. The indicator of a subjective dimension can be, for example, the degree of contentment with the job. It is on the researcher whether he/she chooses the indicators that are relevant for the research. This is the core of his/her work.

In sociology and social sciences we meet with multidimensional phenomena that are comprehensively and qualitatively complex. Social phenomenon as a very complex structure and cannot be ever described in its integrity and complexity. It is not possible, and at the same time, is not needed at all. The process of phenomena analysis is always about realization of certain **conscious simplifications**. It is always the **reduction** with the aim to project and preserve basic aspects of the phenomenon (basic in consideration of the research aim and of the relation to reality) and to abandon those aspects that are not necessary from this point of view so that the mentioned phenomenon is empirically investigable. This simplification need to be conscious and reasoned. (For example, we do not have to take the subjective aspect of a professional aspect into consideration - with the fact that it is not essential for us now. In another case, however, it could not be possible.)

Surynek, Komárková, and Kašparová (2001) present the example of operationalization as a succession of two processes, taking into consideration the fact that besides the ontologically correct operationalization, the researcher must carry out the conscious reduction on the essential units of the phenomenon that will be selected into the research. Two stages of operationalization are according to these authors:

- **decomposition** (of the term on the nearest lowest level of generality)
- **selection** (of units of lower level of generalization, and thus the selection of basic aspects of the phenomenon)

#### Indicators

We are not primarily interested in indicators of social science research as in the basic units of language. Those are the realities that on the basis of their genetic, structural or functional connection with other realities say something about themselves. We thus delimit them from other realities like their specificity, attribute, unit. The indicator is not thus understood as a symbol, we are here interested in the content of phenomena – what they represent in social reality. Every social reality has an enormous amount of expressions, features, and the researcher expresses his creativity in that he selects them for the research. Every indicator occurs in several variants - these are qualitatively or even quantitatively limited (from some minimum to some maximum). Indicators can be understood as derived features. To be able to represent the empirical character, the indicator must have correctly chosen variants (corresponding to or covering the empirical reality).

On the basis of correlations among variants of an indicator or variable, we distinguish following types of indicators:

- a) Qualitative nominal
- b) Quantitative
  - ordinal
  - cardinal

A. **Nominal indicator** – according to its values the unit is divided into parts that we consider not being in any other connection (for example arrangement or distance), they must be mutually exclusive.

- particular example is dichotomous nominal indicator which has only two values (for example, gender, or other indicators with only two variants, such as: professional - amateur, correct incorrect, yes - no)
- particular case is so called simple nominal feature distinguishing identifier of each object that is not a part of analysis (e.g. birth certificate number, questionnaire number, etc.)

For variants of nominal features the following conditions must be fulfilled:

- unambiguity (two values must not be assigned to one unit)
- existence (must cover all possibilities)
- differentiability (determination of particular variants must be different and mutually exclusive)

Operations that can be realized with qualitative indicators: determine the mode, ratio of numerous representations of variants in the data set in absolute values, in percentages or index numbers (e.g. how many people as carriers of the feature are there in 1000 inhabitants, and things like that)

B. **Ordinal indicator** – its variants are arranged in order

We distinguish between comparative and classifying indicators. Comparatives do not have any ordinal scale and particular cases are arranged in order by comparison to what extent is the feature present in them – e.g. I rank the members of the group according to their popularity or their real income. Classifying indicators are those in which I set the order of variant feature according to the scale determined in advance – e.g. satisfaction.

For example: Respondents in marketing research will be asked to rank washing powders according to their popularity. This is comparative indicator. In another research the respondents will be asked to express the degree of their satisfaction with each type of washing powder on a 5 degree scale. This is classifying indicator.

Within mathematical and statistical data processing of ordinal indicators, following operations can be used: addition, subtraction, median and quantiles, but it is not possible to multiply, divide or to use arithmetic means.

C. **Cardinal indicator** – numeric variable, its properties are numbers, quantity, time, proportion or a degree

We distinguish interval and proportional cardinal indicators. Interval indicators are ranked in order, but we are able to measure the distance among them. Proportional indicators are expressed in numeric values of the system in which we recognize the zero value. It is possible to express the value of one variant as a proportion of the second one. For example, salary, age, metres squared, number of appliances, time spent by housework, or so.

With cardinal indicators, the following statistical operations are possible: arithmetic means, standard deviations and coefficient of variance, and even correlation coefficient.

Many other operations and transformations are carried out with indicators, mainly in the phase of data analysis. For example, it is possible to change cardinal indicator into nominal indicator (by classification of its values into categories that will further act as variants of nominal indicator). It is particularly necessary when I want the connection between cardinal and nominal variable. It is true that I can make nominal variable from cardinal or ordinal without any difficulties, reverse procedure is problematic but not completely excluded, however, it is used only in specific cases/situations.

## 6. Selection of Research Sample

As we have already pointed out, in the case of quantitative social research we are able to examine the whole population only in exceptional cases. In other cases we have to strive for the selection of a suitable research sample. Suitability means that we try to be as close as possible to the likelihood of generalization of our results on the entire population (it can be, for example, all the inhabitants of a town, all the unemployed, all students of social sciences, and things like that). Our task in the sample selection will be to create presuppositions for the generalisation of our findings based on the cases from sample on a whole population. At the same time we have to be aware of the fact that this process (that is necessary due to the practical reasons – time and price) further deepens the **probable character** of our findings and it makes us be more careful in their interpretation and in the formulation of conclusions.

What is particularly important is our answer to the question: how should we proceed in a particular case in the sample selection to reach a satisfactory value in our conclusions given the fact that we have formulated them on the basis of the results, with the highest probability very close to those from the whole population.

It is not always true that the larger the sample the better. In achieving our aim we meet with time and financial restrictions. The initial question for us should be <u>whether the sample is well representing</u> the population with the least consumption of time and money. A structural sample (selection set) must in a special and suitable way imitate the structure of the whole population (basic set).

The turning year concerning the tradition of representative researches was 1936. There were two candidates in the presidential elections in the United States: F.D. Roosevelt and M. Landon. On the one hand, there was **George Gallup** with his colleagues that carried out the electoral prognoses on the sample of 2000 respondents using their newly developed method – and as it turned out later, they were successful. On the other hand, there was the redaction of Literary Digest magazine that carried out its own survey while it sent out more than 10 million survey cards to addresses chosen from phone books by mail. It is necessary to mention that surveys in the USA have

a long tradition, dating back to the 19th century. The redaction processed more than two million of the returned survey cards. However, the redaction finally made a mistake in prediction of the winner when they preferred Roosevelt's rival.

How it is possible that Gallup with his team predicted the results more precisely in the sample of "only" 2000 respondents and the redaction of the Literary Digest magazine was wrong although their sample was more than a thousand times larger? The answer is that the lack of systematic procedure of the reduction was the use of <u>unrepresentative</u> survey based on the selection from phone books. The owners of a phone in those times were more often rich Americans that more likely supported M. Landon. The survey was not representative because in the sample there were not included some important groups of voters preferring F.D. Roosevelt.

George Gallup is thus considered to be the founder of the branch of public opinion polls and his methodology, later even more improved, quickly spread around the world and newly founded institutions in different countries started with public opinion polls based on a representative sample. Even in Czechoslovakia after World War II., new institutions for public opinion polls were created - at first in Prague in 1946 and a year later in Bratislava.

What is the methodology of the representative research sample selection and how does one get that sample?

We have already mentioned that the research sample should reflect the qualities of the basic population. Various types of random selection serve for this. One of them is a **quota sampling** that "i<u>mitates well-known</u> <u>qualities of population in the structure of the sample.</u>" Quota sampling means that we set up in a sample a proportion of respondents with certain characteristics according to the distribution of those characteristics in the whole population. If, for example, our sample consists of 1200 respondents, approximately half of them should be men and another half women – because this is approximately their distribution in a whole population. There are populations in which the ratio of men and women can be different – for example, in the population of the retired there will be more women and if we want to do our survey, it is necessary to take it into consideration in our sample and to change the proportion, for example to the 60% of women and 40% of men. However, the technique of quota sampling has some unpleasant characteristics. They are connected to the fact that it is possible to consciously capture only those characteristics of the population that are familiar to us. So we have to base it on some statistics that inform us about the distribution of these characteristics in the target population. Sometimes those data are not available. What if the key factors that differentiate our sample and that influence examined phenomena are those characteristics that are not reflected in statistics?

Disman (1993) points out that quota sampling can be applied only in a population that we are well informed about. For example, we are not able to propose the representative sample for the population of racketeers because we are not informed about their structure, demographical characteristics, etc. Problems can occur even if we take too many characteristics of the given population into consideration. Final requirements on the structure of research sample would be too complicated to fulfil, because all characteristics that we take into account occurs in reality in many combinations.

However, there is a technique that eliminates the disadvantages of quota selection. It is so called **random probability sampling**. A randomly selected sample represents, except the well-known properties, even unknown ones. Disman (ibid) introduces an example: I have colourful marbles in a bag. Let us say that I do not know how many colours there are inside of it, and I do not know their structure. If I randomly pick out the marbles from the bag, in certain amount high enough, there is a high probability that among the marbles I have chosen there are all the colours that are in the bag, even in a similar structure. What is important, it is "only" to state as already mentioned "a number high enough".<sup>6</sup>

The most important thing in random selection is to provide that every member of a population has an equal chance to get into the sample, and thus to eliminate the factor that can influence the results. However, it is not easy to provide it. How to carry out the random sampling?

What will happen if I run out into the street and I randomly ask passers-by? Will it be a well carried out random selection? The answer is: it does not have to be this way. For example, it depends upon the locality in which I run out into the street because the structure of inhabitants in

<sup>&</sup>lt;sup>6</sup> See also the chapter Large Scale phenomena and Law of Large Numbers.

one quarter does not have to be identical with the structure of inhabitant from the whole town or (after all) country. So I have to eliminate the factor of the unrepresentative area. It would also depend on the time I run out into the street. If it were in the morning of a working day, it is probable that in the town there will be people that are not active in terms of work – for example, the retired, mothers on maternity leave etc. It follows that I eliminate or take into consideration the time factor. Another problem can be in connection with a certain kind of stereotyped behaviour of the interviewer in case he/she chooses (unconsciously) a certain type of people that he/she likes or presupposes that they are more willing to take part in the survey.

Let us try it in another way. We will not run out in the street but we will ask interviewer to knock, for example, on the every fourth door in given quarters. However a situation may occur that there are just four flats on one floor in a certain kind of block of flats so the interviewer always knocks on the right door from the elevator. Let's imagine that on that side of the block of flats are four room flats and it is probable that a certain type of inhabitants lives there. It is very probable that there will not be single person households, nor socially deprived groups living there so they will not be included in the sample.

We can conclude that even if the random sampling is a very good technique of representative research itself, it is not quite easy to carry it out - to ensure the real randomness. If the random sampling is well provided, then we can say that with the increasing size of the sample the difference between the structure of population and the sample decreases. We try to get the largest possible sample but not at the cost of severe disruption of the random selection. The size of the sample has its limits, behind which the sample is saturated – i. e. new cases will not bring any new information, nor will they increase representativeness. On the contrary, the costs and the effort for this shift is too large and ineffective.

When determining the extent (size) of the representative sample it is important to search for the optimal limit between the effect of the law of large numbers and the extent of representativeness. This limit is not always the same; it changes its position according to the character of the examined phenomenon or problem. The extent of the representative sample is determined by the homogeneity or heterogeneity of the basic examined group (whole population) and the possibility of a sociologist to take some number of characteristics into consideration when choosing the research sample. The rule that the greater the population the largest the sample is not valid. Heterogeneity – variety in the examined population, is particularly determining the size of the sample. Of course, it can grow with the increasing population, but it does not have to. For example, representative surveys in the USA usually work with the sample of 2000 respondents and in Slovakia it is a little over a 1000 of respondents. Yet, the number of inhabitants in the USA in not only twice as high as in Slovakia, but it is more than fifty times higher. The representative sample in USA is larger, but not in direct proportion to the number of basic population. The reason for the larger size of the sample in the USA is not only the larger size of population itself, but the need to take more factors (characteristics of that population) into consideration when drawing up the research sample.

In random sampling, we are able to estimate the sampling error – and thus determine in what way our sample is similar to the population (considering the characteristics we observe) – in most cases we determine it after the data collection in statistical analysis because for the estimation of this error we need to know the dispersion of the given characteristic in the population.

There are different types of random sampling. For example, so called <u>simple random sampling</u> that is more often based on some kinds of lists, or on selection by special computer software that generates respondents in a random way (e.g. according to the phone numbers). This way is often used by research agencies, for example when estimating the results of elections.

Another type of random sampling is <u>systematic random sampling</u> in which we determine certain criterion of respondents' selection - e.g. we will include the every fifth person from some list into the sample. We get this number when we divide the size of population by the size of the required sample. The condition is that the list cannot be systematically arranged without the awareness of the researcher. For example, the researchers in the USA had once decided to select the sample from the phone book randomly opened on certain page. Accidently, they opened it on a page under the letter M. Since there is a relatively strong population of Irish immigrants with their surnames beginning with Mc., the research captured the opinions of this minority in an increased level and thus he brought a sample error into the survey. Another type is <u>stratified random sampling</u>. Firstly, the population is divided into several homogenous groups according to certain criteria and then respondents are randomly selected from each of these groups. For example, there can be groups based on particular year of study at school, and then the random sampling is provided in each of them. <u>Multi-level random sampling</u> takes place when the groups are randomly selected (e.g. two or three from eight districts) and then also respondents are randomly selected from these groups.

Other types of selection are:

<u>Purposeful sampling</u>: sample selected by purpose does not enable generalization of results because it is not representative. The research concerns only (by researcher or by self-selection) selectively chosen individuals. For example, in a newspaper survey the selection of respondents is based on the decision of an individual to answer the questions published in the newspaper.

<u>The technique of "snowball sampling"</u> is rather a technique of identification of the non-registered population than a representative sampling. We use this technique in cases when we do not have good access to potential respondents – e.g. we want to investigate immigrants from Ukraine that work illegally in Slovakia. They are registered nowhere and we do not have another occasion only to find at first one of them and then get from him contacts to other ones and from those to others etc.. The size of that group increases – like a rolling snowball.

In general, it is necessary to remember that, even if we provide initial representativeness of the sample, many other problems can occur mainly those that are caused by the fact that not all respondents will answer all the questions or they will not answer any at all. It is the problem of returnability. In order to prevent low returnability from destroying the representativeness of the sample we, on the basis of our experience, a little bit overestimate the statistically satisfying number of respondents in Because of this we even the selection sample. check the representativeness of the sample in reverse, after data collection to take possible deviation from the structure of the whole population into account when analysing data and interpreting results.

## Methods and Techniques of Data Collection

As we have already mentioned, the term "method" is defined as a set of rules and operations that generally direct the scientific work, the way of processing of collected data and findings. We distinguish the following methods:

- a. **generally logical** these are used not only in science but also in ordinary life e.g. induction, deduction, analysis, synthesis, generalization, abstraction, analogy, modelling
- b. **generally scientific** these are used in all sciences, not only in social sciences, and they can be divided into:
  - A. empirical methods, e.g. observation, experiment, measuring, description
  - B. theoretical methods, e.g. idealization, formalization, logical method, historical method
  - C. methods of specific science these are, in our case, methods of empirical social science research focused on the procedure of data collection and processing

In the following text we use the term "technique" as a synonym of a term method in this last – narrowest sense – as a method of specific science. Techniques are defined as the way of obtaining and processing of empirical information. They are thus divided into techniques of data collection and techniques of data processing. We can point out that in a really responsible sociological research in which we want to obtain important and applicable information, only one technique is insufficient, and usually the application of several techniques in their combination is needed. If the finding is confirmed by several techniques, its value is higher. (Kusý, Vavrinčíková, 2009: 27-37)

As basic techniques of data collection, on the basis of professional literature:

- 1. **observation** we distinguish participating and non-participating
- 2. **interview** it can have a different level of formalization, it may be more or less standardized
- **3. questionnaire** we mention it as an independent technique because it is characterized by many particularities from the

point of view of data collection organization; active participation of the interviewer is limited to the distribution of questionnaires and possible (rather formal) help to the respondent. The questionnaire is usually sent out and returned by mail, or e-mail and thus without any personal contact with the respondent.

4. **documents analysis -** it is the analysis of written materials (letters, archival materials, official records, newspaper, advertisements, statistical data, architectonic projects, medical reports)

These techniques suit the examination of particular kinds of social phenomena and processes as well as researching intentions and possibilities. This means that we can speak about the unsuitability of particular techniques for the examination of certain phenomena and processes, or about the suitability and form of techniques used by the researcher - the student. The appropriateness of the use of techniques can be classified as very suitable, usable and non-usable, while the criteria for the consideration of techniques are:

- availability to direct observation
- the speed of changes in progress
- binding to an individual, group or structure
- bonding to social or substance and spatial phenomenon
- topicality

As an example we mention several evident incompatibilities of the technique with the examined phenomena:

**Observation** – it is not possible to use it in the examination of past phenomena and processes as well as those phenomena that are not available to direct observation because of some other reasons.

**Interview** -we cannot use it in the examination of the phenomena stemming from material works, and its use is as well very limited in cases of great groups, if the examined process is very fast or we have a lack of time.

**Questionnaire** – it is not suitable in the cases of phenomena that are derived from material works and in case of fast processes; unlike the interview, it is not appropriate when the sample is too small. In case of

final thesis, the criteria for the appropriate size of the sample are due to the practical reasons less strict.

**Study of documents** – we do not use it to investigate fast processes, and in current yet undocumented phenomena and processes.

Aside from the already mentioned techniques that we can find in practice in various forms and combinations, we recognize other techniques but they are of limited use. They are used only in certain fields of sociology, or in certain stages of sociological research, e.g.:

**Sociographic technique** – particularly used in the stage of preresearch while the examined phenomena must be clearly time and place integrated.

**Sociometric technique** – used in sociological researches of small groups, it contains a wide range of attractive procedures.

**Scaling technique** – used in the research of attitudes – with the aim to increase the exactness of analysis and reduction of the researcher's subjectivity

Every one of the mentioned techniques can be classified either to be mostly qualitative or mostly quantitative. The fundamental difference between them is:

a. Quantitative techniques consist of a large number of data that enable the application of various kinds of exact data processing – using methods of mathematics and statistics. To get closer to the exactness of natural sciences we have to pay the relatively high penalty of a certain kind of uncertainty connected with an extent to which the technique inserted between the researcher and the respondent influences the truth of the reflection of the examined reality. Quantitative analysis simplifies problems into a black and white vision of social reality because numerical data do not represent nuances in attitudes of respondents well enough. Their interpretation without practical experience with social reality, without the understanding of relations and differentiated ways of thinking to a large extent shifts and simplifies the final results. We can say that the same data provide space large enough for a different application of sociological imagination and interpretation of these data.

**b.** Qualitative techniques, on the other hand, do not pull out the investigated objects from their time and spatial frame and they do not work with schemes about the given phenomena or process created by the researcher in advance - they are of inductive character, thus they try to generalize particular cases they examine. The low possibility of generalization is one of the disadvantages of these techniques. The obtained data are not representative and they present a deviation from methodological equipment of highly formalized sciences. On the other hand, a question occurs to what extent is the representativeness of the target category of sociological research. Qualitative techniques enable expression of problems of social reality in a more differentiated and precise way, and with deeper understanding to the nature of a problem. It can be equally considered as a contribution to the truthfulness of the results. Nowadays, argument about inaccurate and unscientific rules of qualitative methods can hardly succeed because there are enough of elaborated procedures that enable us to use a computer - a symbol of exactness - although not for calculation and statistical operations with a large number of data, like in quantitative techniques. (Kusý, Vavrinčíková, 2009:37-38)

## Explorative Methods: Questionnaire and Interview

A human manifests himself not only by his actions and behaviour, but also by his way of thinking. We can say that in addition to the outer world and those manifestations visible from outside a person, each human has his inside world - the world of his consciousness. It is the world in which he creates his attitudes, maintains and restores his values, the world his interests and priorities reach out from. This world is invisible from the outside. The only way to peek inside it is to get information about it from the individual itself. His utterance can be of various forms - mimic, graphic, written, oral. We will be mainly interested in those written and oral because they are the basis for two particular techniques of sociological research - a questionnaire and an interview.

One important thing we have to keep in mind is that what is in the human consciousness and what the respondent expresses outwardly in his self-representation or self-evaluation (because this is done in the process of questioning) does not have to be manifested in his behaviour. Between words and actions there can be - and usually always is - some difference. If the research would be focused on the examination of presuppositions for some kind of behaviour in the future, it is highly probable, that answers and results will be different from the real behaviour of people when the given situation occurs. The respondent's utterance is not the final truth about him, his consciousness. It is evidence of his self-observation, his ability to evaluate and express himself, even about his willingness to utter in a correct way. These are thus certain symptomatic manifestations that uncover their true character as late as in confrontation with corresponding objective data.

Since we speak about the inner world of an individual, it would be appropriate if we were aware of certain general experiences concerning the psychology of his behaviour. As for the whole domain of social facts, so for the domain of these psychological aspects it is valid that they are permanent, they undergo the changes in time and they are different from the point of view of space. If we examine the permanence of human behaviour motives, or permanence of his values and interests, we have to be aware of the fact that it is only a temporal.

In general, it is possible to say that a sociologist usually does not have an aspiration to examine individual reasons of the respondent for a given answer. Even though it follows and transmits certain forms of inner logic or reasoning of answers in seeking for correlation relations (if the character of the research enables it). In spite of these circumstances that decrease the value of results obtained by the questionnaire or the interview, these techniques remain a favourite and are broadly used because it is better to have unprecise information then no information at all.

Imagine the second possibility that we have when examining human consciousness - we could determine motives of his behaviour according to his real behaviour. We should observe this behaviour in natural conditions and continue it long enough in order to interpret the observed behaviour in the most reliable way. This research would, however, be very time and personnel demanding, associated with many organizational difficulties. Because of this, the techniques of social-science research most often employed are the questionnaire and the interview, so called explorative methods. Explorative methods are those that as a key technique of obtaining data use oral or written questioning. We can say that the most important property of the question is that **it has to be related to the research subject** and thus to some hypotheses (it must be valid). In the set of questions the most important thing is that the questions should cover all indicators and all dimensions of the examined phenomena that were selected into the research.

Besides these primary ontological problems, the researcher must solve the whole group of practical problems. In the construction of questions he/she must take **the factor of respondent** into consideration and adjust the question. For example, it is important to take into account whether we use the expert surveys, or whether we examine the whole population including the young, the retired, people with different degrees of education and different specialization, in our research. It is necessary to adjust the form of a question in a way that would be understandable for every member of the sample, to take language context (slang) into consideration etc. The important thing is that both, the respondent and the researcher, should understand the subject sphere of questions in the same way.

**Reliability of the obtained information** to some extent depends upon the selection of terms that we use in questions and answers. If the terms are too vague, inaccurate or incomprehensible their interpretation is difficult. For example, the terms "a lot of", "bad", "poorly", "successfully" are together with their opposites unreliable for their different explanations in particular cases (50, for someone is not much, for the other, it is a lot). Then, we just work with the subjective meaning only of the examined phenomena for the particular respondent, but not the objective distribution of the frequency of the given phenomena.

The reliability of information is related to the way of asking a question - word order and intonation (You do not like your job, do you?). The right formulation is neutral (To what degree you are interested in your work?). We get very similar answers to stereotyped questions. What kind of a question that kind of an answer (How are you? Fine). If the answers, on the contrary, are not appropriate to the ordinary expressions, we put the answers into the respondent's mouth that he would not utter

himself. This is mainly related to written questioning but a similar error can occur in the interview.

The sincerity of the respondent (mainly in the interview) is negatively influenced in several ways and the contents of questions. Those are frontal questions - as though investigative (why?), or in other words demanding from the respondent the explanation. "Why do you not like the programme Professionals?" The answer is difficult even for an expertcritic. (Lazarsfeld, 1935: 26-28) Other kind of unpopular questions are those oriented towards privacy or those forcing the respondent to admit that "he doesn't know".

#### **Kinds of questions:**

According to the degree of openness we distinguish following kinds of questions:

- open questions (only a question does not offer standardized possibilities for an answer)
- **semi-open questions** (standardized answers with the open possibility "another")
- **closed questions** (limited variants of standardized answers)

In quantitative research closed questions with precisely formulated answer variants should predominate. We want to get certain standardized and thus unified information from respondents in order to statistically process the answers. In the formulation of answer variants we must keep in mind two things: particular variants must be mutually exclusive (i.e. the categories should not coincide, they should be unambiguous) and at the same time, every member of a sample must have occasion to put his answer into one of the formulated variants. We should avoid the situation when a respondent cannot find an appropriate answer for him.

Semi-open questions have formulated answer variants; however, one open variant "another" is present. The researcher uses this possibility when we are not sure whether we have covered all nuances of occurrence of these phenomena in the proposed questions. For example, we ask about a favourite sport. We formulate in proposed variants sports that are the most popular and the most common in our region. On the other hand we cannot presuppose whether someone from the sample prefers water polo or curling as his/her favourite sport. In the data processing we can decide what we will do with the category "another". If it is only a marginal category in the number of answers, it is possible that these answers will stay "lost" in the column "another". However, if we realize in data analysis that many respondents used that possibility and we can identify the same regularities in them, we can re-categorize the variants of this feature. And for the most often answers with the same content we create new particular categories.<sup>7</sup>

Open questions in quantitative research should be used only rarely, only in case we have good reasons for it. One of these reasons can be that we are interested in a new, not examined problem and we are not able to formulate standardize answer variants. However, if there are more of questions like this, we have to rethink whether our research problem is "ready" for quantitative research and whether it is not better to use some technique of qualitative research in this phase. Another reason for placing an open question into the questionnaire is that in a particular case we are interested for some reasons in respondents' spontaneous reaction, we don't want to influence in any way their answer. We have to remember that if we carry out quantitative research and we do not formulate the variants in the most comprehensive way before carrying out the research we have to do the classification subsequently in order to statistically analyse free questions. The advantage of such a classification is that it is empirical, inductive, and it comes out right from data. On the other hand, with the number of questionnaires over a thousand, a subsequent classification is an unbelievably difficult and demanding activity, particularly in cases when there is a plenty of different but similar answers.

Another classification divides questions into **direct** and **indirect categories.** A direct question leads up to a direct answer, but it is not always suitable, e.g. in case of delicate topics in which we suppose opinions different from those generally accepted. In an indirect question

<sup>&</sup>lt;sup>7</sup> The fact that some unexpected groups of answers occur in advance can inform us about some kinds of mistakes. Either we have made a mistake in sample selection, or we have discovered an important change in our research. If the answer "curling" occurs very often, it is possible that an interviewer by chance conducted interviews in front of an ice arena in which a curling match was held, or curling has become an important item of sports activities. It can lead us to some kind of self-reflection, or to a discovery of a new reality that we describe in our research report and our future researches will take into consideration.

the respondent rather decides to express his/her inner persuasion, even a bit deviant from the mainstream - without feeling discomfort. The most often way of substitution of a direct question by an indirect one is switching from a personal form into an impersonal one. Instead of "What do you think about..." we use the formulation "Which of these opinions is the most widespread?" We suppose that the respondent picks the opinion that he is of.<sup>8</sup>

In the research we also distinguish **basic and control questions**. Control questions are used to specify or complement basic questions. For example, after a certain battery of questions about the given problem a question arises that confronts the respondent with preceding answers. After the general question: "How do you assess problems connected to business activity in SMEs?" we pose in our research specific question: "Have you ever thought about quitting the business?" We can find out what the relation of satisfaction with the conditions for entrepreneurship is with my personal feelings that I have had enough. We can check the **direct question** by an indirect one. The **indirect question** does not precede the direct one. At the same time we can use impersonal questions to check the personal ones, and to check open questions by closed ones, or in reverse.

When researcher formulates questions for the questionnaire or the interview it is necessary to keep several rules. Disman (1993) writes about them in the part "Questions about Questions". He tries to show us which questions we should ask in the case of every question and in a survey or what we should ask about our questions as a whole.

Until now, we have spoken about the questionnaire and the interview together, about the questions that are the common tool of both. However, there are several particularities that need to be discussed separately and that can influence the researcher's decision whether to use the interview or the questionnaire in a particular case.

<sup>&</sup>lt;sup>8</sup>The degree of personal perception of the questioning basis that is deliberate helps to state the degree of personal interest in the problem. In 1943, American soldiers answered the question about female participation in the army. In the series of questions, personification of the answer consequently deepened and opposition to the women participating in the army increased (from a generally acceptable need to increase the fighting efficiency in war time to the approval of the decision of their own daughters to enter the army).

#### Interview

Interviewing is a research technique used to get vocal replies from respondents by oral language communication. As a particular research technique it has its specific rules and procedures of preparation, use and processing. Generally we distinguish two types of interview: **standardized** and **non-standardized**.

Standardized interviewing means that particular interviews are carried out on the basis of the given structure - all respondents are asked the same questions in the same way and their answers are recorded onto record sheets. These written materials for the standardized interview are similar to the questionnaire with the same sequence of questions, the structure of them and preparation of possible answer variants. Interviewer marks those variants that capture the real answer in the best way.

The interview prepared in this way has a clearly limited frame of research intention and this standardisation enables easier data processing. Standardized interviewing can be carried out when we enter into familiar milieu and issues with our research and we do not expect any new information. If we use this method we obviously want to confirm already known realities.<sup>9</sup>

Non-standardized free interviewing is also to some extent bound to the main topic and frame of the given issue, but it enables the researcher to carry out a wide, and in new information rich investigative research. The role of the interviewer is to maintain the interview in the given topic, but he/she acknowledges that every respondent has something else to say on that topic and it is necessary to give the space for these answers. In this way, a lot of personal experience, knowledge and evaluations that help us understand the topic gets into the results. This type of the interview is used if our research issue is not explored enough and researcher enters into new field.

<sup>&</sup>lt;sup>9</sup> In the selection of appropriate standardized answers we meet with the issue of determining their appropriate content. Categorization (the determination of possible answers to the questions) can correspond to the research aim, or categorization used in statistics, or another requirements and possibilities of processing of some statistical operations. All possibilities have their good reason and meaning - we have to choose what helps us more. The possible thing is that we can't get particular answers (about age) due to the structure of the question. We then classify them into subsequently created categories.

In between of these two types of the interview there is a whole scale of intermediaries - semi-standardized interviews. In a semi-standardized interview there is usually a firm list of questions but as well it offers the freedom in their formulation, order and recording.

For its apparent simplicity based on the similarity with every day activity, questioning attracts a whole spectrum of professionals and nonprofessionals. Sociological questioning is thus the formalising of some part of social communication that everyone of us daily participates in. To not dismiss this research technique the important thing to say is that the interview enables us to obtain equally reliable information as from the observation, the questionnaire, and study of documents when keeping some rules.

Standardized interview should have is always marked by the role of the researcher and his effort to direct the interview and at the same time to give the impression that the interview flows spontaneously. The basic issue (the condition of acceptability) of this research technique is to find the way to reduce the disturbing influence of the researcher.

At first, the researcher must try to perceive the personality of the respondent without prejudices and to register all his answers. A serious communicative barrier can be the difference between the researcher's and respondent's education. The researcher should adapt to the respondent in the pace of speaking. The age and gender of participants also influences the interview (the experience say that the optimal is similar age and the opposite sex). The interview should not be carried out in busy places, within the presence of foreign people or at too official working places. The researcher should point out the importance and the significance of the interview's results. He should be shielded by a credible institution to increase confidence and reduce fear from possible abuse of the information. To record the interview the researcher should ask for permission from the respondent. The manner of recording should not influence the fluency of the interview.

The interview is a technique that is based on direct contact with the respondent. Many-sided demands on the abilities of the interviewer emerge from this:

- the ability to enter into conversation
- make a trustworthy impression and
- to create a pleasant atmosphere,

- to keep the necessary level of equality of conditions in which the interview with particular respondents was carried out,
- together with the ability to react and to be flexible with situations, the researcher should be able to get back to the main topic of the interview, or rather
- he should be able to systematically cope with unplanned, but relevant digression from the interview topic.

If it is not an intentional selection, it is necessary to explain the principle of randomness to the respondent and the importance of his answers to get objective results. The important thing is the selection and handling the way of making the first contact - input information and "breaking" of the resistance.

## Questionnaire

Questionnaire as a technique of quantitative research has several common features mainly with the standardized interview. On the other hand, it has its own particularities. Disman (2002: 141) speaks about several pros and cons regarding the use of a questionnaire in opposition to an interview:

- The Questionnaire enables us to get information easily from a vast number of individuals in a relatively short time and with relatively low costs.
- The colleagues doing the field research are needed only sometimes (in the use of personally collected questionnaires) and the requirements on their experience are low.
- The costs of research on the diffused sample are relatively low.
- Anonymity is relatively convincing.
- The conditions are formally very similar in different cases, the researcher's influence on the answers is very improbable.
- On the other hand, the questionnaire in comparison to the interview has several disadvantages:
- It places high demands on the willingness of the respondent. It is easy for him to skip over some questions or to not answer at all.

- It places high demands on preparation before beginning in the field research.
- A situation can occur whereby someone other than the one chosen for the sample would answer the question e.g. another family member, or the whole family team would do that.
- The returnability of questionnaires is very low.

We will further analyse the returnability of questionnaires. In the technique of the interview there is the danger of low returnability – caused by the situation when a lot of respondents do not accept the interview or would interrupt it later on. The reason for this can be the content of the technique, the personality of the investigator, insufficient knowledge and awareness of the respondent. Time and place of the interview can also influence the research, but there are other outer impacts.

In the questionnaire technique there are a lot of factors that influence the returnability – practically, it depends on every stage – whether it is in the formal or content aspect of the questionnaire (title page, degree of anonymity, scope of questions, topicality and disagreeableness of questions), in organizational security of filling-out the questionnaire (degree of questioner's presence, the difficulty with filling it in), dramaturgy of questionnaire, and even in the way of filled in questionnaires, the level of voluntariness of the respondent, in the support of returning a securing their collection.

The returnability of questionnaire and even the acceptability of the interview are also influenced by their time demand. The interview that is longer than one hour is inacceptable. The same thing is valid for the time that is needed to fill in the questionnaire. The length of the interview enables us to work with more variables (200 and more), however in the research practice their number is limited to only several tens. We can adequately extend the interview, for example, when it is realistic to presuppose that the topic of the interview is interesting even for the respondent. It is generally valid that the returnability of the questionnaires decreases with their increasing difficulty. Low returnability of questionnaires means serious deviation of the results. Focusing on the preciseness of the research project, and thus on the thoroughness in formulation of the research object with corresponding

and actual hypotheses, an attentive statement of the structure and the extent of the representative sample along with the proposition of quality, effective obtaining and processing of the results is not only a question of saving and increasing the effectiveness of the research activity. It is mainly the respect to the demands on really great scientific work with the opinions, needs, interests and values of the citizens. It is a part of our understanding of democracy.

If we know the conditions and possible endangering of returnability, sociological research practice offers lots of procedures that can help to decrease the risk of low returnability.

## **Observation**

Observation is the oldest research procedure. Many chronicles, descriptions of specific historical events and the first technical texts were built on the basis of observation. Observation as a scientific technique (method) of research has been adopted by various scientific disciplines, even sociology. In sociology, it is mainly the observation of people's behaviour – in various settings and situations. Its results could be a basis for analysis of motives for this behaviour. However, we can imagine what should be a part of the documents recording the content of the observation (a detailed description of the observed situation in particular milieu and particular time, positions of respondents in it, detailed description of the respondent's behaviour in the particular situation). Expected reasons for this behaviour and our own immediate evaluation of the whole social phenomena or process should be also recorded.

Observation is the ubiquitous component of the cognitive process (in everyday life and in scientific practice), so the observation is to some extent a universal method. Observation as a specific sociological method then is under some rules and verified procedures. We use observation as a method to investigate the phenomena that are directly available to senseperception and were not provoked by the intervention of the researcher. The advantage is that the observation of objects is carried out in their natural setting. The disadvantage is that we can record only those phenomena that can be observed from outside. The observation does not have to be used as an individual method in the research. It can be combined with another methods and techniques.

Scientific observation is different from the everyday observation by:

- subordination to the clearly formulated research aim
- planned procedure
- systematic recording of the observed
- checking of information

Similarly to other methods, the observation can be **more or less standardized**. Non-standardized observation is then a method of qualitative research. We will mention it in the third part of our publication as a part of the ethnographic method. If we want to quantify some observed elements in the analysis, we have to make a plan of observation and choose the unit of observation. Procedure for the observation then consists of the following steps:

- 1. We choose the object of observation, define the problematic situation (variables and variants that can be observed)
- 2. We choose the unit of observation particularly definable demonstrations of behaviour (what, who, when, where, etc.)
- 3. We choose the way of observation:
  - a) standardized, non-standardized (only general aim, preliminary units of observation, that I do not have to strictly follow)
  - b) participating and indifferent
  - c) secret and evident

As we have already mentioned, there are some rules in standardized observation:

- precisely calculated and defined elementary phenomena that are observed
- precisely classified values to which the phenomena are classified
- precisely determined features and their categories
- the results of standardized observation of various units observed by various observers are comparable.

On the contrary, a non-standardized is characterized by:

- only the aim of observation is given
- tools of data collection are not formalized,
- it is possible to adjust the choice of the observed, fundamental phenomena or proposed hypothesis during the observation
- there are high demands placed on the observer

In the selection of the way of observation, we at the same time decide between participant/non-participant and secret/evident observation. In non-participant observation, the observer is an evaluator separated from the situation (he is not a part of it). In participating observation it is the opposite. The observer is a part of the group, situation or place that he observes. The example of non-participant observation, for example, is a school inspector present in lessons. He is present in the class but he is not a part of relationships or the situation of the lesson itself and he does not enter into conversation and does not interact with the observed. The example of participating observation can be a situation where the researcher becomes a part of the observed group and integrates into it. He can do it in a secret or evident way. I.e. the observed either know or do not know that they are the object of the research. The advantage of secret observation is a greater chance to capture the natural behaviour of the research objects. However, some ethical questions that the researcher has to answer himself are connected to it, particularly when it is not possible to keep the anonymity of the observed.

Observation as a method is useful for the first orientation in the topic. It can capture the <u>dynamics</u> of quickly changing phenomena and their continuity. These are indisputable advantages. On the other hand, several errors of observation can occur in the application of the method. They are usually called the "halo effect" and they are connected to subjectively conditioned inclinations and the capacity of the researcher to evaluate the recorded phenomena.

## Document analysis

Under the term document we understand the object created especially for the transmission and keeping the information (any information fixed in hand-written or printed text, on the audio tape, film, etc.) Documents are from the point of view of the research the unintended information sources, and thus they were not created for research purposes. Because the documents were originally created for other than research or diagnostic purposes, it is necessary to separate those information from documents that are in relation to the research aims.

We differentiate:

- primary documents: direct records register office, statistics, etc.
- secondary: already processed and they contain some kind of interpretation of primary documents

Documents can be divided into personal (letters, diaries) and public (concerning bigger structures – organization of groups, municipalities, etc.)

When working with documents, it is necessary to verify their reliability and validity – so called source criticism. It can be

- external: we consider whether the source is reliable
- internal: we consider whether information in documents are true

From the data analysis point of view we distinguish several techniques or ways of document analysis. One of the most important classifications of these techniques is their classification into quantitative and qualitative. It is necessary to say that they complement and combine with one another in the research.

Qualitative and quantitative document analysis:

- 1) **qualitative:** it is a necessary prerequisite for further quantitative analysis, it is used when the purpose is the content interpretation of materials or a type of document that is available (e.g. medial analysis: in what way are the Roma people presented in the media, etc.)
- 2) quantitative: it is used if a lot of material is available and if it is necessary to classify it (e.g. within medial analysis we would be interested in using of slang in the media – we would choose
the sample and calculate the frequency of occurrence of certain catch words or phrases, etc.)

A specific kind of qualitative analysis of documents is the discursive analysis that we will mention in part III, dedicated to qualitative methods.

Quantitative analysis of documents has spread primarily after the 1940's. A special method was developed, so called **content analysis:** the transfer of mass text information into quantitative indicators.

- 1) it creates a set of mutually exclusive and comprehensive categories (the content unit must be delimited by the primary qualitative analysis) for example, at first we look through some editions of a magazine, and we find out that the articles we can characterize as pro-American often use the term "freedom" or "transatlantic relations", consequently we follow the articles in subsequent editions that contain that terms more than 10 times.
- 2) we record the occurrence (distribution) of particular categories

## **Data Analysis and Interpretation**

After collecting the data, when writing the final thesis at school, or research report from professional sociological research, it is necessary to distinguish two processes: analysis of the data and interpretation of results. Interpretation follows data analysis and these steps cannot be interchanged and have to be distinguished. After the interpretation of the results in logical order of scientific research the recommendations for the practice may follow. By this, research process is closed because the researcher is not responsible for the real use (within abuse and non-use) of its results.

#### Analysis

Data analysis is the process in which we make basic summarization and systematic operations that prepare the researcher for the "reading" of these data. We arrange data into charts and graphs, we can calculate some means, percentages or indices, and we test the correlations among variables and look for an optimal form in which we introduce the data to the reader. The first reader will be a research worker who transforms arranged and processed data into compact ideas (sentences) and then a wide range of interested people from the professional and nonprofessional public can follow. Data analysis should show to everyone what the research has brought. It should concentrate obtained results into certain sets that relate to particular hypotheses and to state whether and to what extent the given hypothesis was confirmed or verified.

We would like to highlight some frequent bad habits and mistakes in students' theses when students use the research part to broaden the extent of their thesis without adding anything new to the explanation:

**a.** It's useless to repeat all the data from tables or graphs again in written text. The researcher should focus only on those numbers that show remarkable deviations from the normal (average) distribution of data, or they confirm expected data or links presupposed in the hypothesis. The output from the analysis is thus pointing out those results that we will interpret in further cognitive process. In this phase, it is the output that could be interpreted in different ways by various professionals and non-professionals; and because of this, it should not

contain any subjective claims – instructions or directions according to which the future interpreter should proceed. Analysis is thus an objective process of data systematization.

**b.** If we work with average values it is necessary to say that a mean as a tool of synthesis of the group of data, is to some extent a very misleading number. In the final theses there are means from means. What is the expressive value of such kind of data?<sup>10</sup>

**c.** Students often overestimate the real value of the data, and they forget to relativize their findings in regard to the size of their sample, or other important properties of the data. They interpret as interesting those numbers, differences and deviations that are of a marginal character. In general, it is possible to say that unreasoned analytical conclusions are caused by incorrect data processes.

**d.** We think that it is insufficient if the author has recorded only that "those who are unemployed because of a low education have low education". The author has probably lost his train of thought, which is not rare in cases when the thesis is written in haste or the author does not manage to cope with data analysis.

**e.** In formulation of conclusions, the author often does not realize the fact that he/she works only with a small sample that is representatively restricted and he is only writing a school thesis. In the school thesis there are different rules from those of a real research and cognitive process.

**f.** The orientation of data analysis often deviates from requirements for hypotheses verification. It is evident in cases when the student has ignored the orientation function of hypotheses or the hypotheses were formulated incorrectly.

For responsible and transparent data analysis it is necessary to become familiarized with at least a minimum knowledge about visualization research techniques and their creation. It involves the creation of tables and graphs.

#### Tables

Long-time practice confirms that students often have problems with the formation of tables and they make certain mistakes. Because of this we

 $<sup>^{\</sup>rm 10}$  It is possible to read more about the mean in the corresponding part of mathematical and statistical analysis.

have to start with a short introduction of the table form. Every table has its name that characterizes its content. If it is not stated in another place, the name has to contain the information about time and place that are related to the table data. If there are several tables in the document, in order to keep the transparency, we mark them by an ordinal number: "Tab. no. X" e.t.c.

An actual correlation table (a table presenting a relation of two variables, two features) contains a specific number of rows and columns. This information is one of the characteristics of the table – we can say that it is the table of 2x4 or 7x3, etc. The stated numbers are basically variants of the value of the first and the second sign. For example, one of the variables is education. In this case the sign education can have different variants. For example, we can divide the set of respondents into those whose highest education is with a secondary school and those who have higher educations. The same set can also be, for example, divided into four education categories – elementary, secondary education with a graduation exam, secondary education without graduation exam, and university education. Educational structure can be further precised and the number of categories may increase. The important thing is to categorize every variable so that it suits our needs and requirements for verification of hypotheses.

In the table both variables have to be named (e.g. education and satisfaction with income, while one sign presents a dependent variable and the second one an independent variable) and at the same time, particular variants of both signs (categories) have to be named in the first row and the first column. The last row and the last column contain total values – in these (outer) columns and rows there are sums of all values of the given column and row. If in the table cells there are the absolute numbers, it is the **table of absolute frequencies**. If there are relative numbers (percentages) calculated either in rows or in columns, it is the **table of relative frequencies**.

In a table that serves as a basis for further analysis, absolute and relative values can often be found. Absolute values are very important in correlation tables, for example, for verification of number of cases of the given category. If there are enough respondents in the research as a whole, there may be a problem with number of respondents belonging to particular category – e.g. of those who took part in the election and have an elementary education. If this category is underrepresented, it is necessary to take it into consideration during the analysis and interpretation of the data. Relative frequencies we use as a base for analysis, and thus for the verification of hypothesis about the relation of two variables. As long as the researcher stays at the level of so called univariate statistics and does not proceed to bivariate statistics (use of the statistical testing procedures), it is up to his/her evaluation which differences in the table he/she points out as interesting and proving or refusing the hypotheses. In relative values it is important to pay attention to the whole from which the percentages are calculated. In the following table, there are two percentages: 91,7% and 12,4% assigned to the absolute value of 11 in the first cell of the table. 11 respondents thus have elementary educations and at the same time they did not participate in the elections. 91,7% of respondents that have elementary education did not participate in the elections, but of those who did not participate in the elections altogether, the respondents with elementary education create only 12,4%. Of course, a table as a basis for analysis and a table that we use in the final research reports should not look the same. During writing of the research report we try to simplify the table in a way so that it would contain exactly that data which we want to point out and that we refer to.

The table as a result of analytical processing of the data has to follow the mathematical and statistical logic. The student or the researcher has to choose appropriate variables and make use of the opportunities to present the calculated data. Even in the simplest modification it is possible to easily make a mistake. If there are, for example, incorrectly calculated percentages (it matters whether I choose row or column percentages), or if the percentage calculated from the whole number of respondents (the differences among percentages only copy differences in absolute frequencies which is not necessary in the analysis and interpretation because it does not bring any new cognitive quality), if the dependent and independent variable are interchanged, if we incorrectly calculate some means and things like that, the internal logic of mathematical and statistical operation is lost. The results analysed in this way are not possible to be interpreted without the risk of debasement of the original results.

## Example: Table 1

		Electoral participation		Total	
				I do not	
Education		No	Yes	remember	
Elementary	Count	11	0	1	12
	% within Education	91,7%	,0%	8,3%	100,0%
	% within Electoral Participation	12,4%	,0%	12,5%	4,3%
	% of Total	4,0%	,0%	,4%	4,3%
Lower secondary	Count	14	28	1	43
	% within Education	32,6%	65,1%	2,3%	100,0%
	% within Electoral Participation	15,7%	15,6%	12,5%	15,5%
	% of Total	5,1%	10,1%	,4%	15,5%
Higher Secondary	Count	48	95	3	146
	% within Education	32,9%	65,1%	2,1%	100,0%
	% within Electoral Participation	53,9%	52,8%	37,5%	52,7%
	% of Total	17.3%	34.3%	1,1%	52,7%
University 1.degree	Count	10	22	1	33
	% within Education	30,3%	66,7%	3,0%	100,0%
	% within Electoral Participation	11,2%	12,2%	12,5%	11,9%
	% of Total	3.6%	7.9%	.4%	11.9%
University 2. degree	Count	6	35	2	43
	% within Education	14,0%	81,4%	4,7%	100,0%
	% within Electoral Participation	6,7%	19,4%	25,0%	15,5%
	% of Total	2,2%	12,6%	,7%	15,5%
Total	Count	89	180	8	277
	% within Education	32,1%	65,0%	2,9%	100,0%
	% within Electoral Participation	100,0%	100,0%	100,0%	100,0%
	% of Total	32,1%	65,0%	2,9%	100,0%

It is thus necessary to be precisely occupied with the principles behind the table's creation and with its transparent drawing up to create a suitable basis for the analysis of the data included. We must never forgive the role of analytical data processing in our cognitive process. We should not add new results of different mathematic operations into our report simply to show we know how to do it. On the contrary, we have to learn to do those operations that we need for directed data analysis and that serve for the verification of our hypotheses.

During data analysis in the table, errors often occur even in the simple addition of numbers, in calculation of relative frequency (percent) and other mathematical operations. We can say that it is often a consequence of a poorly arranged or inappropriately designed table.

#### Interpretation

Interpretation of the data is process in which the researcher uses his knowledge and experiential background. With this awareness he/she tries to present to the reader his/her own ideas about the source, reason for the given relation, or for recorded distribution of data, circumstances that could influence it and meaning of that data. At the same time, researcher should try not to change unscientifically the meaning and the value of results. Interpretation of a research is process highly marked by subjectivism but it does not mean that there is a place for fiction and fabrication of different meanings that does not come from correct data analysis. Trying to adjust the results to those expected and confirm the conclusions prepared in advance is big interpretative mistake. Unfortunately, this procedure occurs in practice and it is only a question for the scientific community whether it is able to eliminate similar errors from professional discourse.

The authors of the interpretation (not only students) often exceed (repeatedly, mildly or largely) the interpretational opportunities of the data. The control question in detecting such cases is: "How do you know it? Which of your data or other facts can confirm this claim?" During the interpretation it should not happen that the presented theses (statements) do not emerge from the data and they present only the researcher's intended meaning that is not justified at all.

In a professional practice, the situations may occur when the researcher suddenly realizes that his/her data set is too small for the

interpretation, or the data that would reliably clarify the meaning of the results are missing. The opposite situation can occur as well – the researcher has too much information that often do not relate to the problem or hypotheses and he/she is not able to cope with them, to evaluate them in real time and include them into analysis. Both extremes result from violation of the principle of gathering only those and at the same time all relevant data needed to verify the hypotheses, no less no more. We should not forgive any important quantity – some factor that can significantly classify dependent variables or verify formulated hypotheses. At the same time it is useless that the research would contain something extra – that is not related to the hypotheses.

In practice, for example, students put into questionnaires (as the most often used research technique) a large extent of social and demographic characteristics of the respondent that are not often used in further analysis. Students do not realize that these characteristics should be included as independent variables that should (according to their presupposition) interpretatively meaningfully classify the answers to substantive questions.

#### **Recommendations for Practice**

This is often the target item in the cognitive process. We investigate social reality with its problems to make some decisions that will help to eliminate or minimize these social problems, or prevent those that we can presuppose.

When writing a school thesis the student formulates recommendations for practice only as a part of the training – nobody will apply them into practice. This seemingly restricting fact is actually an opening of possibilities to propose whatever the student considers to be appropriate, however, it has to be supported by the data ant thus, the research findings.

# Methods and Techniques of Data Analysis: Statistics in Social Science Research

Statistics in social sciences or in sciences working with large-scale phenomena means improvement of methodological tools. It enables the investigation of relations of various kinds between phenomena and processes. Applied statistics should therefore be a part of the qualification of every social scientist.

To define statistics in more detail, we can say that it is a formal science. If we would like to define it as an independent science, it must have its own subject - as any other independent science. When considering statistics a subject-matter science, its subjects are large-scale social phenomena. But this would mean putting statistics into social sciences. In such case, statistics as a social science cannot confirm its activity (use) in natural sciences. When we consider that social phenomena are the subject of economics, legal sciences, sociology and other social sciences, it must be applicable in all these sciences.

A solution to this is to consider statistics a **methodological science** in this case statistics helps all subject-matter sciences (not only social sciences, but also e.g. physics, chemistry, biology, sociology etc.) to deal with the large scale phenomena and with their proper analysis. Wallis and Roberts (1956) state that statistics is not a set of meritorious knowledge but a **set of methods** for knowledge acquisition. As such it must be confronted with the base of general methods of knowledge acquisition, i.e. with general scientific methods.

Statistics has a logical and mathematical part. Logic is necessary for every mental activity, therefore it should be commonplace for every university student to deal with statistics and cope with its procedures and principles.

Statistical study of society (of social phenomena and processes) has certain characteristics. Three viewpoints are important:

- 1. The quantitative viewpoint is peculiar it is expressed by terms like plurality or "large-scale phenomenon" which occurred in the definitions of statistics.
- 2. Another one is collectivity which cannot be equated with a social group but is not just a sum of units. For example the population of a country has certain characteristic features which cannot be

inferred from the features of individual inhabitants. Statistical collectivity is usually called "population".

3. Statistics observes the variability of phenomena. This occurs only by a sufficient amount of units. Then it is possible to speak about a relative stable regularity or even about rule.

The goal of every scientific study is to reveal regularities in certain areas. It is the case also with social sciences, although here it is not possible to reach regularities in the form of a rule. In social sciences fulfilling of certain conditions does not have to lead to certain phenomenon, we cannot forecast social phenomena with absolute certainty. Many of the regularities appearing in social sciences have statistical character - they are based on the observation of large-scale phenomena. The key to understand statistical regularities and rules is the law of large numbers.

## Statistical Theory and Practice

As we already mentioned, statistical theory is mostly methodological theory with a mathematical and logical base. This logical base is closely connected with statistical practice. Statistical practice makes up an important information system necessary for ruling the state (especially in the case of centralized governance) and other social units. The theory says that every administrator should be able to use statistical information in its analytical and synthetic form. But practice says that statistical information speaks about social reality too objectively and definitely, which makes it impossible to adapt the reality to the concrete interests and goals.

Statistical data are useful not only for planning but also for analyses of various kinds and purposes. Before any kind of our own empirical research we should get familiar with relevant data about natural or artificial environment and population on the basis of statistical data. It helps us to be informed about what to expect and to put new information into an objective context. We can determine the causes and better estimate the consequences.

Nowadays, it is much easier to find required data. Various issues are statistically analysed, so these data can be gained from many different sources. Data about things, people, human behaviour and various operations are available. We get these data mainly from population and housing census which is done every ten years, but also from selective surveys (micro censuses and family budget surveys) that are more specialized but faster and cheaper so they are able to bring new data very flexibly.

Statistical data have a fundamental characteristic which should be regarded as a disadvantage. They refer to realities which have already happened and their informational value is decreasing because their processing takes a long time. So these are **ex post facto** data. Despite this, in practice, statistical data are used **propter hoc** which means that we turn to the future. It is evident that the bigger the distance between processed data as starting points and as prognoses is, the bigger are problems with their validity. If we estimate and plan for the future on the basis of outdated data, we cannot rely on their adequacy.

The speed of social changes determines the vitality of a data. The quickly changing social phenomena have to be predicted without using statistical data. Statistical data are more important for prediction of slower social changes. Statistics takes this dynamics into account and the branch of statistical dynamics investigations is here very important. It means, inclusion of the variable of time and the creation of a time series made of findings about social reality. This enables the depiction of large-scale social phenomena, reveal long-range tendencies and increase in the probability of fulfilling the estimates of social changes.

## Statistics as Science and as Methodological Tool

If we consider statistics an independent science, its peculiar subject would be large-scale phenomena, especially social phenomena. The largescale phenomena could be then for example consumption, election results, criminality or various addictions. From this point of view statistics is interconnected with other social sciences, especially with sociology, political science, economics, demography or legal sciences. But the large scale is not sufficient precondition for creating an independent science. When stressing the subject-side of statistics it can be at the expense of the methodological side. From the methodological point of view statistics is a formal science. The emphasis on this side of statistics leads to quicker practical use in other fields of science. From this point of view greater attention is paid to the logic-mathematical apparatus of statistics. This apparatus is always separately applied to particular sciences - on the basis of their special needs. The application of statistics means the emergence of new individual fields of statistics, for example social statistics, health statistics, or the most important demographic statistics.<sup>11</sup> An example for published output is the annual material of the Ministry of Labour, Social Affairs and Family of the Slovak Republic - Social situation in the Slovak Republic.

#### **Dynamic Statistical Description**

Statistics can be divided into statics and dynamics. Dynamics means that we consider certain movement. Because movement takes place in time, it is time which is the most important variable. Time can mean longer time units - years, decades, centuries – in these cases we speak about historical dynamics, or it can mean short time units - hours, minutes, seconds - we speak about dynamics of the functioning.

In sociology, various ways of recording of certain states of phenomenon in different time periods are recognized. It is a discontinuous observation by which we get so-called "cross-sectional data". When we line up these data according to the time, we can come to certain image of the development of the phenomenon in time. The concept of time is connected with other concepts that determine whether particular change is:

<sup>&</sup>lt;sup>11</sup>Demographic statistics is a statistics of population. Sociological research cannot do without it. Sociologist uses from the demographic statistics following data: distribution of population, demographic structure of population (according to gender, education, age, health, economic activity, nationality, and some other characteristics) and these indicators of population dynamics: marriage rate, divorce rate, natality, mortality. Besides these population phenomena we distinguish population processes such as: ageing of population, feminization (of education, health care), changes in function and structure of family, population development in given area, migration of population. Also other applied fields of statistics provide important and interesting data for sociological processing. Most desired are data about market. Other applications are social statistics, judicial statistics and cultural statistics.

#### a. continuous or

b. discrete or discontinuous.

Continuous process is, for instance, upbringing, the consumption (physical and moral wear) of different objects. To study the continuous social change is actually impossible. It is not possible to observe how the change of media communication content influences personality (e.g. value) changes. Even the individual himself is not able to realize gradual changes in the personality under permanent influence of an external impulse. And it is more difficult in the case of social group. In the social sciences we must do a compromise and it is the discrete, discontinuous observation of social processes. In sociology repeated measures (e.g. market research or public opinion polls), statistical reports (quarterly, annual) and also census (every 10 years) are used.

After lining up the results of these measures we get a line of values and we have the tendency to create a curve. But it is not correct - because we have no information about the points between the data. We just assume that the development of given value was as we have drawn mostly rectilinear. That's why it is better to depict the development of social phenomenon in columns instead of line graph although the second one is nicer for the common user.

The study of actions, events, situations, processes and other dynamic phenomena differs from static observation. Therefore we distinguish between static structure (e.g. structure of social group) and dynamic structure with its constant changes. Their observation and depiction exceed the possibilities of mathematical statistics and there exist various mathematical models for that.

#### **Time Series**

Time series are the most common example of dynamic statistical analysis. Their aim is to record an observation series of certain phenomenon in time. We distinguish four types of movement in time:

- 1. secular trend or general long-term movement
- 2. periodic movement cyclical
- 3. periodic movement seasonal
- 4. irregular variation

In sociological practice there is always a compound of all four movements and therefore the final change of a variable in time is defined as a product of all four components. A time series should be analysed according to following operations:

- 1. long-term trend is specified
- 2. seasonal variation are defined
- 3. irregularities and accidental variation are equalized
- 4. cyclical movement is found out

Usually, only the first two operations are used. Certain corrections might be necessary to be done before the analysis of the time series. We should consider:

- 1. if the way of its recording cannot be specified for example by recalculation of monthly records on day (because months have different number of days),
- 2. if it is not more appropriate to use relative data instead of absolute data about the size of given phenomenon it means per capita consumption, divorces per 1 000 inhabitants etc. or
- if it is not necessary to consider changes in population structure e.g. ageing. That (or another factor) can be the actual cause of observed change. In this case we divide the sample into significant subgroups - e.g. we observe separately suicide rate by youth, adult and old, or drug addiction in towns and countryside.
- 4. then we should examine if data standardisation wouldn't be better (to reflect changes in structure), or
- 5. if the defining of observed phenomenon didn't change (e.g. infant mortality in relation to the definition of a child born alive, unemployment in relation to the people looking for job, or property offence in relation to financial calculation of damage, changes in administrative structure which often have a destructive influence on showing changes for certain territorial units - e.g. districts)

We have talked about how statistics depicts the structure of given phenomenon (static statistical description) and of course, the change of given phenomenon (dynamic statistical description). Besides these functions, statistics has a third function - it can record and use the relationship between phenomena and deepen the way we get to know them. The relationships can be static as well as dynamic. It is actually a statistical understanding of causality.

## Charm and Problems of Statistics

When we become aware of all the initial aspects of statistics as a science, we realize that statistics creates a beautiful and surprising landscape in which it is a pleasure to move if we move with understanding. Statistics, as many other things, is misused and becomes a means to manipulate facts, even reality itself. Researchers that come through different statistical tools, although they keep all the principles of scientific correctness, may come to the different results. A very simple example is the calculating average value. In statistics it is natural that by calculating mode, median, arithmetic mean, harmonic mean or geometric mean, we come to different figures - although the mean is always calculated by using the same units. (Frisch<sup>12</sup>, In Swoboda, 1977)

Statistics brings to the cognitive process of social sciences both of its components - logic and mathematics. The sense of mathematical exactness and perception of logical links and their meanings (also in using terms) and the willingness of the student to comply with procedural requirements are important factors already from familiarization with the basic principles of statistical procedures. It is not possible here to plead that the student does not have a "talent for mathematics", because it is about the general intellectual skills of the individual that should be on an adequate level by university students.

Ragnar Frisch in the above cited publication by Helmut Swoboda emphasizes that the statistical verification of hypotheses (according to Jerzy Neyman, as well as E. S. Pearson) is actually an attempt to reject the given hypotheses. By the statistical verification it is therefore not possible to say that the hypothesis is not correct. We can just state that it is not possible to reject the hypothesis. And this is a difference that strongly influences the possibilities to use the gained results in social practice. We would like to emphasize that not only in mathematics and statistics but

<sup>&</sup>lt;sup>12</sup> From the introduction of Nobel Prize winner prof. Ragnar Frisch

also in social sciences the correct way of expression and subsequently the adequate understanding of used expressions are very important. It means that if we would have to act in the situation when the hypothesis was several times rejected, we would act **as if** the hypothesis was correct. But it is completely different situation than the one, in which we would be able to prove clearly the hypothesis **is** correct.<sup>13</sup>

In the area of quantitative research statistics has a broad application. But we do not want to say that a numerical data has a higher value and weight than qualitative data. Such comparison is not appropriate here, as well as the boundless respect for statistical data. They are seemingly convincing and give the impression that it is not possible to object to them but statistics as a social science (if we would understand it as such) and statistics applied to social sciences brings results which correspond to the character of social phenomena - they are unstable, unsteady and are not sufficient support for pronouncing final verdicts.

The inner dispositions of statistics and its social-scientific application are the reasons why people believe that "statistics lie". We want to learn how to understand statistics, so that we can consider the use of a particular way of collection, analysis and partial interpretation of figures, the likelihood of incorrect (consciously thereby reducing or unconsciously) use. Equally, we hope to be able to assess the adequacy and acceptability of other statistical data and analyses. We will constantly encounter them because they are a part of different kinds of planning, organisation, professional and public communication, understanding and decision making as a part of functioning of the modern coexistence. Through statistics we can learn more about regularities of our life. (Swoboda, 1977: 17-18)

If we should insert the meaning of statistics into social-scientific cognition through the prism of sociology, we could state that scientific work in sociology has four basic phases:

a. description of the reality - empirical collecting of information, data and knowledge

<sup>&</sup>lt;sup>13</sup>To complete this problem it is necessary to say that by testing the hypotheses in the socialscientific research it is possible to confirm or to reject the hypothesis.

- b. explanation of the gained knowledge and formulation of hypotheses
- c. deduction from these explanations empirically acquired knowledge is processed to generalizations
- d. verification of knowledges confrontation with reality.

In different sciences these phases have a different weight and importance. It depends on the role, phases of development and methodological tools of the given science. Sociology went through a longer speculative period in its development but even in those times it couldn't ignore the empirical knowledge. The personal experience of sociologists served for this confrontation. The need for sociological research and mass collection of data came later and the importance of statistics for support of the scientific character of sociological cognition began to grow. This process was gradual and was strongly accelerated in the last fifty years especially thanks to the new technical opportunities.<sup>14</sup> One of the impacts of this huge progress and time benefit was unfortunately that people have become unwilling to think independently. To avoid this, we should be able to reveal and use all information that we have acquired, or we should try not to exceed the depth and extent of data collecting we are able to deal with.

## Basic Population and Representative Sample

A complete population or statistical set can be made of any kind of units, but it must include all the units of the set or population. For example all married men in the given area, all traffic accidents in Slovakia in the last year, or all borrowed books from the town library in July and August.

<sup>&</sup>lt;sup>14</sup> The first half of the 1970s the computer technology emerged and it was the beginning of qualitatively new and more exact processing of quantitative results and large scale data. Until this time, results were processed manually and it worked so well that the scientists didn't see the need for innovations. Development of this tool was fast and its use was quickly broadened. This was thanks to the usage of punched tapes which replaced punched cards, then the magnetic tapes came and were quickly replaced by floppy disks and later USB flash disks. Of course, the growing capacity and speed of personal computers and the development of generally accessible software are also significant. One personal computer today is able to replace computational workplaces (specially located, equipped with air-conditioner and organized) of several scientific research institutes of the 1980s.

To observe the entire population - all its units - can be sometimes too expensive or even impossible. Therefore, statistics has some application for selection of a sample, which is with its composition and size able to cover some features of the basic population. Through the observation of the representative sample we can come to very reliable findings about the whole population.

The objective of sampling techniques serves to organize the sample so that the studied sign (e.g. quantity of children in rural family) in a representative sample is distributed in the same way (or at least similar way) as in the basic population. Then we do not have to study all the families living in the Slovak countryside but only a sample that is large enough. The individual size groups (childless families, families with one child, two children...) should be in the same proportion as they are in the complete sample.

In such representative sample we can carry out a research about incomes and expenses in families of different sizes, research about the reasons for certain types of leisure activities, or research about the educational aspirations of children in families of different sizes.

Some biological features, e.g. body height or bust line, show very characteristic division of this feature: there are fewer people that are very tall or very short but there are very many people with the height about average. It is so called "normal distribution" that plays a significant role in statistics. In many cases, we are able to forecast the occurrence of a certain feature according to its approximation to this normal distribution. But there are different ways which make up the main goal of statistical activity which is: to estimate through calculation what distribution of certain feature we can expect in the basic population if we are aware of its distribution in the representative sample.

"Statistical sample" is one of the basic statistical terms. It is a set of statistical units. A statistical sample can be divided into smaller units - subsamples, parts and units.<sup>15</sup> On the statistical unit we distinguish feature, indicator and unit value.

<sup>&</sup>lt;sup>15</sup>Units of statistical sample in e.g. sociology are: people, social groups, social units (institutions and organizations), objects (magazines, drugs), operations (payments, services), events (accidents, births), behaviour (purchase of consumer goods, visit of cultural facilities), attitudes, opinions, wishes, needs (socio-psychological units the observation of which hasn't got only sociological character.)

A statistical sample is a set of homogenous units according to certain feature(s) which is defined in terms of place and time.<sup>16</sup> According to the extent, we distinguish small samples (up to 10 to 30 units), middle-sized (30 to several thousands of units) and large samples - macrostructures such as ethnic groups, social classes within the state.

## Absolute and Relative Numbers

Absolute numbers inform us about the size of a given phenomenon, about the number of units that carry an observed feature. We can compare absolute data between the samples (e.g. men - women) only on condition that these samples are the same. But usually it is not so and the samples have different sizes, so we must find such indicators that express the size of the observed phenomenon in relation to the size of the sample. They are **relative indicators**. Relative indicators supplement the information that is provided by absolute numbers. We distinguish three types of relative indicators:

- 1. intensity indicators inform us about how often a given phenomenon in a given sample occurs
- extensity indicators inform us about the composition of observed phenomenon or sample - what part of the whole the observed phenomenon represents
- 3. index numbers are used for comparison of data expressed in time, differentiated locally, etc.

ad 1. For example, work injury rate related not to all workers of a company but only to those who work in risk departments. Marriage rate related not to the whole population but only to those who are capable of marrying (minus inhabitants under 18, married, alternatively those who are not able to set up a family with children because of advanced age).

ad 2. Extensity - structural indicators are proportional figures that express a relation between two quantities of the same content, from

<sup>&</sup>lt;sup>16</sup>For example: a statistical sample of motor vehicles accidents in Slovakia in the last year, statistical sample of subscribers of the magazine Politology, statistical sample of the population of Slovakia till 31 December 2014

which one has besides identical features also a differentiating feature or its modification. For example, all students and male students. It is always the proportion of the part to the whole. These proportional figures are called **relative frequencies in the strict sense**.

ad 3. These are proportions which make comparisons more illustrative and clearer. We express individual data in relation to a certain base that we choose in advance. This choice depends on the nature of the phenomenon and on our goals. By time series it is important to choose the year or other time unit that represents a certain turning point for us. Today a lot of time series have been created since 1989 or 1990 because it was a turning point in the development of society. A comparative value can be for us such value that we just want to achieve. All the previous data will then express the rate in which they came closer to this value.

We distinguish the creating of index numbers to one period, one value that does not change - it is called fixed base index. We can proceed in that way that the base is always previous data in terms of time. It is called a chain base index. Data about infant mortality in Czechoslovakia since 1937 (the last year before the war) can be analytically modified in this way, for example which to be calculated as the numbers of deceased children for every 1000 live births for children. By the interpretation of chain base indexes we can point to the pace of change. Periods when growth took place are easy to identify and we can also easily estimate the average annual change.

By the term "indexes" we also call indicators, which we use to process the scale data into one figure. We ask for example, about satisfaction and the respondents have to choose from a 5-point scale<sup>17</sup> (e.g. excellent - very good - good - fair - poor, or: very good - good - fair - poor - very poor). We can express the overall assessment through a calculated "index" and compare its value with the value of the index which was calculated from the assessment of another social group (assessment

<sup>&</sup>lt;sup>17</sup> The creation of verbal expression of these grades is often problematic. We should also consider possible associations which the respondent could connect with these terms. For example, "good" is, in the school evaluation, an average grade (between excellent and insufficient) and here it expresses the average (between excellent and poor), but also the grade between very good and fair. To minimize these differences we use reference to some (well-known and unambiguous) scale. In the case of 5-point scale we can tell the respondent to evaluate "as in the school". In another case, we can request to use 5 points in which 5 means the maximum assessment.

of women and assessment of men). So it is a comparison of two values which substitutes comparison of the whole assessment scale. To compare two (and often it is necessary to compare even more) scales is actually impossible.

Imagine that one of the hypotheses of the election satisfaction survey is that women are more satisfied with the correctness of an election campaign than men. We have analytically introduced the answers to this question in the table 2.

This analytical base enables us to state that male and female samples are almost equal but there are slightly more men, which does not correspond with the average in the population. This will influence the results. We can also say that the most frequent was the middle value of satisfaction and that the overall satisfaction will be positive because the frequencies in categories "very good" and "excellent" clearly prevail over "fair" and "poor". But we are not able to say if in this positive evaluation of correctness of election campaign, it was the men or the women who were more satisfied.

Table 2:	Satisfaction	with the	correctness	ofa	n	election	campaign	according	to	the
gender -	absolute frec	quency:								

Satisfaction rate							
Gender	Excellent	Very good	Good	Fair	Poor	Total	
Woman	28	20	32	12	7	99	
Man	19	37	40	16	5	117	
Together	47	57	72	28	12	216	

Table 3 Satisfaction with the correctness of an election campaign according to the gender - relative frequency:

	Satisfaction rate								
Gender	Excellent	Very good	Good	Fair	Poor	Total			
Woman	28,3	20,2	32,3	12,1	7,1	100,0			
Man	16,2	31,6	34,2	13,7	4,3	100,0			
Together	21,8	26,4	33,3	13,0	5,6	100,0			

We calculate the index which will clearly speak about this:

We match the grades of satisfaction with the values 1-5 (it does not matter if it is 1-5 or 5-1 or if we use absolute or relative frequencies). With these values we "weight" the frequencies (or proportions) in the columns and we come to the result of how many points men and women got. We use now the scale 1-5 on absolute frequencies:

Women: 28 + 40 + 96 + 48 + 35 = 247 Men: 19 + 74 + 120 + 64 + 25 = 302

Men got more points but now we must take into consideration that there were 117 men and only 99 women. We divide the points by the maximum number of points, i.e. points that would men or women get if they considered the correctness poor:

Women: 247 : 495 = 0,499 Men: 302 : 585 = 0,516

These results are the searched values of indexes. By the analysis we must take into consideration the reversed scale of values. In our case we matched the lower values with positive satisfaction, so the lower value of index will express higher satisfaction. Needed conclusion which serves for verifying the hypothesis is the statement that women were in fact more satisfied than men. By the 5-point scale the index has values in the span of 0,200 - 1,000 and the average evaluation has a value of 0,600 (we calculate these indices on three decimal places). It is also possible to assess the distance of calculated values of index from this average.

#### Means

While a layman often considers different "means" a very important and sufficient statistical data, a statistician knows about all their insidiousness and deficiencies. When a layman sees certain mean, he can think that is confronted with the informatively rich data that can replace at once all different data hidden in vast number of tables. From this it is only one step to consider means to be some kind of a magic number which is used very seriously in media often also by politicians. It is the reason why we want to show to what extent it is correct to rely on average figures.

The mean (that should be admitted) brings a certain order to the quantity of numbers - and to make things clearer is one of the basic roles of statistics and statistical analysis of social processes and phenomena. For an expert, mean is only a tool for a very simplified depiction of a phenomenon.

A non-expert usually imagines a figure to which we come after adding up the values (e.g. results of measures) and dividing this result by the number of addends. But this is only one of the means which statistics uses - the **"arithmetic mean"**. If we would use it for depiction of reality with one figure, that 100 families have a monthly income 50 000 euro, we could say that there is on the average in every family, 500 euros monthly. Inside this depiction of monthly income with one figure there can be hidden cases of substantially lower or higher incomes. The average figure hides these cases which are important to know. Such mean is only an orientation aid for arrangement of quantities and is by no means a prognosis for the case. On the basis of long-term statistical investigations we know for example that the life expectancy of 30-year-old man in Slovakia is on average another 40,9 years. But this does not tell us that all 30-year-old men (or great part of them) will die shortly before their 71st year.

In this case we unite through the mean different data and we must be aware of it by using such figure. A 30-year-old man can be a victim of a car accident tomorrow, or he can live in health till 2070. His personal destiny is included in "life expectancy" regardless of the actual course of his life. We can determine the life expectancy for a huge number of people and then it has great importance. But for an individual the importance is low.

The mean can hide rather important differences in the composition of a group of values that we have included in the calculation. Unfortunately, in such common areas as were the beginnings of production of ready-to-wear clothes we have faced the magic meaning of mean - they produced average sizes and there was no demand for it. We can rarely find ideal figures (given through average proportions). When we are shopping today, we often have the feeling that since the beginnings of production of ready-to-wear clothes very little has changed in producers' and traders' thinking. Another example: If we would like to choose an ideal place for living with an average annual temperature of, for example 12°C, we would find out that such average temperature is typical for cities like Beijing, Milan and Quito. Despite the same average temperature, the winters in Beijing are cooler than in Stockholm and summers are so hot as summer months in Rio. The mean for Peking is result of the summation and division of extremes. In Milan, the temperature fluctuations are usual for middle Europe and the annual mean is result of mild winters and mild summers. And in Quito (the capital of Ecuador) the temperature during the year is almost the same. It is the result of high altitude (3 000 meters above sea level) and the proximity of the equator. Only after looking at the values from which the same means were calculated could we choose the most suitable place for living.

On the basis of this example we can say that the mean has the greater meaning (it is more close to the reality), the smaller is the dispersion of values in individual cases (they are nearer to the axis). And contrarily with the mean and its derivatives (e.g. division of a phenomenon) we can skilfully hide the reality. For example there were the statistics of General Motors which stated that 78% of shareholders owned fewer than 50 shares and only 8% of shareholders owned more than 100. Division given this way arouses the idea of a broad democratic division of shares. But it can hide the case that the 8% of shareholders is actually only one single shareholder who would thus own an absolute majority of all shares. This would be a misuse of statistics and its trustworthiness.

For these reasons, statistics tried to find another, more real representative of the larger amount of numbers which wouldn't be so easy to misuse. Today it is e.g. "the most frequent value" - "**mode**". In many cases it is more suitable for a better understanding of value distribution. This value does not speak about the extremes - in which directions they lie and how large they are. But mode can uncover a mistake in statistical construction – informs us that we made a mistake in what we have considered together. An example is the bimodal distribution which occurs when we mix two different groups together - e.g. men and women - by the determination of the most frequent (typical) body height of the population because one most frequent value is typical for men and another one for women.

Here is another example:

We will ask the graduates of political science in 2012 what their incomes are:

Number of respondents	Monthly income in €
1	10 000
2	6 000
3	4 000 each
4	2 500 each
5	2 000 each
1	1 800
12	1 300 each

Table 4

The arithmetic mean is 2 550 (71 400 : 28). Nobody actually has such income. The most frequent value (mode) is very distant from the mean - these are the 12 who earn 1 300 each. So what now?

It would be appropriate to consider the "middle" income as the average income - according to the order in the table. We find a person among the observed 28 who earns more than the 13 poorest and less than the 13 richest - and it is the person with 2 000 euro. This middle value is called the **"median"**.

A median has an interesting advantage - we can use it where the arithmetic mean makes no sense. But this "middle value" has its disadvantages too. By small samples the middle value moves to neighbouring categories with any slight increase or decrease of the quantity of units and the value of median changes rather noticeably. But in large samples we can rely on it.

One of the features of means is that there is always something that is not expressed. It is natural because they synthesize and they cannot express everything. Every mean effaces the extremes although it is rather influenced by them. Therefore it is appropriate to use complementary figure about dispersion.

Although we have introduced three middle values (means), we can use none of them to calculate such a common average as the average speed of a car. Let's assume that we drive the distance of 30 kilometres. The first 10 kilometres we drive at a speed of 60 kilometres per hour and another 10 kilometres at 80 kilometres per hour and the last ten at 100 kilometres per hour.

The arithmetic mean is seemingly clear: 60 + 80 + 100 = 240 : 3 = 80

Median confirms these 80 kilometres per hour and we do not have the most frequent value (mode).

But in reality it took us 23 minutes and 30 seconds to complete the trek.

- 1. leg: 60 kilometres per hour ..... 10 kilometres in 10 minutes
- 2. leg: 80 kilometres per hour ..... 10 kilometres in 7,5 minutes
- 3. leg: 100 kilometres per hour ... 10 kilometres in 6 minutes

The average speed is 76,6 kilometres per hour (because 30 kilometres per 23,5 minutes).

We do not need to come to this average speed in such a complicated way but we can work with the speed directly if we use the **"harmonic mean"**:

- a. at first we count up the reversed values of speeds 1/60 + 1/80 + 1/100 = 47/1200
- b. we put this number into the denominator of the fraction, the number of measurements is in the numerator:

$$\frac{\frac{3}{47}}{\frac{1200}{1200}} = \frac{3\,600}{47} = 76,6$$

#### Standard Deviation and Variance

From what we have talked about mean, it's clear that middle values (we have mentioned four types of means) need one more dimension for a correct evaluation of their meaning. It is information about how distant the extremes or real values are from the average (which is in the case of the arithmetic mean an imaginary value). This information, this complementary value is the measure of variance and it is called the **standard deviation**.

We can explain this in the example of a shooting target. If there are 50 hits in the target, they are dispersed around its midpoint. If a good shooter shot, the distances between the shots and the midpoint are small. If a bad shooter shot, these distances are larger. In statistics we speak about small and large variance.

How do we calculate standard deviation?

We take the example of two shooters who are the same in average the first one hit 6,4 and 2 (average 4) and the second one hit 1,1 and 10 (probably by chance) with the average of 4. At first we compare every figure with calculated arithmetic mean and we find out the difference:

5 is distant from average:	2
4 is equal to average, so the distance is:	0
2 is distant:	-2

-3
-3
6

With this we have determined negative and positive deviations from arithmetic mean which we cannot count up (or we can but the result will be always 0 which we cannot use further). So we make all these figures positive figures in that we square them.

Then we calculate:

Table 5

	1. Case (squared values)	2. Case (squared values)
	4	9
	0	9
	4	36
Together	8	54

Despite the same average of 4, we come to a relatively huge difference in the calculated value: 8 and 54. We are on the right course but we must be aware that these figures (these differences) still lie - the differences will be the larger, the more shots (values of hits) we have. To apply the knowledge about the number of measures (shots), we divide the acquired figures by this number of measures. With this we get the first measure of variance which we label with "var X" or sigma squared.

var 
$$x_1 = \frac{8}{3} = 2,67$$
 var  $x_2 = \frac{54}{3} = 18$ 

Once again, we stress the importance of this calculation: firstly we exclude the negative values and secondly by squaring them we emphasize the extent of deviations.

To get a figure which would enable us to say that the average distance of shots was so and so, we must extract the root of these acquired values. That is the measure which characterizes these values substantially better. After extracting the root of our calculated values of dispersion we get the **standard deviation.**<sup>18</sup>

- 1. The square root of 2,67 is 1,63
- 2. The square root of 18 is 4,24

We must match both these figures with the already known average (middle value) - then they express in the first case that:

- most figures (values) deviate from middle value by fewer than 2 in both directions - so they lie in the range from 2 to 6 in the second case:

- most figures (values) deviate from middle value by more than 4 in both directions - so they lie in the range from 0 to 9

We used only three figures for the simplification. But imagine that there are 50 of them, or  $1\ 000$  - here the standard deviation offers us a

<sup>&</sup>lt;sup>18</sup>Since "var x" is actually sigma squared, sigma is the designation for standard deviation

good image of the concrete sample. So we do not have to observe all individual figures (findings) but only two measures of data are enough (mean and standard deviation) and we know what the situation in the sample is.

If we put these two data into a fraction (standard deviation is in numerator and mean is in denominator) and divide by 100, we get the coefficient of variation:

1,63/4 x 100 = 40,75% 4,24/4 x 100 = 106,00%

A very rough rule for reading this figure is that if the coefficient of variation is higher than 50% it means strong heterogeneity of the sample. Such heterogeneity excludes in effect the application of an arithmetic mean as a representative value. (Swoboda, 1977)

## Statistical Dependencies

In statistics we use the term "dependency" instead of "relationship" so in cases where we will observe the relationship between two variables, we will speak about statistical dependency. It is not only a formal requirement because what is important is that here we have dependencies between large-scale phenomena. In statistical analysis we do not think about individual cases but about statistical series of certain things, phenomena, processes, events. Since the basic feature of a statistical series is variability, the dependencies between them will also be variable. Statistical dependencies are based on probability theory and the relationships are stochastic (probabilistic). (Swoboda, 1977)

**Example:** Is there a statistical dependency between income and consumption of beer?

Some observations show that there is a dependency, other observations do not. For example, A has an income of 800 and drinks half a litre of beer per day and B has an income of 900 and drinks one litre a day. From these observations we could conclude that there is certain dependency. But there is also C who has the income of 700 and drinks 3 litres of beer per day. His income is lower but his consumption of beer is higher.

We can come to an exact conclusion only through statistical data about the whole population or data from a representative sample which fulfils the conditions of representativeness. We compile two series of data from our given data - there will be inhabitants ordered according to the height of income in the first one and according to the quantity of consumed beer in the second one. Then we study the dependency between these two series. If we confirm a statistical dependency, we can interpret it only as probable - it is valid only for the whole, not in individual cases. On the basis of the inhabitant's income we can estimate his behaviour only with certain probability - in this case drinking beer. In general, we can say that a specific structure of the population in a given area (e.g. according to income) is a condition for a determined behaviour.

Restrictions of statistical dependencies:

- 1. These are samples which are defined in terms of time and place and therefore also restricted in terms of time and place (they do not have to be valid in another time and place).
- 2. It is necessary to validate the discovered dependencies they must undergo the control of validity either by comparison with results from other similar research or by repetition of research. Only after such confirmation of dependencies, we can consider them relatively permanent or regular and they can be used for the creation of partial theories. It is a relatively time-consuming procedure and that's the reason why some sociologists are sceptical towards the possibilities of empirical cognition. Discovered dependencies are then used only in an application area, i.e. for practically usable conclusions.
- 3. Since there are dependencies between complex phenomena, we can assume that dependencies are as well of complex character they are blocks of factors and many of these factors are hidden for us. Statistical dependencies depict only partial aspects of mutual relationships between phenomena. Sometimes they are only

superficial - they are evident at first sight but they do not have to be important. Statistics offers for example **factor analysis** which respects the fact that these are dependencies of many variables and it is able to order them according to importance and choose the most important factor.

4. In discovered dependencies the "spurious" variables can be found. If we are not able to exclude them, they can significantly distort the results (e.g. relationship between residence size and divorce rate - but M. Disman (1993) points out their different forms and relationships).

#### **Dependencies of Nominal and Ordinal Variables**

To find out the dependency between two nominal variables means to answer the question whether the distribution of frequencies according to changes of variable variants A is connected with the distribution of frequencies according to changes of variable variants B. The simplest case is when each nominal variable has only two alternatives (e.g. yes-no). A complete table of absolute frequencies with four fields has this form:

Ta	h	e	6.
ıα	v		υ.

Variable A	Variable B		
Valiable A	Yes	No	Together
Yes	а	b	a+b
No	С	d	c+d
Together	a+c	b+d	a+b+c+d

We call the frequencies a+b, c+d, a+c, b+d "marginal" frequencies. If any of the observed variables has more alternatives, we use a table for the recording of their values (recording of values of one variable) which we call a "contingency" table. A table which expresses the relationship between values of one variable and values of the other is called a "correlation" table. Measures of dependencies of these nominal features are test X<sup>2</sup> (chi-square), Fischer's test and specific measures of dependencies. These tests do not show the level of dependency between two qualitative variables but by means of them we can confirm whether it is true that there is no dependency between them. Our starting point is the assumption of independency between observed variables. If we come to  $X^2$  which is higher than in the table, null hypothesis about independency is rejected and the dependency between two observed qualitative variables is proved. For the calculation of  $X^2$  we must work with a representative sample which is large enough (by the table 2x2 it means more than 40 units, or to reach that the theoretical frequency is not lower than 5 in any field).

**Example:** Is there a dependency between listeners' satisfaction with a certain programme and their achieved education? There are four variants (categories) of education and three levels (categories) of satisfaction. The results are in the following table:

Level of		Satisfaction with programme				
education	Satisfied	Undecided	Dissatisfied	Total		
Primary	22	15	133	170		
General secondary	25	11	74	110		
Specialized secondary	20	8	42	70		
University	30	7	13	50		
Together	97	41	262	400		

Table 7:

For the particular fields in table the theoretical values, differences towards real values and quadrates of these differences are calculated. Finally these quadrates of differences are divided by theoretical values. We get a value in every line and when we count them up, the result is the searched testing characteristic X<sup>2</sup>. We compare this value with the table (critical) value by degrees of freedom v = (m-1). (n-1) and if it is higher, the null hypothesis about independency of variables is rejected.

i	j	fo-ij	fij —f 0-ij	(f <sub>ij</sub> - f <sub>0-ij</sub> ) <sup>2</sup>	<u>(f<sub>ij</sub> — f <sub>0-ij</sub>)²</u> f <sub>0-ij</sub>
1	1	41,225	19,225	369,600625	8,965
	2	26,675	1,675	2,805625	0,105
	3	16,975	3,025	9,150625	0,539
	4	12,125	17,875	319,515625	26,352
2	1	17,425	2,425	5,880625	0,337
	2	11,275	0,275	0,075625	0,007
	3	7,175	0,825	0,680625	0,095
	4	5,125	1,875	3,515625	0,686
3	1	111,350	21,650	468,722500	4,209
	2	72,050	1,950	3,802500	0,053
	3	45,850	3,850	14,822500	0,323
	4	32,750	19,750	390,062550	11,910
Together		400 000	0.0	_	53 581

Table 8:

Table (critical) value  $X^2$  for § = 0,05 and the number of degrees of freedom where v = (4-1). (3-1) = 6 is 12,592. Because the calculated value is higher, the null hypothesis about independency between listeners' satisfaction with a certain programme and their education is rejected. It is obvious from the course of theoretical frequencies that the largest differences are between the extreme groups of education. Respondents with primary education were more often dissatisfied with the programme than would correspond with a theoretical distribution of frequencies.

Respondents with a university education were more often satisfied with the programme.

Note - correctness check:

- a. the sum of theoretical frequencies must be equal to the number of observations (400)
- b. the sum in the next column (if we record values + and -) must be zero

## **Dependency of Ordinal Variables**

We replace the real measured values by their order which means that we compare two orders. By Spearman's rank correlation coefficient we proceed this way:

- a. We order all the values, which both variables acquire, according to their order from the lowest to the highest and we match them with order numbers from 1 to n. So every observed element is matched with two order numbers.
- b. For every pair of values which belong to one element, we calculate the difference of orders. We will move between two extreme cases: the first is when both orders are the same and the differences in orders are equal to zero, and so are their squares and the whole fraction of the expression R (R = 1 minus fraction). Then R = 1. The second extreme case is when both orders are opposite to each other. Then the fraction has the value 2 and expression R = -1. If both series of order numbers are independent of each other, rank correlation coefficient is R = 0.
- c. Then we calculate the quadrate of these differences and we count up the results of all elements.

Example: The number of people dying in medical institutions has increased. They looked for the cause and hypothesis were formulated that in districts where there are more beds per 1 000 inhabitants and therefore also a greater probability to hospitalize the incurably ill, there will also be a higher percentage of death in these institutions.

Districts	Number of	Order xi	% of the	Order yi	Difference	di <sup>2</sup>
	beds per		dead in		di=xi-yi	
	1 000		institutions			
	inhab.					
Ι.	8,50	6	40,9	6	0	0
П.	9,24	4	46,7	3	1	1
III.	10,12	3	34,6	9	-6	36
IV.	5,78	9	42,4	5	4	16
V.	7,73	8	42,8	4	4	16
VI.	11,49	1	40,2	7	-6	36
VII.	8,51	5	47,0	2	3	9
VIII.	7,94	7	36,6	8	-1	1
IX.	10,78	2	49,2	1	1	1
Х.	4,00	10	31,6	10	0	0
Together						116

Table 9:

$$R=1-\frac{6\sum_{i=1}^{n}{d_{i}^{2}}}{N(n2-1)}=1-AR\sum$$

$$R=1 - \frac{6,116}{10,99} = 1 - \frac{696}{990} = 1 - 0,703 = 0,297$$

On the basis of the value of Spearman's coefficient we can claim that there is a dependency between the number of beds and the rising death rate in institutions but it is very low. It is necessary to search for another explanation for the rising death rate in institutions.

## Large Scale Phenomena and Law of Large Numbers

Statistics is based on the knowledge that large-scale phenomena are fundamentally different from individual and typical phenomena. According to this we distinguish between statistical and individual approach and statistical and monographic procedure. Let's show the differentiation of large-scale phenomena and individual cases in some examples:

X got hepatitis. It is an individual case. It has many individual features. The environment where X lives, also has many characteristic features. We assume that his illness has a relationship to these characteristics. A statistical approach considers this case one of many, similar cases. Of course it is necessary to determine what connects these similar cases - what characteristics they must have to work with them as with a statistical sample. When this sample of similar phenomena (cases with hepatitis) is defined and separated from other phenomena, then a statistical approach searches for relationships to the rest of the population. Statistics is interested in the proportions of those similarly ill in the population of some territorial units (what is the proportion of people with hepatitis in particular regions in Slovakia). The result of this comparative analysis can be a statistical indicator which shows the intensity of the phenomenon in the whole population. The individual phenomenon (the reality that X got hepatitis in the Banská Bystrica region) can be something quite normal from the point of view of occurrence or something abnormal (if there is no hepatitis in other regions). We see that statistics uses the terms of normality and abnormality.

Another example: One day, we buy certain magazine. Although it may be unusual in **our behaviour**, from the point of view of **saleability of the magazine** the usual number of sold magazines does not necessarily deviate - it is close to the average. But it might also be the opposite case and our behaviour might be part of a broader social phenomenon, some extraordinary wave of increasing sales of that magazine. The reasons then aren't individual (our current idea) but have social character because other people act similarly in different places - it is a large-scale phenomenon.

Another example shows other aspects and possibilities of use of a statistical viewpoint: our neighbour usually drinks three beers a day. Statistics studies this phenomenon as a part of a large-scale phenomenon which we call consumption of beer in Slovakia. It states how much beer is consumed a day and then it is easy to calculate how much it is for one inhabitant of Slovakia (let's say inhabitants over 18 and of both genders). It can be, for example, one beer for one inhabitant. Then we can put the individual case of our neighbour into this large-scale phenomenon and we
can see that this phenomenon isn't normal in relation to the figure calculated from the entire statistical sample - the neighbour exceeds three times the daily consumption of beer for one person. But if we could choose only those who sometimes drink beer from the sample, we could see that the neighbour's three beers correspond with this new average. If we chose only those who drink beer daily and are male, the neighbour could be under average.

And another example of statistical perception of the world around us: a doctor deals with individual cases and he diagnoses and proposes a therapy. These cases are, of course, wholly individual but the doctor can rely on statistical data of large-scale character which are actually records of all individual cases of a given illness. From them he finds out what the typical course of the illness is, in which environment it occurs, what treatment is the most effective, what the possible side effects are etc. The doctor must transform this information in the individual cases and **the way** he does this, is the key to his success.

These examples show clearly that statistics enables us to see our individual cases in context through observation of large-scale phenomena. It means that we are able to assess and evaluate the meanings of our individual phenomenon in a broader sense. Let's stress that large-scale phenomena that we can observe statistically, are connected with human behaviour (shopping, consumption, culture, interests ...), their membership in various social groups (they are inhabitants, voters) or their opinions of various problems around us that we observe through public opinion polls.

It is necessary to say that all statistical data are related to certain territory. It is ideal when it is a territory in which the given statistical feature occurs in as homogeneous quality as possible. Then the average and other statistical values are least confusing and most of the individual cases are getting closer to them. Statistics therefore uses such administrative territorial units as is the basic statistical unit - statistical area, residence, district, region, state.<sup>19</sup> The recorded statistical data must

<sup>&</sup>lt;sup>19</sup>From this point of view and for these needs the cancellation of some territorial units means the loss of certain cognitive levels and frequent changes in their definition make it impossible to compare the data (e.g. for districts) backwards - we lose the opportunity to observe the (historical) development of given data in time.

always be definitely assigned to specific territorial units. Without this we cannot analyse them. Besides this territorial categorization, the time dimension is equally important for the analysis of statistical data.

French mathematician and statistician Pisson defined the law of large numbers this way: phenomena of any kind are subordinate to one common law which we can call the "law of large numbers". This law tells us that by observation of a large number of phenomena of the same kind that are dependent on irregularly variable causes, we can find numeral ratios which are almost invariable. These ratios have a particular value for every phenomenon. The empiric data are getting the closer to it, the larger number of phenomena we observe. If we could broaden the number of observations endlessly, we would get this value.

According to the law of large numbers the statistical regularity or rule expands towards an endless number of observed cases. The curve which expresses the approximation of empiric data to theoretic figures expressing the value of phenomenon, reaches by a certain number of cases the first turning point from which the regularity grows and the information begins to have importance. It is still not stable regularity, but we use this knowledge for the determination of a number of cases for pilotage (through which we verify the quality of techniques of sociological research - e.g. the comprehensibility and order of questions in a questionnaire).

The curve comes to the second turning point after the previous rapid growth of regularity with an increasing number of cases. At this point the curve is getting balanced and the growth of regularity by the same growth of number of cases is no longer so considerable. So in this point we reach a relatively high stability of regularity, or statistical rule by a relatively low number of cases. Here we find an optimal number of statistical selection which has the required accuracy by acceptable number of observed cases.

The law of large numbers is closely related to the term "coincidence" which is the base of the probability theory. We have used in this text the terms "random" or "probabilistic" to mark different ways of sampling. But we should distinguish this "randomness" from the "randomness" of results of individual observations or of a small number of observations which is more likely equated with the term "indefiniteness" or also "entropy" known from the information theory.

# Understanding Qualitative Methodology in Social Sciences

In his publication, Miroslav Disman (1997: 283-322) calls the chapter which is devoted to qualitative research "The Other Research or To Prove or Understand?" In this part we will talk about that "other" research which has, according to Disman other goals and rules than quantitative research. He summarizes the basic difference between quantitative and qualitative research: while quantitative research offers us a limited extent of information about many cases, qualitative research offers us lots of information about a very small number of individuals. (ibid)

In other words, quantitative research is a very useful tool for cognition of social reality but the price for "as reliable and representative as possible results" is often too high. Reduction of information which is necessary in quantitative research means the loss of further information which can be an important aspect of social reality. Quantitative research tries to reach the maximum possible objectification which means ignoring the uniqueness of each case and a reduction of the observed field. Qualitative research tries to grasp the subject of its research in its uniqueness and in its natural environment. On the other hand, it can happen that the observed uniqueness will not depict the side of problem which prevails in the population.

Let's take a questionnaire as an example. As a result of reduction of content it can happen that we lose the subjective meaning to which the respondents attached their answers. Quantitative research does not enable us to understand why the respondent answered in the way that he did and at the same time the standardized schema of questionnaire "forces" respondents to fit their answers (a statement about their experience, reality and life) into the predefined categories. Consequently, there is a shift between how the researcher registers the reality and the reality itself.

Imagine the situation where in public opinion polls in which 1 200 respondents took part, 30% respondents expressed confidence in the government, and 70% respondents did not trust the government. If this information is put into a time series, i.e. we know if confidence has increased or decreased, it is certainly useful and our attention is drawn to a particular problem. Then the experts have opportunity in the media to comment on the problem and to think about the reasons why confidence in the government has decreased. They do it by including various macrosocial factors or they analyse specific steps of the government or

behaviour of politicians. 70% of respondents have something in common in this case. All of them were inclined to answer "I do not trust" regarding the two or three possible variants.

Despite this common feature, there are still many differences between them. The first one might be the different extent of this noconfidence. One absolutely does not trust the government, the other one hesitated and finally ascribed to this option. This incompleteness of information can still be solved within quantitative methodology - for example by using a scaling method through which the respondents can express their extent of no-confidence. However, each of these respondents has his "own story" about why he does not trust the government and the difference is not only the extent of no-confidence but also the factors responsible for this no-confidence. Here we could object that it is no problem to create a questionnaire which would be aimed at revealing the factors responsible for no-confidence in government. But the researcher would have to create a typology of these factors which he would probably do on the basis of his own assumptions, knowledge or other similar researches. This research, if it were done correctly, would certainly help to advance our cognition of the problem of no-confidence in government. We would certainly know more than before but we still wouldn't know to what extent the researcher's categories corresponded with those of the respondents. We wouldn't know if there were some other unexpected factors and we wouldn't be able to grasp the mutual relationships of these factors and the meaning which the respondents attached to them.

Qualitative research should eliminate this problem. This research works with **emic categories** - categories which are used by people themselves, they are "their" concepts with which they describe the world around and the observed problem. At the same time, qualitative research gives the respondents freedom to formulate the relationships between these emic categories, to formulate their story leading to the reasons "why I do not trust the government". The consequence of the richness of information in this case is abandonment of the requirement of representativeness which is not the aim of qualitative research. This research would be realized on a smaller sample by which we would be interested not only in the internal mutual relationships of individual categories but also in regularities in the stories of the individual respondents. The result of qualitative research could therefore be, for example, typology of the "distrusted", of course without the aspiration to representativeness.

It must be emphasized that although the preference of quantitative or qualitative research is often paradigmatic - i.e. there are the exclusive supporters of the one or the other type of research in particular theoretical orientations and schools, even workplaces - our attitude is that the choice between quantitative or qualitative research is mainly the **function of research intention** and depends on the problem we want to study and on what we want to find out. It depends also on other practical circumstances and possibilities (financial means, number of professional staff, amount of time). But between these two types of research there isn't a contradictory relationship and their usage can be effectively **combined** and they can **complement each other**.

We can get back to the previous case and imagine a researcher who decides to actualize the research of the factors of no-confidence in government through the version of a quantitative questionnaire. He/she wants to formulate a series of factors which would be the base of the research. If he/she had the results of a qualitative research on similar topic, they could be useful and help to put together a series of facts and subsequently to put together questions (especially the closed ones). Researcher could even have the objective to verify the validity and occurrence of individual types of the "distrusted" revealed in the qualitative research - now on a broader sample and through one of the techniques of quantitative research. The situation can also be reversed. The results of quantitative research can be the base or impulse for qualitative research. They may indicate some social problem. In our "imaginary" qualitative research of distrust to the government, there were the results of quantitative research at the beginning of our interest, they attracted the attention of our imaginary researcher to such extent that he has decided for deeper study of the problem using particular method or combination of methods of qualitative research.

We can complete and summarize our previous reflections about qualitative and quantitative research this way:

Quantitative research is representative; its results are reliable although they have only a probabilistic character. However, quantitative research reduces information and is thus characterized by high reliability but low validity. Qualitative research, on the other hand, deepens our cognition and understanding of a particular problem, it works with emic categories (which are made by the respondents' themselves). But the conclusions of qualitative research are not generalizable on entire population.

It is important to emphasize that while quantitative research is aimed at the testing of hypotheses and pays attention only to data connected with the tested hypotheses, proceeding basically deductively; qualitative research tries to pay attention to all data and to capture any regularities and relationships that can be sociologically relevant even if the researcher didn't expect their occurrence. It is a more inductive approach to data.

Silverman (2000: 19) characterizes qualitative research this way:

- 1. Qualitative data are preferred in a simplified way it is rather the analysis of words and images than numerical figures.
- 2. Naturally occurring data are preferred rather observation than experiment, rather unstructured than structured interview.
- 3. Meanings are more preferred than behaviour attempt to document the world from the point of view of observed people.
- 4. Natural sciences as a model of conducting scientific cognition are rejected.
- 5. Inductive research connected with generation of hypotheses is more preferred than the testing of hypotheses.

Silverman summarizes also the criticism of quantitative research; the arguments emphasizing the difficulties of quantified investigation in social sciences are often used to stress the importance of qualitative research:

- 1. Quantitative research can lead to "fast fixation" with the consequence of weak or no contact with people or "field".
- 2. Statistical correlations can be based on "variables" which are arbitrarily defined in the context of naturally occurring interactions.
- 3. Processes of common thinking which the science tries to avoid, can then be subsequently employed in speculation about the meaning of correlations.

- 4. The effort to get "measurable" phenomena can lead to creeping of unregistered values in the research only because we work with highly problematic and unreliable concepts such as "delinquency" or "intelligence".
- 5. Despite the importance of testing hypotheses, the purely statistical logic can transform the developing of hypotheses into trivial issues and can fail in generating of hypotheses from the data.

# Philosophical and Theoretical Sources of Qualitative Methodology

Very important in qualitative research is **its interconnection with theory.** The interconnection of research and theory is, of course, not the privilege of qualitative research, it relates to social-scientific research as a whole but in qualitative research this problem is more urgent. Because qualitative research does not provide representative results, its scientific contribution is emphasized in relationship to the theory which it verifies, develops or directly generates.

Without theoretical reflection of the problem the sociological cognition is very flat and there is a danger that assumptions and starting points which are behind this cognition remain hidden and un-reflected. The interconnection of theory and methodology is in the case of qualitative research closer because the danger of lapse into "abstracted empiricism" is lower - the methods used in qualitative research are more difficult to transfer from one problem to another without deeper reflection.

Drulák (2008: 19) points out that while quantitative research is based on clearly structured methods providing mostly technically exacting but clear-cut research instruction, qualitative methodologies are in greater extent dependent on the research context and their application is almost always a creative act. For this reason, qualitative research is in a greater extent influenced by the subjective abilities and ideas of the researcher. While quantitative research is defined by the testing of hypotheses, what qualitative research "cannot do", qualitative research often leads to the formulating of new theories or new hypotheses an thus develops the sociological knowledge.

Close interconnection with theory is also related to the development of qualitative methodology, to the fact that they were not as generally accepted as quantitative methods and therefore they were developing in dependency upon concrete sociological orientations and schools. Qualitative methods in social sciences are perhaps as old as quantitative approaches, however quantitative methods dominated for a long time and qualitative methodology begins to assert itself in social sciences in the second half of the 20th century only gradually. The long dominating positivist approach which was connected with the use of specifically quantitative methods had its opponents from the beginning.

Let's get back shortly to the origin of social sciences (or human sciences as they were called). Originally, philosophy comprised the unity of all knowledge from which, as a result of extending of knowledge, particular areas thereof gradually began to separate and specialized sciences began to arise. At first it was mathematics, astronomy, physics and then other sciences. In this sense, we can talk about social sciences as about relatively young sciences. Most of them "became independent" of philosophy only during the 19th century. Social sciences deal with problems that were of interest long ago (e.g. by Aristotle or Plato), however, systematic research of these problems began with the gradual institutionalization of social sciences. Many authors connect the origin of social sciences (mainly sociology) to the fast changes in the modernizing societies. The quickly changing character of social phenomena and processes asked for their systematic reflection. Of course, there were other factors that also contributed to their development - e.g. decrease of the influence of religion in public life or new political ideologies.

As soon as the social sciences were constituted, researchers ran into problems with the use and the application of methods which so far had been used in natural sciences, now being applied to social sciences. There were several reasons, but the most serious seemed to be the complexity and variability of social reality which was very hard to fit into general rules and propositions. This complexity of social reality lies in the greater number of variables which play a role here and in the dispersed structure of social phenomena, then in dynamics of social processes and last but not least in the fact that social reality is constantly shaped by the activity of individuals on the basis of their free decision making. Because of this character of social phenomena and processes, a question arose whether there were any social rules which were repeatable and stable or the social reality is shaped by the tangible activity of free individuals. This problem is not only a methodological problem but also an epistemological and theoretical one and according to some authors (e.g. Giddens, Bourdieu) it is the basis of paradigmatic division in sociology.

Human Sciences as Sciences about Man and Culture: Wilhelm Dilthey and Max Weber

The approach according to which it is necessary to build the social sciences on a basis of natural sciences (which also means to prefer using of quantitative methods in social-scientific research) is called positivism. A key opponent of the positivist approach in social sciences was **Wilhelm Dilthey** (1833-1911).

Dilthey claims that the role of human sciences is to understand life, the human and his history. He understood history in compliance with German historicism as the history of unique and unrepeatable products of human spirit which we cannot reduce to material processes, to a summary of objective facts and we cannot derive them from abstract metaphysical principles. To understand the historical reality, the experience and ability to empathize is inevitable. Rational analytical procedure leads us to knowledge which is exact but only outer and on the contrary the repeated experience enables us to understand life from the inside. In this sense, Dilthey created an important argument in favour of the qualitative approach to the social science research. Qualitative methods are decisive for social-scientific cognition not because social sciences are not able to reach the level of natural sciences and formulate general rules, but exactly because this understanding cognition is more suitable and fruitful in the case of social phenomena.

Dilthey even turned the positivist argument when he claimed that we cannot understand nature and he showed how unsuitable the opposite situation would be if someone would try to apply the procedures of social sciences on natural sciences and would try to "understand" natural phenomena subjectively. According to Dilthey the difference between human and natural sciences is therefore crucial. The basis of naturalscientific method is the reduction of individual unrepeatable phenomena to general natural laws and it wants explain them this way. Human sciences, on the other hand, want to understand individual unrepeatable phenomena from the inside as products of spirit. They interpret their meaning in certain whole and they reveal the meaning of the whole on the basis of understanding of meaning of its parts. **Max Weber** (1864 – 1920) followed in Dilthey's definition of human sciences. Weber had to deal with the tradition of German historicism which proceeded from the fact that every event in the history of mankind is unique and unrepeatable and cannot be subordinated to some universally valid law. It was difficult to build up a new science of society - sociology - on this principle. He copes with this tradition with the help of Neo-Kantians - Windelband, Rickert and Dilthey - but he builds his own position. Neo-Kantians influenced him especially by the effort to give a general methodological explanation to human sciences because they cannot use the procedures and methodological apparatus of natural sciences. Here by the analyses of human behaviour, it is important to keep in mind that this action is determined by a goal and this goal is related to certain value - so the action is conditioned by value. It follows that it is necessary to study also certain **inner experience by means of understanding.** (Keller, 2005)

In his book Economy and Society (Wirtschaft und Gesellschaft) Weber defines sociology as a science based on "understanding" and trying to analyse the meaning of **social action** through interpretation. Under social action Weber understands action oriented to others. So such action that is "according to its subjective meaning attached to it by acting individual related to behaviour of others, is thereby oriented in its course and this behaviour can be clearly explained from this subjectively intended meaning," (Weber, In: Novosád, 1997; 22) Social action is determined or codetermined by expectations that another human or humans will behave in a certain way. (Novosád, 1997: 22) Sociology is interested exclusively in this type of action, social action which is in its course oriented on other people. Sociology must explain all social types of state, e.g. feudalism etc., as products of the action of involved individuals as a manifestation of **socialized action**, when the actors are oriented to certain general rules which regulate their behaviour. Sociology must be able to reconstruct the meaning that people give to their action (not only the level of the individual, like psychology, but also on the level of group).

If we want to understand the motives of a group - the best way is to understand a typical representative of this group - someone who represents the **ideal type** - and we describe it in detail and then we can compare it with the reality. Weber speaks about ideal type as a basic methodological tool of social sciences. Ideal type arises by one-sided accentuation of one or more traits of a phenomenon. But he warns against the mixing of ideal type with reality - he says that this is what the economists in particular do (supporters of free market but also Marxists).

In his methodological papers Weber tries to interconnect certain elements of German historicism and principles of positivism but rejects weak places of both theories. He rejects the positivist conception according to which the cognitive procedures in natural and in social sciences are identical and he also rejects the doctrine of German historicism according to which it is not possible to do any generalizations in the area of culture and history. According to him science must always strive for an abstraction or generalization. In the area of human sciences it is necessary to build up **intersubjectively verifiable cognition** - but he does not agree with the thesis that this action is irrational and unpredictable. The basic principle of his understanding of sociology is the possibility to explain human action through the understanding of it, which means that we match observed action with established inner motives which we can reconstruct from this action.

#### Phenomenological Sociology of Alfred Schütz

Weber's ideas in the area of methodology of social sciences were forgotten for a period of time after his death. Interest in Weber revived with the development of the interactional approach in sociology, in the paradigm of phenomenological sociology. But the ideological source of phenomenological sociology is not only Weber's conception, but primarily phenomenology as specific orientation in philosophy. Its founder Edmund Husserl tried to bring the world of objectified scientific interpretations back to life-world, which means the world of everydayness as the actors people - perceive and interpret it. The goal of phenomenology is to get to know the world as it really is, to uncover regularities in its experiences and so reach the objectivity of cognition. The phenomenological orientations in sociology are based on this philosophical starting point.

**Alfred Schütz** (1899 -1959) in particular tried to bring Husserl's thoughts into sociology. The basic goal of phenomenology in sociology is to search for the answer to the question of what makes up the term entitled **lifeworld** and how can we get to know it? Lifeworld is the world

of common sense and of so called natural attitude. We inhabit this world together with other people and it is therefore a shared and intersubjective world. The basic structure of lifeworld and basic knowledge about it are common for all of us. We draw our whole experience from lifeworld and through this experience also all the meanings of this world. The lifeworld is by its participants considered natural, unproblematic and familiar. At the same time, the lifeworld makes up horizon of our ideas about it. But in this world the individual behaves actively, and he intervenes in it this way and changes it. To act practically in this world (e.g. solve problems) the individuals must understand the world. So they must have a certain stock of knowledge. It builds the framework to which the actors relate all their interpretations. A stock of knowledge is the base for how to deal with things, how to cope with various life situations and to communicate with others. The contents of this stock of knowledge are standardized interpretations and orientations - they enable us to use predetermined learned and proven ways of behaviour and action in other similar situations. People do not have to think about every situation of their everyday lives over and over again.

People are engaged in an on-going process of making sense of the world, in interaction with their fellows and we, as scientists, are seeking to make sense of their sense-making. In doing so, we must inevitably make use of the same methods of interpretation as does the person in his or her 'common-sense world'. What distinguishes the social scientific enterprise, however, is that the social scientist assumes the position of the disinterested observer. He or she is not involved in the life of those observed - their activities are not of any practical interest, but only of cognitive interest. (Wilson, 2002)

#### Symbolic Interactionism

Symbolic interactionism is one of the orientations of phenomenological sociology. It is such view on social reality that understands it as a result of each interpersonal interaction realized by symbols. The author of the term is Herbert Blumer. Symbolic interactionism proceeds from the philosophy of pragmatism and was influenced also by psychological behaviourism. In its centre there is the specific view on the human as a creative and active individual who creates social reality through his actions. Vladimír Drozda emphasizes that the base of social reality in symbolic interactionism is not an individual, but the relationships between individuals. "We must understand the human as a product of interaction in society, not as a source that creates this society." (Drozda, 1996: 13)

The founder of symbolic interactionism is George Herbert Mead (1863 – 1931) and especially his work "Mind, self and society"<sup>20</sup> in which he formulated key theses of symbolic interactionism. Before the formulation of the interactionist conception Mead was devoted to the detailed analysis of the human mind and he came to the conclusion that subjectivity is an inseparable part of the mind. And "because subjectivity is related to the mind of the individual, to that phase of mind that is characteristic for the individual as such, it must have a place in the process of reflection" (Drozda, 1996: 18). Through a detailed analysis of what is mind and what is mental, Mead prepared his groundwork for creation of the interactionism concept. According to Mead, to consider an action conscious it is necessary that gesture becomes a symbol. Mead adopts the basic general idea about what gesture is, from Wundt: It is the "part of the social act, so that he takes the attitude of the other person who cooperates with him" (ibid). Unlike Wundt he noticed that the emotional or intellectual attitude given to the gesture by its sender towards the addressee, does not have to be reproduced in it. Each of them perceives it subjectively. And the gesture (Mead used the term gesture also for word) that is a symbol - it means that people agreed on certain meaning enables this meaning to be understood by all participants of the interaction. Symbols provide us with the interpretation of meaning, which the sender has attributed to his gesture or act, to its recipients who then react on the basis of this interpretation.

This process of symbolization is at the same time the process from which the human "self" arises. "Self" is formed in the process of interaction and "social process has a priority in the relationship to mind" (Drozda, 1996: 28) - individual "Self" can be explained within and from social process but we cannot explain the social process or interaction

 $<sup>^{\</sup>rm 20}$  This book was not compiled by Mead himself; actually it is the collection of his lectures published by his students after his death.

through understanding of "self". "Self" isn't here by the birth of the individual but is formed in the process of socialisation. The awareness of our "Selves" arises in us on the basis of the reactions of others towards us. Mead thinks that the human individual has his own experience as "Self" "only in so far as he/she first becomes an object to himself just as other individuals are objects to him... and he becomes an object to himself only by taking the attitudes of other individuals toward himself within a social environment in which both he and they are involved." (Mead, In: Drozda: 36) Mead calls this "taking the role of a particular other". To understand social life - its specificity, and the reason why people live in large societies we must go further to Mead's concept of the "generalized other".

How the ability to take over the role of "other" and then of "generalized other" is created in the process of socialisation is given support by Mead through his analysis of a "play" and a "game". While the first one is an unorganised playing and imitating of the role of someone else, the second one is an organised, regulated game in which it is important to cope with more varied but precisely determined roles. It is characteristic also for life in society. The term "generalized other" also denotes a process. And "this process provides us the identification of individuals within the social groups" (Drozda, 1996: 38). With this, Mead came from analysis of human consciousness to the study of life in society.

The term institution plays the role of intermediary between theory of consciousness and theory of society. The formation of institutions is caused according to Mead by process of social control. We mustn't understand the term control only in the sense of coercion but also in the sense of creating the life of individual. Institutions are therefore sort of visible form of "generalized other" and are as important for the life of individual as anything else. (ibid: 43).

An important successor of Mead's ideas was **Herbert Blumer** (1900 – 1987). Blumer tried to find sociological sense of Mead's work - of his philosophical approach to study of society. According to him, Mead points out what most sociological theories do not take into consideration - and this is the regard to activity of human "Self" by explaining an activity. According to Blumer, these theories try to explain the action from the point of view of factors that determine it and not from the point of view of human.

As we already mentioned, Blumer is the author of the term symbolic interaction and this term served him for several things. By means of it he explained mutual adaptation of behaviour in human community and also relationships such as cooperation, conflict, dominance, exploitation, consensus etc. But Blumer's fundamental question is the explanation of joint action that is for him a starting point for thinking about forms in which people live their lives. His view on joint action emphasizes that the base of society lies in a constantly running process of activity and not in a postulated structure or relationships. Then he emphasizes that joint human action has on the one side its history that determines conditions for this action, but on the other side an important characteristic of joint action is its openness - the fact that it can be re-identified and transformed by active influencing of human individuals. (Drozda, 1996)

Blumer from his point of view on human society, symbolic interaction and activity draws several methodological requirements that sociological theory should fulfil:

"a) a processual treating of social factors, b) the need to integrate human "Self" into scientific research, c) the requirement of complexity and compatibility of particular factors in interaction and d) the empirical world is a criterion for truthfulness of our conceptions, not a required shape of theoretical models" (Drozda, 1996: 64).

Symbolic interactionism is a still developing paradigm, especially in the USA. Today it is also called "an interpretative paradigm" which is a term invented in 1982 by T.P. Wilson and which is used as the counterpart of a normative paradigm. Symbolic interactionism also stimulated the formation of other orientations that partially share certain basic starting points with it - especially focusing the attention on reality as a formed interaction of individuals, but they are partially singled out from it. We speak about the ethnomethodology or about the approach of Erving Goffman which he himself called "dramaturgical sociology". These are sometimes seen as parts of symbolic interactionism and sometimes as separated sociological orientations. Erving Goffman's Dramaturgical Sociology and Herold Garfinkel's Ethnomethodology

Erving Goffman (1922 – 1982) didn't consider himself a symbolic interactionist. He was interested in the microsocial level, in the study of everydayness and primarily he worked with the concept of roles which he developed in a very original and inspiring way. Goffman's metaphor which compares situations of everyday life to acts in the theatre is an understandable explanation of the principles of human behaviour in everyday life. In his work "The Presentation of Self in everyday life" he proceeds from the idea of an individual "trying to induce a certain impression with which he would influence how others define the situation. It is therefore in every individual's interest to control the behaviour of others - especially their reactions to him." (Šubrt, 2001: 73). The individual exercises his right to be taken as he introduces himself. The ability to control impressions Goffman calls "the art of impression management". (Ibid) The participant is in Goffman's metaphor the actor who performs on the stage and demands that the audience takes his performance seriously. Mostly he tries to show that things are as they seem to be and he plays his performance "for the good of others". The spectators can identify themselves with his performance completely or not at all. Techniques of controlling the impression are aimed at strengthening the credibility of the impression that is given but also to prevent various faux pas in communication.

Many interpreters today see the importance of Goffman's work in that it demonstrates how to keep meaningful interpretable order in the elementary communication situations. It is an order that is kept thanks to a specified consensus and is reproduced by everyday routine. (Šubrt, 2001: 75) This order can be disturbed by different influences but people have established means for how to "save" the situation and course of communication. "Stability of interaction is not something given but we must constantly strive for it. If this regulation breaks down, the social experience is no longer shared." (Ibid)

A thematically and methodologically similar approach to dramaturgical sociology is **ethnomethodology.** Ethnomethodology represents the study of "ethnomethods" i.e. folk or lay methods which

people use to understand what others do and say. (Giddens, 1999: 91) For this understanding it is not enough to know the meanings of gestures and grammatical rules of used language but it is also important to note the context in which the interaction takes place. Harold Garfinkel (1917 -2011), the founder of ethnomethodology, dealt with this important finding in detail. He tried to reveal the shared meanings which enable us to engage in interpersonal communication and thus the functionality of the social order. He based his ideas on the assumption that we do not perceive many of these rules of everyday life because we take them for granted. But the absence of them can, according to Garfinkel, uncover these rules and together with them also the fragility of social order. In comparison to symbolic interactionists and Goffman we can see in Garfinkel's work a definitive shift in understanding of what constitutes the core of interpersonal interaction. From his point of view, people do not need to understand mutual thinking and motives to make the interaction between them possible. Interaction is enabled through their mutual relying on that the interaction will take place by established convention, on the basis of stabilized patterns of behaviour. (Light, Keller, Calhoun, 1989: 63) The ethnomethodological standpoint requires not only studying social reality from the point of view of participants but also "to pay systematic attention to the details of human behaviour... and also to competences on which such behaviour depends" (terms, thinking, knowledge). (Jalbert, 1999: 514).

Symbolic interactionism, Goffman's dramaturgical approach and ethnomethodology were criticized from the beginning. They were accused of the inability to reflect and explain broader social phenomena and processes. It is true, that it is really problematic to show whether and how it is possible to use findings of these orientations on the macro-level of social-scientific research – in what way the interpersonal interactions manifest themselves on the macro level of social structures. These structures are often studied apart from such things as subjective meanings. The reason is simply that nobody has shown how it would be possible. It seems that on the macrosocial level we can do well without these views so far, but the truth is that the "lower" we move on analytical levels of society, the more fundamental the existence of these factors (human autonomous thinking and consciousness) for social-scientific research is.

Part of this criticism referred to how human action is understood in these orientations. Some critics didn't like that it gives the impression of being based on hypocrisy and greed. This was ascribed mainly to Goffman but his work does not have to be understood only as a cynical view on human behaviour – man as a double-faced actor. His ideas can be interpreted as emphasizing the necessity of social consensus - participants of the interaction strive for a common definition of the situation, and as Goffman says, socialisation not only changes us but also unifies us - we learn to put the images of ourselves together coherently - and to create such a "Selves" in the interaction with which we would feel good but which would be at the same time acceptable for others.

#### Social Constructivism of Peter L. Berger and Thomas Luckman

Social constructivism is not only one of the orientations in social sciences. Social constructivism is an approach to the research of social reality which today has become almost generally accepted and is typical for many sociological orientations, especially those which prefer qualitative research. Since the 1960s social constructivism has become more and more important in social sciences. The work "*The Social Construction of Reality*" by **Peter L. Berger** (\*1929) and **Thomas Luckmann** (\*1927) from 1966 is considered ground-breaking work for the social-constructivist approach.

Social construction relates, broadly speaking, to every experienced social image about which people are convinced that is natural, given or self-evident, and they do not realize that it is a product of human activity - norms, conventions or traditions. Reality depends on the context in which we are moving or more precisely on that what is in society considered as knowledge, as something given. Therefore, we can find many various realities. Every individual or society has its own reality. And these realities are meeting and interacting. The "face-to-face" meetings are the most intensive perception of these realities. (Kabele, 1996: 323)

What is important is the relationship between human, as a creator, and the social world which is a human product (Berger - Luckmann, 1999: 64). These two subjects are in a mutual relationship which must be

understood dialectically. This relationship is made on the basis of three processes: externalization, objectification and internalization. We can express this triad with Berger and Luckmann's words: "Society is a human product (externalization). Society is an objective reality (objectivation). Man is a social product (internalization)." Whole process is then stimulated by institutionalization. The human being externalizes itself within an activity. It means that he manifests himself in society. With the repetition of these activities we arrive at their objectification. Reality created by us becomes an objectively given. We consider it completely self-evident, independent of us. Within internalization the human receives back new factors and impulses from the objectively given reality without realising that this objective reality was created by him in the first phase of the whole process (externalization). To pass the objective reality on to the next generations it must come to legitimation, which is the process of "explanations and justifications" (Berger - Luckmann, 1999: 64). The whole process happens again and again.

Berger and Luckmann distinguish four levels of legitimation. The most exacting is the fourth level where it comes to the creation of symbolic universes where "all human experience can now be conceived of as taking place within it" (Berger - Luckmann, 1999: 97). A symbolic universe can be compared to a universe that gives a life order and system. A symbolic universe gives the opportunity to become a part of some higher order (Berger - Luckmann, 1999: 97-98; Kabele, 1996: 324). Berger and Luckmann emphasize the importance of language "that provides the fundamental superimposition of logic on the objectivated social world" (Berger – Luckmann, 1999:67).

From the above-mentioned it follows that the social world is a world of meanings ascribed to it by people. Identity and culture are no objectively given. Identity is something which is constructed and can be in principle constructed in different ways. Identity is a variable and man as a subject is determined by his environment. Each of us is therefore a kind of construct. To be a human means to be the product and the creator of society at the same time. Findor (2007:17) points out that social reality is in ordinary thinking (of common sense) understood as objective, independent of people – a world created by the human is understood as a natural givenness. People do not usually realize their authorship of their social world, i.e. that they jointly form or construct the reality of social relationships, roles and images. Social reality gains its permanency, constancy and usability only when people forget that they have created it.

Key elements of social reality are, according to Berger and Luckmann, **institutions** present in every type of social action, in every social situation. Social institutions ensure elementary understandability of social relationships and action which, however, exceeds a given social situation. Social order based on the understandability and predictability of social relationships and action is thanks to institutions not only describable but also sustainable. People understand institutions as social facts independent of man and these institutions independent of human interference determine the borders of human action and content as well as the form of their relationships. If somebody oversteps them, his behaviour is incomprehensible and seen as a threat to social order. Disruption of institutional order is often considered a denial of natural order or deviation from reality. (Ibid)

Many qualitative researches try to work with this idea and to view the phenomena which are subject of their study as socially constructed. Sometimes it is only a basic ontological assumption of researcher, another time it is directly the goal of research - to uncover socially constructed origin of certain phenomenon.

#### Postmodernism and Poststructuralism

One of the orientations which have had a significant influence on the development of a qualitative paradigm is postmodernism, or poststructuralism. Some ideas and assumptions of these originally philosophical orientations were reflected also in social sciences.

Postmodernism as a term has multiple meanings and is therefore understood ambiguously even controversially - completely differently in various spheres of social life. It started to be used intensively since 1980s but we can distinguish two most frequent and most general meanings of postmodernism or postmodernity:

- 1. Postmodernity denotes those orientations that question basic elements of modernity in art, philosophy, psychology and social sciences.
- 2. Postmodernity also denotes the contemporary period in the development of society, the lifestyle and life feeling typical for it

that according to postmodernists prevail in contemporary society and also efforts for reflection of this period. With the term "postmodern society" we denote the spiritual and social climate in contemporary western societies.

Already the prefix "post" -(modernity) shows that it is a semantically negative term, i.e. it is defined through absences - through that what it is not. Postmodernity distances itself especially from modernity. To understand postmodernism better, we must ask about modernity.

The term modernity usually denotes the type of societies that followed after the traditional society in connection with the industrial revolution, progress in science and social changes with their consequences but also the type of thinking and idea that inspired these changes and that was expressed by the Enlightenment. During the 18th and at the beginning of the 19th century new social and political ideas gained importance in Europe. Founders of the Enlightenment (F. M. A. Voltaire, D. Diderot, Ch. Montesquieu, J. J. Rousseau, P. Holbach, Helvétius, B. Franklin, D. Hume and others) asked new questions and proposed new methods for their solutions. The representatives of the Enlightenment fundamentally criticized previous opinions about nature, society, humanity, the system of government and religious teaching. By their criticism the only criterion for everything that can be assessed was reason. The basic orientation of the philosophy of the Enlightenment was an emphasis on the natural equality of people, tolerance, searching for ways to secure a real opportunity for mundane happiness. Such ideas of the Enlightenment as objectivity of science, emphasis of connection between scientific and social progress, unity of mankind, liberation of man from restriction from state etc. helped to create the history of western Europe and northern America (the West). Leitmotiv of these ideas was the hope that people can create a world of freedom and justice for everyone by means of science and democratic institutions. (Sopóci, 2007)

Since the 1950s many of the ideas of the Enlightenment lying in the bases of modern society were questioned. Critics of modernity point out that the practices of western society such as colonialism, different forms of minority discrimination, exploitation of workers, poverty and social exclusion etc. resulted from ideas of the Enlightenment. So the Enlightenment didn't bring freedom and justice for everyone, especially not for Non-European, other races, women, the weak, the disadvantaged etc. (Ibid)

That caused other doubts and reflections: Is modern society so as it was considered to be? Is the society in which we live still modern society? Shortly after World War II some thinkers (e.g. historian A. Toynbee)<sup>21</sup> began to pay attention to vast changes of social and political life in western societies. According to them the industrial economy oriented to the production of goods was replaced by a post-industrial economy oriented primarily to the production of services. In culture the emphasis is no longer on hard work and accumulation of wealth but on emphasizing of consumption, delight and self-realization. (Lipovetsky, 1999) The interconnections between politics and science became unclear. Media and new information systems changed public space and our everyday life. These were the beginnings of discussions about the present and future of modern societies that influenced sociology. From this point of view, postmodernism а theoretical (philosophical) reflection is of spontaneously arisen events in society, culture and various areas of them. Postmodernism is the manifestation of disputes about the nature of modernism (modernity). After the age of optimism resulting from modern technological progress, the latest achievements of science and technology, the belief of the Enlightenment in the omnipotence of human reason and unlimited abilities of humanity, it comes to disillusionment. The society and its perception and reflection have changed.

In the diversity of postmodernist orientations we can also find the common features. They are: resignation on the idea of progress; dissolving of the subject as a centre of knowledge, action and experience; scepticism towards terms describing reality (in philosophy and social sciences), refusal of general terms and general principles and overall refusal of logocentrism of modern thinking (unilateral rationalistic world projection); proving of invalidity of theories and "stories" that claimed universality; refusal of any uniformity of thinking and totalitarianisms of any kind; looking for alternative ways of thinking; acceptance of radical plurality of lifestyles.

<sup>&</sup>lt;sup>21</sup> A. Toynbee writes in 1954 about contemporary stage in society development that it is a period of transition from national-state thinking to global interaction.

The relationship of postmodernity to modernity is in postmodernism reflected in a different way, in spite of its criticality. Most often, postmodernism understands and presents itself as an attempt to overwhelm starting theses and determining conceptions of modernity. postmodernists understand postmodernism Some radical as а fundamentally new paradigm, completely different from modernity. Others understand postmodernity either as a new dynamic phenomenon inside of modernity or as its contemporary variant (alternately called postmodern modernity). W. Welsch (1993: 59) rejects the opinion that postmodernity is anti-modernity. It is only a radicalized modernity, not the "breakup with modernity but rather radical asking about modernity. It is not separated from modernity by a break but it has specific interconnections with it".

The term **poststructuralism** is closely related to the term postmodernism. Most often, these terms are used as synonyms but sometimes they are terminologically distinguished. The differences are related primarily to different areas of contemporary societies and culture. In this case, the postmodernist analysis is related to reflecting contemporary society, social theories, postructuralistic analysis is situated in the area of literary theory, philosophy and history. Postmodernists (among sociologists especially J. Baudrillard, Z. Baumann) point to the exhaustion of modernistic theories and their inability to explain the specific character and development of contemporary societies. Some (e.g. Z. Baumann) believe in the possibility of creating postmodern sociology as an intermediator between various societies and cultures in postmodern plurality. Poststructuralists (e.g. J. Derrida, M. Foucault, J. F. Lyotard) pay attention to the questions of thinking and knowledge, e.g. to those such as the crisis of representation and with it a connected obscurity of meanings, absence of fixed bases of knowledge, role of language and discourse or study of how an individual becomes an object of action and controlling in modern or postmodern societies.

Poststructuralism is usually defined in dictionaries as an orientation in contemporary philosophy that split off from structuralism in the late sixties (in France). Poststructuralism refused a systematic concept of structure and principle of binary opposition (from formal logic: "yes - no") of older structuralism, from structuralism it took over only one idea (and radicalized it) - that a language system has priority over "self". The effort to unmask the logocentrism of western philosophy in which reason as means of power and control suppressed the minority, partial in the name of general and universal, is related to poststructuralist philosophy. On the other hand, poststructuralists methodologically proceed from the principle of plurality and from the fact that we must expect endless amount of mutually interchangeable meanings in each language expression. To catch the particular (partial) and the unidentical (unequal) as irreducible givenness, we must use a hermeneutic procedure of deconstruction. The deconstruction means decomposition of the polysemous language phenomenon into all its possible meanings and links, in all possible contexts and texts (not only scientific or philosophical).

Postmodernism is an important object of study for sociology. Sociologists are, of course, most interested in the question of the character of contemporary - postmodern society. There are such topics as the postmodern family, postmodern mentality of the young generation, postmodern art, etc. But sociology itself is not influenced much by postmodernism, the discussions are mostly about methodological plurality. (Jandourek, 2001:188)

Alieva (2006) evaluates the situation in contemporary sociology in which modern and postmodern orientations coexist: Modern and postmodern sociology complement each other, there is a specific division of labour between them: postmodern sociology can fulfil the role of "searching, trying and experimenting theory from which anti-scientifically oriented sociologists could draw inspirations" and also thinkers of other branches. But as a "shadow", modern sociology always accompanies it and occurs "always when it is necessary to do a solid report about the state of society or its areas" (ibid).

#### Methods and Techniques of Qualitative Research

With the goal to build up knowledge in the spirit of premises of those orientations that were crucial for the development of qualitative methodology, many various methods and techniques were formed. In this part we introduce only some of them – those that are used the most frequently.

## Case Study

Pascal Vennesson (2008: 223) states that "a significant part of what we know about the social and political world comes from case studies. Case studies famously contributed, for instance, to uncovering the tendency towards oligarchy in political parties, the inner working of the exercise of power in democracies, the dynamics of international crises, the logics of authority and control in organizations, the interplay between values and institutions in the Indian caste system, the sources of success and failure of deterrence, and the causes of social revolutions". Besides using case studies in these classic works, case study is nowadays still popular method in social-scientific research. However, we can find periods in the history of particular social sciences when case study was very popular or contrarily - damned. The main objection to case studies has always been the inability to offer theoretical generalizations from one case. That means an objection from the side of positivist conception of social sciences. As Kořan (2008:29) states a great part of history of case studies can be understood as philosophical and methodological competition with the objection that "we cannot generalize from one case".

This fight took place in two basic forms. The first one tries to prove that it is possible to gain reliable information about a broader category of phenomena through the detailed analysis of one case. In other words, that it is possible to create general theories without studying larger numbers of cases. This approach is also called **instrumental or positivist** and it sees the case studies as a useful tool for further work with theories. The second one holds the view that detailed, expert and often practical knowledge offered by case studies have their scientific value themselves without theoretical generalization. This approach is called the **interpretative approach**.

Case studies appear in various disciplines, in law, pedagogy, history, medicine, psychology, sociology or anthropology. The most influential school responsible for spreading of case studies in the world is American Chicago School, influential mainly since the 1920s through the first half of the 1930s. But the gradual penetration of positivism into the American social-scientific environment in the second half of 1930s and in 1940s brought the decline of the Chicago School in favour of statistically and theoretically oriented quantitative approaches. The main argument against case studies was their inability to provide clear tools for verification of research results. (Ibid)

Thanks to their permanent effort to reach the scientism of the quantitative approaches, supporters of case studies since the 1950s tried to build up their tradition on as solid and precise foundations as possible. Paradoxically, the dispute with quantitative approaches caused that case study gradually became more elaborated with the ambition to equal strict criteria of scientism. A very good example of this development is so called **grounded theory**.<sup>22</sup> Since the beginning of the 1980s a renewed interest in case studies occurs. The pragmatic spirit of this period brought the opinion that quantitative and qualitative approaches do not have to be in opposition and the discussion about which one is more "scientific" is unproductive and useless. (Ibid)

For a successful and meaningful creation of a case study it is important to know what a case is. Kořán (2008: 32) considers case to be a "sufficiently bounded or limited aspect of historical episode or this episode as such." Vennesson (2008: 226) claims that "a case is a phenomenon, or an event, chosen, conceptualized and analysed empirically as a manifestation of a broader class of phenomena or events". A case study is then according to him "a research strategy based on the indepth empirical investigation of one, or a small number, of phenomena in

<sup>&</sup>lt;sup>22</sup>Grounded theory is actually a very strictly inductive approach toward building a theory. Unlike positivist research, a study using grounded theory is likely to begin with a question, or even just with the collection of qualitative data. As researchers review the data collected, repeated ideas, concepts or elements become apparent, and are tagged with *codes*, which have been extracted from the data. As more data are collected, and as data are re-reviewed codes can be grouped into concepts, and then into categories. These categories may become the basis for new theory.

order to explore the configuration of each case, and to elucidate features of a larger class of (similar) phenomena, by developing and evaluating theoretical explanations" (ibid).

The "boundedness" in terms of time or topic should be so important that it encourages with its character to define it as a case. Case can be defined also as a certain object or closed system with clear boundaries, inner logic of functioning and its specific nature. A case can be the process of negotiation about a definite international declaration but the general phenomenon of negotiation of international contracts is not a case anymore. It lacks clear boundedness and definition. A case can be a specific Ministry of Foreign Affairs as a participant in the economic dimension of foreign politics but the economic dimension of foreign politics as such is no longer a case. A case study is a detailed analysis of case chosen as an object of interest by a researcher. Its aim is to provide deep understanding of this case, so it must take account of its context (social, political, historical) and must provide a complex view of the studied case. (Kořán, 2008: 33).

Case studies come in different shapes and forms, and they can serve a variety of purposes, often simultaneously. First, the descriptive case study (configurative-ideographic) is a systematic description of the phenomena with no explicit theoretical intention. It is common to label this kind of research as simply suggestive and to dismiss its social scientific contribution. Still, while the work of many historians and anthropologists might lack an explicit theoretical framework that does not mean that a theory is altogether absent. Furthermore, in any type of case study there is an unavoidable descriptive dimension. Case studies sometimes explore subjects about which little is previously known or phenomena in need of an interpretation that sheds new light on known data, and their descriptive aspect is invaluable. (Vennesson, 2008: 227).

Second, the interpretive case study uses theoretical frameworks to provide an explanation of particular cases, which can lead as well to an evaluation and refinement of theories. Third, the hypothesis-generating and refining case study (heuristic) seeks to generate new hypotheses inductively and/or to refine existing hypotheses. The researcher can clarify the meaning of certain variables and the validity of empirical indicators, suggest alternative causal mechanisms and identify overlooked interaction effects. A deviant case is especially useful to generate new hypotheses and/or to adjust theoretical propositions. Fourth, theoryevaluating case studies are used to assess whether existing theories account for the processes and outcomes of selected cases. (Ibid)

Individual types of case studies reflect ontological disputes about the character of the world, and epistemological disputes about the character of our knowledge about the surrounding world. If we accept the standpoint that the world is made of variable material and completely unique events, then we incline to those case studies in which it is important to understand only one case without the ambition to say anything theoretically relevant or generalizing. But if our ambition is to uncover (at least partially) causal mechanisms acting in the surrounding world, then the case study has a rather instrumental function for us and the requirement to "communicate" with theories it hen relevant. The value of such instrumental case study does not lie in studied phenomenon as such but its benefit is broader and more general theoretical knowledge it provides. Here the case is studied deeply and it is set into context but with the aim to confront various features of that case with certain theory.

An infinite number of factors react in the social world. The case study has as its aim to record this multiplicity as faithfully as possible. As we have mentioned, it can also be useful for verification of existing theories or for the generation of new theories. But a case study does not have comparison at its disposal and so cannot check the influence of factors so as it is supposed by the experimental method and therefore we can expect from case studies more likely tentative conclusions about the influence of a given variable on final result. However, case studies have their strong point by the identification of the extent of circumstances under which the assumed causal mechanism works. (Kořán, 2008: 40).

There are several specific procedures for testing the relevance of a certain theory. One of them is testing by the crucial, the least probable and the most probable case. A crucial case is a concept which was delineated by Eckstein (1975, In: Kořán, 2008: 41) and which corresponds most with the assumption of the existing theory. If the prediction or result fails, the whole starting theory is probably incorrect. A test designed in this way has two basic problems. Firstly, we cannot find out if the incorrect result is a consequence of the theory itself or of contextual conditions that make this theory inapplicable to the given problem. Eckstein (ibid) proposes therefore to take the cases which are most probable and least probable,

and study them in relationship to the tested theory, but even to the alternative theories. We get the best possible proof for strengthening the theory if we find a case that is the least probable for the tested hypothesis and probable for other theories. If such case shows aspects congruent with assumptions of the least probable theory while other theories fail in prediction of result, we can say that tested theory went through the strictest possible empirical test. We find out the best possible way to question the theory when we find the most probable case for tested theory and alternative theories turned out to be applicable to this case as well.

Another function of case studies is the creation of theories. It comes to this situation when the studied case defies the existing theoretical explanations we want to apply to it. It forces us to create new theory which would explain this case and in which broader validity is a subject of further research. Case study often points to new variables or hypotheses. For this purpose, we can use a specific method called process tracing. Process tracing is a procedure for identifying steps in a causal process leading to the outcome of a given dependent variable of a particular case in a particular historical context (Drulák, 2008: 18). Using process tracing, the researcher assesses a theory by identifying the causal chain(s) that link the independent and dependent variables. The goal is to uncover the relations between possible causes and observed outcomes. This procedure can be used in theory testing as well as in theory development.

In a positivist perspective, the main goal of process tracing is to establish and evaluate the link (or the absence of a link) between different factors. Through the use of histories, archival documents, interview transcripts and other sources, the investigator examines whether the causal process of the theory that he is using can be observed in the sequence and values of the intervening variables. In an interpretivist perspective, process tracing allows the researcher to look for the ways in which this link manifests itself and the context in which it happens. The focus is not only on what happened, but also on how it happened. It becomes possible to use process tracing to examine the reasons that actors give for their actions and behaviour and to investigate the relations between beliefs and behaviour (Jervis 2006, In: Vennesson 2008: 233).

We can state that case studies are often beneficial for the knowledge not only of concreted studied case which is often interesting as such but also for work with theories. Of course, the right choice of a case and further work with it is crucial. Therefore we can also say that by the analysis of case studies we must be cautious, avoid inappropriate generalizations and realize the possibilities and limits of chosen approach. Finally, it is important to recall that during the case study, we can also use the quantitative data.

#### Qualitative Interview

Denzin and Lincoln (2012: 47) state that we live in an interview society, whose members seem to believe that interviews generate useful information about lived experience and its meanings. Asking questions and getting answers is much harder than it may seem at first. The spoken or written word has always a residue of ambiguity, no matter how carefully we word the questions and how carefully we report or code the answers. But interviewing is a popular way to get any systematic information and this not only for social scientists but also for others - e.g. journalists. The interview is a traditional technique of quantitative research, as we have shown in previous part. Interviewing, as a part of a qualitative research, has several distinct features and characteristics.

The first distinct feature is its bigger laxity and lower structuring. While the interview in quantitative version is based on the same series of questions that all respondents answer and these questions have usually limited and given number of answers, qualitative interview is structured more loosely, sometimes only the overall lap of questions is given and it is considered to be an advantage when the researcher reacts to the situation during the interview and is able to adapt the questions.

But it does not mean that the preparation before entering into the field is not important. Of course, it is good to have a **preliminary list of questions.** Although according to Kaufmann (2010), in case of qualitative interviews this list should be a very flexible manual. It helps the researcher not to forget to ask something important. For the researcher it is important to identify himself with his questions which enables him to ask the questions during the interview freely and in an appropriate moment. At the same time the researcher must be creative in the field which means to reformulate the questions or formulate new ones

promptly on the basis of new findings. It depends on the researcher and on the existing problem that he/she wants to solve, whether to prepare definitive questions or only rough spheres of questions. Kaufmann (ibid) points out that it is good to order the questions according to particular topics to avoid digressing from one topic to another.

Although it is not a representative research, creating adequate research sample is very important in qualitative interview. Although in the case of qualitative interviews, the sample can be completed also during the research, it is good to have at least an approximate idea about what criteria of choice the researcher will take into consideration. strict representativeness according to the Although it crucial characteristics of population is not insisted on, in some cases it is logical that we will strive for certain representation of specific categories in the sample. We can say that it is good to avoid the evident imbalance in the sample. For example, not to omit men and also women in the sample, or different age groups if the studied problem requires this. One method of sampling in qualitative research is to proceed so that we have good representatives of the given phenomenon in the sample. In the case of less accessible populations (e.g. immigrants), we use the method of snowball sampling.

If the sample and the list of questions is prepared, we must "only" meet the respondents and realize the interviews. The beginning of the interview, which means addressing the respondent and ice breaking, can be sometimes difficult but it is possible to go through this phase relatively fast to the deeper phase of the interview which goes relatively smoothly in the case of qualitative research. But it is also in the hands of the researcher. This type of interview requires establishing the situation in which it comes to what Kaufmann (2010) calls "**disruption of hierarchy**". It is necessary to establish as natural situation as possible - so as to disrupt the hierarchy of power (especially on the side of researcher) to reach as loose and honest answers of the respondents as possible. Then it is of course necessary to listen and think carefully about given question which should either follow from what the respondent said or should be connected to it.

Most ideal is (because of the data proceeding) if we could record the interview but still we must listen to the answers during the interview. In the case of recording it is necessary to get the agreement to record the respondent. Of course, there are topics by which recording could disturb the course of the interview but usually it is so, that the respondent can be a bit nervous by the recording but mostly he stopped paying attention to it during the interview. It also depends on the researcher's behaviour. Empathy and interest are important.

The particularity of qualitative research and also of a qualitative interview is that the phase of data collecting and data analysis take place simultaneously. The data should be analysed continuously. And on the basis of this preliminary analysis there might emerge some implications for further data collecting. Final conclusions (e.g. in form of final report, publication or thesis) are then fine-tuned. The crucial goal of the analysis is to give a compact image about our findings - but this does not mean that it should be any detailed description of all what the respondents said. We compare the information received about the respondents with our theoretical knowledge and on the basis of this, we try to point out what we consider inspiring and interesting - what indicates some theoretical conclusions (so we either make some new theoretical conclusions or verify existing theoretical knowledge).

In qualitative research we shouldn't look at the acquired information about respondents only from the point of view of what objective information they offer but we should also analyse "between the lines" - if the respondent said something, in which context did he say it, what does it mean and how the respondent understands the problem. Why he had such long pauses, or to analyse some contradictions in respondent's answers. But our view of what the respondent said is a kind of understanding interview, so we do not judge his answers but we try to understand his point of view.

In building a theory we proceed on several levels, at first we just compare some basic summaries with a concept, we create a middle range theory and in the end our findings can be situated into a broader context. For instance we find out that by the approach of young people towards sport, the family (e.g. motivation of parents etc.) plays a role. So we can connect this with broader knowledge - e.g. that the role of sport is probably underestimated in society because there is not sufficient access and motivation for young people to participate in sports from the state, leisure centres and schools. We can go further and say that it is the result of the current globalised neoliberal capitalism.

## Ethnography

Ethnography is at the heart of qualitative methods in the social sciences, in relation to the descriptive and interpretive approach. It is a method based on the ethnographer's long-term stay in a studied group. The base is called a thick description<sup>23</sup> of this group, its way of life, culture, organisation of relationships - especially on the basis of observation and through understanding. Ethnographers immerse themselves in the context of the phenomena they are studying.

The original aim of ethnography was to describe ethnic groups distinct from the culture of industrial society by their culture and by their way of life. The first ethnographers often accompanied missionaries or travellers on their journeys of discovery and they observed and described new different cultures. Over time, the description changed to the effort of understanding and interpretation of the behaviour of these people. The development of ethnographic methods and strategies has been closely associated with anthropology. Anthropologists continue to take the lead in insisting on linking their inquiries to the understanding of culturally based behaviours and values, thereby distinguishing ethnography from the broader category of qualitative research methods. (Chambers, 2003)

With the increasing influence of industrialization, sociologists from the American Chicago School came-up with the idea to use the research tools of already established anthropology by the study of social inequalities and life in towns. The crucial question was for them the origin of social problems and the probabilities to solve them. An advance to solve questions and problems of modern society was also seen in anthropology, which moved through this to the sphere typical for sociological research. Some anthropologists began to study, for instance, habits of different communities and ethnic minorities. They used their specific approach to "different" culture study also by the study of administrative and bureaucratic institutions where they had to defend their approach in the sphere of interest of traditional political science. This defence was based on specific methodology or optics with which the anthropologists are able

<sup>&</sup>lt;sup>23</sup> A thick description of a human behaviour is one that explains not just the behaviour, but its context as well, such that the behaviour becomes meaningful to an outsider. The term was used by the anthropologist Clifford Geertz in his *The Interpretation of Cultures* (1973).

to look at these institutions - to decode specific **political culture**. This approach was partially used also as practical tool e.g. by the preparation of cultural profiles of different social groups that became bases for politics creators. (Tulmets, Střítecký, 2008: 150 - 151).

The ethnographic approach is naturalistic, in that it attempts to work with society as it is, without trying to influence or control it. The goal is to understand behaviour in its habitual context, as opposed to an abstract or laboratory setting, and to interpret how people give meaning to their experiences. Ethnographic research involves an exploration of a society's cosmogony, of the way in which people make sense of the world they live in and how, acting on the basis of their beliefs, they relate to each other and to people different from themselves. Through descriptive generalizations and the development of explanatory interpretations about how societies work, in particular contexts and time spans, the researcher seeks to account for the commonalities and variations among societies and their trajectories over time. (Ibid)

Chambers emphasizes several of the most important approaches to ethnographical research. The first one proceeds from the essay by Bronislaw Malinowski "*Practical anthropology*" (1929) which offers the view on how ethnographic understanding might assist British colonial administrators in Africa. Malinowski's argument for the usefulness of ethnographic data in colonial administration rests principally on the anthropologist's presumed grasp of what is sometimes called "the native point of view" (Chambers, 2003: 392). The researcher who participates intensively and for a long time in the studied field (in studied culture) should be, according to Malinowski, able to adopt this point of view. At the same time, the ethnographer should be able to observe behaviours and to explain their significance in relation to their functions in a larger institutional and cultural context. (Ibid)

Another rationale for applied ethnography was offered by Clyde Kluckhohn in his article "Covert Culture and Administrative Problems" (1947) written on the basis of his experience with the U.S. Indian service. He argued that further progress might be made if ethnographers were to pay more attention to the covert dimensions of culture or to those cultural configurations that lie somewhere beyond a people's conscious awareness of their own culture (Chambers, 2003: 393). So this approach assumed that a researcher can be someone who finds out and says about the others
even more than they know about themselves. The third specific approach is the one by Sol Tax (1958) which he called "action anthropology". By the study of Fox Indians in Iowa he focused especially on problems that the Indians face by the contact with Caucasians. According to Tax, the anthropologist should be not only the one who studies but also helps and mediates the contact of two cultures.

In all these cases the basic and common method remains what we call **fieldwork or participant observation**. This basic method can be realized by using defined techniques such as observation, in-depth interviews or study of documents. The result of the observation is the material which the researcher acquires on the basis of his own participation on the events in given community or group. It assumes, of course, closer or even personal relationships with the observed that the researcher makes during research. This method requires much more than the other methods, to win the trust of the observed. During the research the researcher lands in conversation with the objects of his research either purposefully or not and therefore the data acquired this way might represent valuable part of his stay in the field. The researcher often comes into contact with different objects of material or non-material character, notes, diaries, visual material which can tell us much about the studied culture.

From the point of view of practical realization of ethnographical method, the **field notes** depicting important moment of the research are crucial for the researcher. They should be as rich and detailed as possible because things that seem to be of little relevance can become important later by the interpretation. Of course, the most important part of research work is to interpret the findings and to get them in some form over to the scientific community or broader audience. It is like by qualitative methods in general. The researcher must show incontrovertible creativity by the data processing but he must remain faithful to the facts. His interpretation should be rooted in observation but at the same time, it must exceed the horizon of simple facts. Like by qualitative methods also here it is not clear in advance if this will succeed, however, since it is a really intensive contact with studied population it is probable that the researcher will answer this question in the field.

At first sight it might seem that the method of ethnography is not very usable in political sciences but Elsa Tulmets and Vít Střítecký (2008:

164-173) give some examples of such application. For instance, the study by Marc Abéles "*Ethnologist in Parliament*" from 2000, or the research by Jean Joan and Andy Smith of the European Commission.

# Discourse Analysis

One of the methodological approaches which have broad prospects for application in social sciences and specifically by the study of politics is discourse analysis. From different philosophical sources that saturate assumptions and applications of discourse analysis, the paradigm of poststructuralism mentioned in the previous chapter is crucial.

In general, different sources and different variants of discourse analysis agree on one crucial assumption that "the language matters". Beneš (2008: 92) points out that: "Traditional methodologies consider spoken and written discourse only an ordinary (often unreliable) mediator of information about the world "behind" the text." But discourse theorists consider the expression by means of language a basic human meaning-making activity. Then we can understand the study of discourse as a "study of human meaning-making practices" (Ibid). The most basic feature of discourse from the point of view of discourse theory is the fact that discursive practice is always social practice. A classical work that points to this feature of language is the work by John Austin "How to do things with words" (Austin, 2004). In this work, Austin analyses the basic feature of human language and expression - its performativity. Performativity means that expressing by means of language has its secure, tangible effects in relation to social reality. Definitely employed terms, concepts and statements, different "texts" produced in the world have the performative effect.

For instance the current shape of geographical borders of states is to a large extent the result of language, discursive activity. Within this activity also the meanings of such concepts as nation, state, border, national interest, international community were "negotiated". The discovery and "negotiating" of meanings of these terms and concepts had palpable performative effects vis-a-vis social reality, it built up the organization of the world in which we live and such understanding of state that is dominant at the present. But of course, current understanding of these terms is by no means something definitely finished because the process of negotiation of meanings is never finished. There are always alternative discourses that formulates in other way what is relevant, they bring new demands, other criteria of good and truth, other seeing of the world and its priorities. On the one hand, the dominant, hegemonic discourse of nationalism formulate the idea of a state as a national state with dominance of one nation, ethnic group, and on the other hand are the alternative discourses such as multiculturalism discourse that formulates the idea of state as a multinational whole built up by individual cultures and groups which should have equal rights and duties.

These discourses compete for **power**. Power is the key word of discourse analysis because language is the main arena of power struggles. To take control over particular discursive space is the key to reach hegemony. If I assert definitions of key words and concepts which become widely accepted and little questioned, they will push the alternative concepts and explanations out of the dominant space or, as Foucault (2006) says "out of the truth", and if these definitions are institutionalized, i.e. anchored in social practice, then we can say that the discourse reached hegemonic position.

We could, for example, ask which one of above mentioned discourses - discourse of nationalism or multiculturalism is the hegemonic discourse in the discursive field of understanding of the state. Such discourse analysis wouldn't be based only on language analysis - which means the analysis of e.g. speeches of politicians, parliamentary debates or media performances with the aim to uncover what terms and concepts are here frequent and which assumptions and rationalities these statements formulate, but we would have to look at the institutional practice - to see which of these discourses is institutionalized in the social reality in which we live. For example, what concrete rights and duties do the members of minorities or foreigners have in our country. Is the dominant ethnic group, Slovak in this case, seen as the dominant and constitutive element of the state etc. Although to analyse discourse means to proceed from the assumption that "the language matters" it means also to analyse how the formulations and definitions in language practice are transmitted on a non-language practice.

Within different traditions from which current conceptions of discourse analysis proceed (especially Foucauldian discourse analysis, inspirations from linguistics, inspirations of microsociological and qualitative orientations in sociology as symbolic interactionism, ethnomethodology, etc.) there are several approaches to the study of discourse. N. Phillips and C. Hardy (2002: 19) talk about approaches oriented more on text or context. It means approaches that attach bigger importance to context (and are from this point of view closer to Foucauldian analysis or sociolinguistic analysis) and orientations that are concentrated more on text - they are closer to linguistics. The next criterion of classification of individual approaches by Phillips and Hardy is the extent of inclusion of power dynamics in the process of discourse analysis. From this point of view we can distinguish more critically oriented studies versus studies oriented more on process of social construction that constitute the social reality represented in texts and discourses.

In our case the most relevant are approaches aimed more at social aspects and aspects of the power of the discourse. Here belongs an orientation called Critical Discourse Analysis (hereinafter referred to as CDA) and the Essex school of discourse. The postulate shared by both these schools is social constructivism - the idea that our communication plays an active role by creation, stabilization and transformation of social reality. But in the question whether the discourse shapes the whole sphere of social reality or whether it is only one of its aspects, these two schools diverge. It is obvious that clarification of this point is significant from the point of view of the approach to analysis. If the social dimension does not overlap with the dimension of discourse then there also exists the sphere that works on the basis of other principles and must be therefore analysed in other way. But the discourse should be analytically graspable without adverting to economic, political and other causalities built-in the society.

#### **Discourse Theory of Laclau and Mouffe: Essex School**

Every practice that establishes a relationship between its elements in such a way that their identity is due to its being modified is called by E. Laclau and Ch. Mouffe articulation. Discourse is then for them "structured totality resulting from articulatory practice". Discourse is an attempt to fix a web of meanings within a particular nodal point. Formation of discourse includes the excluding of further meanings and reduction of some likelihood for the benefit of others. Because of this, formation of discourse is always the execution of power. Every discourse is constituted as an attempt to take control over the discursive field, attempt to stop the flow of differences, to construct centre. E. Laclau and Ch. Mouffe call these privileged points, centres, nodal points (Laclau, Mouffe, 2001: 112).

Discourses try to fix a web of meaning through constitution of nodal points which bind them. This way the discourses give meaning to various markers. For instance in the communist ideology and discourse, the marker "communism" is a nodal point that connect already existing markers as "democracy", "state" and "freedom" and gives them a new meaning: democracy as excluding the class oppression, there are economic connotations given to the term freedom which changes the content of the term (Žižek, In Torfing, 1999: 99).

A differential position articulated by discourse is called "moment" Difference not articulated by discourse is called by Laclau and Mouffe "element" (Laclau, Mouffe, 2001: 105). But no discursive formation is a "tied" totality and the transformation of elements into moments is never finished. Elements have the character of floating signifiers that cannot be completely articulated into a chain. This unstable character penetrates through every discursive identity (Laclau, Mouffe, 2001: 105-113). Torfing sees it similarly and he defines discursivity as a "field of irreducible surplus of meaning" (Torfing, 1999: 92). According to him the discourse is a "differential ensemble of signifying sequences in which meaning is constantly negotiated" (Torfing, 1999: 85). But it never comes to complete "negotiation" and closure of meanings and the fixation of meanings within a specific "nodal point" is always temporary, there are always meanings that escape this fixation.

If complete totalization of meanings were possible, there would be no allowance for politics. All identities would be fixed as necessary moments of one discourse and conflict would take place between various intradiscursive variations. But a fixed centre cannot exist and therefore there will be always something which will escape the seemingly endless processes of signification within the discourse. That what will not be fixed as differential identity within a designated discourse is therefore outside the discourse but it is discursively constructed in the arena which escapes the stabilization (Torfing, 1999: 91-92). In the understanding of Torfing, everything is discourse and the distinctive feature, by the definition of discursivity is the level of meaning fixation. But what is important is that he offers us asignificant definition of politics which follows from his model. We can understand it as a sphere where the producers of discourse struggle for the meaningful fixation of individual elements of discourse in an effort to reach hegemony.

This point of view is topical by the research of historical turning points as was for example the change of social-ideological paradigm in post-communist countries in the late 1980s. The stable hegemonic discourse was confronted with new events which it was not able to explain, represent or to domesticate in other way - <u>it was dislocated</u>. This process opened up room for political struggle for definition and solution of the problem that occurred and led to the articulation of a new hegemonic discourse stabilized through construction of a new set of political borders. During the hegemonic struggle the producers of discourses tried to fix the floating signifiers and to "bind" them to a new nodal point (Torfing, 2005: 16; Žižek, 1999: 92-94).

#### **Critical Discourse Analysis**

The term Critical Discourse Analysis (CDA) comprises many various approaches, various authors that have something in common. First of all, CDA approaches are interested in "naturally occurring language" as used by the real participants. They are also interested in functions (social, cultural, situational and cognitive) of the contexts of language usage. Of course, the basis is the study of language and analysis of various textual grammatical and language forms: coherences, topics, macrostructures, speech acts, argumentations, rhetoric, mental models and other aspects of speech (Wodak, Meier, 2009: 2).

The next common key feature is connected to the term "critical". The critical accent lies in the deconstructivist strategy which has the aim to uncover hidden meanings and power interests that constitute certain discursive strategy. CDA theorists take an explicitly "evaluating" stand - it means that they leave the ideal of "neutral" science and they see their scientific activity as engaged. R. Wodak speaks about the fact that some

CDA authors play advocatory role for groups socially discriminated against (Wodak, In: Wodak Meyer, 2009: 19) and the line drawn between scientific research and political argumentation sometimes gets blurred. Norman Fairclough states that CDA is a form of critical social science oriented toward a better understanding of the character and sources of "social wrongs", and of the obstacles by their identification and the ways to overcome these obstacles (Fairclough, In: Wodak, Meyer, 2009: 167). He defines social inequities as "aspects of social systems, forms or orders, which are detrimental to human well-being, and which could in principle be ameliorated if not eliminated, though perhaps only through major changes in these systems, forms or orders." (Ibid) As examples he gives poverty, inequality, lack of freedom or racism. Fairclough realizes that to define what is "social inequity" is controversial; CDA brings therefore contribution to this controversial discussion. Identification of these "social wrongs" and the obstacles by the removal of them is therefore by Fairclough a part of discourse analysis.

N. Faircloughn who within CDA worked up the discourse theory, speaks about it in connection with the methodological framework for his research, as a complex of three components: social practice, discursive practice (text production, distribution and consumption) and text (Fairclough, 1995: 74). The connection between text and social practice is seen in his three-dimensional model as being mediated by discourse practice, processes of text production and interpretation are shaped by the nature of the social practice, the production process shapes (and leaves 'traces' in) the text, and the interpretative process operates upon 'cues' in the text (Fairclough, 1995: 133). Fairclough points out that discursive practices, which he understands as practices mediated by language, are only one of the forms of social practice (Fairclough, 1992: 95). So he acknowledges the existence of processes without discursive character (e.g. institutions, technologies etc.) which influence the shaping of society.

The basic difference in approach between CDA and the Essex School of Discourse is the study of the connection between discourse and nondiscursive social practice which CDA understands as two spheres, whereas for Essex everything is discourse. The second important difference is that within CDA they work more with various linguistic elements of language such as argumentation schemes and syntactical means which express unity, sameness, difference, singularity, continuity, change, autonomy, heteronomy, etc.

Discourse analysis is always based on analysis of a set of "texts" where "text" is understood very broadly as any written or spoken expression including visual elements. Institutional practice is always taken into consideration, because it is understood as part of the discourse or because the aim of the analyst is to study their mutual relationship. Therefore it is hard to give the student instructions for dealing with this analysis, but it is like qualitative research in general. We choose either inductive approach and try to identify individual discourses from the data without any initial assumptions or hypotheses. The discourses begin to "emerge" from the data by the study of these texts. Often we know about the studied problem in advance and have some expectations what the data might bring. Then we formulate and frame research questions and try to answer them and to develop them by means of the analysis of relevant texts.

We can convince ourselves about the relevance of discourse analysis as a strategy for political scientists or experts on international relations through many studies which have used this method. Besides the above mentioned analyses of nationalism there is, for example the study by David Campbell (1998) "Writing Security: United States Foreign Policy and the Politics of Identity" or other analyses of the war against terrorism, e.g. analysis by Phil Graham, Thomas Keenan and Anne-Maree Dowd (2004) "A Call to Arms at the End of History: A Discourse-Historical Analysis of George W. Bush's Declaration of War on Terror".

# **Stages of Research and Research Project**

Helena Kubátová (2006: 183) points out that to write a quantitative research project is not easy at all, but it is easier than to write a qualitative research project. It is because qualitative research is not standardized (and its steps are not definitely given), and all the research steps within qualitative research take place simultaneously. Every subsequent decision changes the initial research frame and therefore it is very difficult to design qualitative research. Besides this, it is not possible to forecast the results of qualitative research - they are uncertain. In other words, we know that we will find out something - we will either confirm or reject the hypotheses. But in quantitative research the researcher enter the field and does not know if he will find something interesting about which he could write in his final work.

# Preparatory Phase of the Research: Qualitative Research Project

In spite of this, Kubátová (2006) tried to propose the specific structure of the qualitative research project but she points out that the order of individual steps does not correspond with their real position in the research, because in the qualitative research it is necessary to work on all three steps simultaneously:

- 1. Formulation of the research topic
- 2. Formulation of research paradigm. Description of the theoreticalempirical starting point and choice of the topic: formulation of the frame idea about studied topic
- 3. Formulation of research question and research goals. What riddle will I try to solve?
- 4. Proposal of the case that represents the chosen research problem and the explanation of its choice.
- 5. Proposal of the data collection technique and explanation of the choice

On the basis of what we already know about qualitative research, during the research the basic research question might and probably will change and so will the frame idea about the studied problem. Despite this, a defined frame project is very important for qualitative research.

#### **Formulation of Research Topic**

We have already mentioned that the choice of methodology (but then also of concrete method and technique) is the function of research intention - so the most important is to have an idea about what we want to study, or to find out - we will take closer look at the choice of topic within qualitative research methodology, but also at the choice of a predetermined research strategy with regard to the topic, and how we should proceed when we have the idea that we want to study something qualitatively. Just as in quantitative research, the source of research topic is here also either theory or practice. D. Silverman (2005) gives many examples of the choice of research topic for qualitative research motivated by the practice by way of examples of his postgraduates.

First of all, qualitative research requires doing "much from little", it requires a great deal of inventiveness. Since in qualitative research there aren't hypotheses set in advance, we often start the research and do not know if we will find out something relevant and worth publishing. Of course, it is necessary to have an idea and intention before we enter the field - it helps us to finish the research successfully. But the principle of qualitative research is that we must be prepared to give up our idea, to change it or adapt it according to the continuous findings. From this point of view, there are two approaches within the qualitative methodology. One is more radical and states that it is not necessary to have any hypotheses before we enter the field - Glaser's Grounded Theory, and the hypotheses should arise in the field. This theory is grounded because it lies in experience from which it directly proceeds - it is therefore strictly inductive approach to theory formulation. The main advantage here is that the researcher minimizes the influence of his a priory expectations and he lets the "data speak". Other qualitative researches are not so radical and they usually acknowledge that the researcher enters the field with certain idea (which can also fulfil the function of hypothesis) but it is necessary to work further with the hypothesis by means of information from the field - i.e. to develop, elaborate, change it or even radically reject it and formulate a new hypothesis.

# Formulation of Research Paradigm. Description of the Theoretical-Empirical Starting Point and Choice of the Topic: Formulation of the Frame Idea about Studied Topic

The research topic can arise in various ways. It can be even inspired by acquaintance with an interesting theory or paradigm that interests the researcher to such an extent that within the familiar topics and problems he will see new dimensions of them and new opportunities of research. In qualitative research, the research problem itself is very closely connected to the theory. Also, if the theory was presented by the genesis of the research topic, it is necessary to try to combine this theoretical standpoint with the research topic and to carry a view on the problem that will lead to a more specific research problem. If the theory was not present at the conception of the research topic, it is even more pressing to try to find an appropriate theoretical starting point within which this topic would be meaningfully graspable, and within which the formulated research problem would be meaningful and solvable. Finding of the appropriate theoretical starting point also often indicates the used method. If the discourse theory interested me and I think that this theory is good and usable within my topic, and it guides me to interesting research questions, it determines which method I will use in my research. It will definitely be some form of discourse analysis. But as we have learned in the part about discourse analysis, there are many ways to approach the discourse analysis. It is therefore necessary to get to know these variants and to decide for the set of techniques which will be most appropriate not only with regard to my research interest but also with regard to practical possibilities of research realization.

There can exist several **concepts** within the chosen theoretical perspective. Concept is more established than theory and the choice of concept depends on its applicability on given research problem and especially on my research data. If it is suitable, we can combine concepts from different theories in the research, it can even be seen as an innovative and helpful step if we show in our research the combinability of certain theories and concepts which was so far not uncovered and used by empirical analysis. It must be pointed out that in qualitative research we often work with **middle range theories** which means not to work with such theories that try to explain principles of functioning of society as a whole or of social behaviour as such, but only its part.

The role of preliminary reading and finding the appropriate theoretical starting point is therefore crucial. Kauffman (2010) points out that we have to read neither little nor much (which is not easy to determine). If we read too little, we are not familiar enough with the theory or concept which we want to work with and then we do not know what we should focus on and what questions we should ask in the field. If we read too much, we lose time because we do not confront what we have learned with the empirical findings and it can happen that we will go in the "wrong direction". Kauffman recommends to shorten the preparatory phase in the qualitative research (in comparison with e.g. quantitative research) to avoid the problem that after detailed theoretical and methodological preparation it emerges that the chosen perspective is not appropriate and applicable to the data that we will acquire in the field. It is therefore suitable to go into the field relatively early and to confront the concepts and assumptions with the experience.

#### Formulation of Research Question and Research Goals

If we have formulated the research topic we have to transform it into more specific research questions, or research goals. To define them it is necessary to draw from theory but also to go into the field. Research questions will become more specific only by contact with the field, i.e. with the studied problem. During research, these goals and research questions can change, some will be abandoned, some added. The reason for this can be, for example, finding that one of the questions cannot be answered by means of given data.

# Proposal of the Case that Represents the Chosen Research Problem and the Explanation of its Choice

The choice of research case/sample can stand in some cases at the beginning of the research. It means that the whole research was motivated by interest in the given case. But in other situations the choice of a specific case or sample is not a primary impulse and the choice of it is not so simple. We can proceed in various ways. We can choose a typical case - such a case which we assume will have relevant features. We can then compare this case with another one that we consider untypical. Much depends on the chosen research method and research questions, and on

the elements of the research, whether they are people - respondents, organizations - some broader groups, or e.g. texts as in the discourse analysis. If we analyse texts it is important to define what texts these will be - e.g. we should define the time of its origin or to define in other way the concerned sphere. If we work with individual respondents - it is like by the interview - we should avoid the evident imbalance of the sample. If we choose organization(s) various criteria can be taken into consideration: place, size etc. depending on research problem. In all cases we should take into account the practical possibilities of research realization: the demands in terms of time and finance or simply good or worse access to field. If we want to do research in an organization it is important which organization allows us to do the research and offers the best conditions.

# Proposal of the Data Collection Method and Technique and Explanation of the choice

The idea of method is present very often already by the birth of the research subject, however the gradual specification of research tool takes place during the whole research because all the mentioned methods have more variations and techniques which can be combined. Combined can be also the individual methods. But we must give good reasons for the choice of methods and techniques in the project, particularly with regard to research problem and we must constantly get back to the question of what the relationship between our research question and method is i.e. is it possible through this method and its techniques to come to the answers of our research questions? Do we not make a methodological mistake: unjustified generalization or unjustified comparison, etc.

# Data Analysis and Interpretation in Qualitative Research

While with quantitative research we divided the research process into three phases: preparatory phase, realization phase and phase of analysis and interpretation, in qualitative research these phases take place simultaneously. David Silverman quoted Coffey and Atkinson (1996, In Silverman 2000: 131): "analysis is an activity going through the whole research project. Analysis simply isn't one of the later phases of research after which an equally independent phase of writing the results follows". Silverman recommends analysing the data right from the beginning, even before the beginning of the research itself. How is it possible? Silverman recommends confrontation with the existing data either in the form of existing records, documents concerning the problem or of data from other researches with the same or similar topic. This is important from the point of view of a closer acquaintance with the problem and its broader links, but also from the point of view of "training" in approach to data and from the point of view of acquisition of certain research self-confidence, the confirmation that we can work with data and find interesting conclusions.

If we do not use these "other" sources of data and try to get our own data, it is equally important to analyse them from the beginning and also continuously. If we do for instance in-depth interviews we must confront the data with our theory or concept already after the first interview and also to confront the used research tool with research question. If it is necessary, we can modify them both, although we shouldn't expect too much from the first interview and we shouldn't give up our initial intention too early only because we haven't found anything relevant for the research yet. But if we run into the same problem during further research, something is probably not working. It might be necessary to change the approach to data, or the nature of data, or just to search another more appropriate theory or concept. It is the same also by other methods and techniques of qualitative research. If we analyse texts, it is to some extent natural that we analyse them continuously and so it is by the stay in the field.

On the other hand, we must keep certain distance from our data. What we analyse and interpret during the first phases of research, will probably not be substantiated and conceptual enough to write the final work of the research. It is only the first step. It is good to put our data, notes or partial analyses and the first attempts to interpret them, later to repeated critical reassessment in the light of new research conception. The analysis takes place in several phases, or cycles.

Specific in qualitative research is the view of studied data. Qualitative methodology, as we have already mentioned, is a form of understanding sociology - and its primary intention is to reconstruct the meaning attached to their action by actors themselves, or to analyse the ways how this meaning and "truth" is negotiated and constructed in a certain society or culture. Therefore the studied data are seen as not only statements and descriptions of outside reality (facts, events etc.) but also as ways of description and construction of an outside or inner world, as on individual or "shared" and negotiated stories. The aim is to understand how and why this story is constructed this way, to take over the view of those we study. On the other hand, the initial understanding "from inside" requires later taking the stand of a "foreigner" and being able to extricate material from the seemingly commonplace or uninteresting-ness of data in what, for example, Schütz calls a natural attitude. If the researcher would stay in this natural attitude he would never find anything interesting because everything would seem natural and sociologically uninteresting to him. The metaphor of a foreigner is therefore ideal for understanding the meaning of interpretation in the qualitative paradigm: to be able to look at seemingly natural things through the optics of the uncommonplace, to search a socially rooted origin and to problematize it. The foreigner enters the "other world", other culture. It does not have to mean to cross the borders and enter a different culture in geographical sense, but the view on studied community and society through the view of a foreigner. For instance when we enter the world of football fans for the first time, it can seem strange, the gestures, slang and rules of behaviour might seem strange or at least not familiar to us. If we decide to do an ethnographic research in the community of football fans, our intention will be to break this barrier and to try to "identify" ourselves with the position of those we study, to understand their world from their point of view. And also by the analysis and interpretation of data and by the writing of the final work we have to grasp the view of the "foreigner" again but this time the view of an informed foreigner who can tell us more about the others than even they can do. And it is thanks to the distance and other point of view. As we have mentioned before, this point of view is

connected with a definitive theory of concept by means of which we grasp the reality.

#### Validity and Objectivity of Knowledge in Qualitative Research

We have already spoken about validity as the truthfulness of research or its parts in the section about quantitative research. In qualitative research, validity has its specifics. In qualitative research, the feeling of non-validity is often connected with the unconvincingness of research results, or with doubts regarding the research procedure. The audience to which the results are presented can for example question the choice of data that should confirm the conclusions. They can state that the researcher has chosen only that data which confirms his conclusion and left out the others. As Silverman (2000) writes, this problem is called "anecdotism".

Mehan identified three weak points of ethnographic researches:

1. Common field researches usually have an anecdotal character. Research reports include several exemplary examples of behaviour that the researcher has chosen from field notes.

2. Researchers only seldom give reasons or criteria why they have chosen certain cases and not the others. As a result of this, it is hard to judge the typicality or representativeness of the cases and findings drawn from them.

3. Research reports do not keep the materials used for the analysis. When the researcher abstracts the data from raw materials for the purpose of presentation or summarizing findings, the material in original form gets lost and therefore it is not possible to think about alternative interpretations of the same materials (Mehan 1979, In Silverman 2000: 189).

Silverman proposes 5 ways of critical thinking about qualitative data analysis for the problem of anecdotism. These are:

- Refutability principle
- Method of systematic comparison
- Complete data processing
- Deviant case analysis
- Use of appropriate tables

The refutability principle is very close to what Popper called "critical rationalism". It requires that we try to disprove assumed relations between phenomena. Only in the case where we cannot disprove the existence of the connections between them, we can speak about objective knowledge. Popper's requirement is relevant in the relationship to quantitative research where it is applied especially as an attempt to disprove "illusory correlations". Silverman explains how the qualitative researchers can meet Popper's criterion through the combination of use of the next four ways of critical thinking about qualitative data analysis. The method of systematic comparison requires the exploration of all fragments of data that appear within one case, or eventually in comparison with other appropriate cases. Complete data processing means that all cases from data are included in the analysis and not only those which confirm or disprove the hypothesis. Deviant case analysis means focusing on deeper exploration of cases that defy our analytical conclusions and detailed analysis or explanation of this deviation or the reformulation the research conclusions on the basis of these deviant cases. With appropriate use of tables Silverman means the use of simple quantification where it is appropriate (Silverman, 2000: 191-198).

Despite the fact that qualitative approaches in general take over postmodern relativism in the sense that they exclude the chance to get to know social phenomena "objectively", it does not mean that the question of objectivity of social-scientific knowledge here is irrelevant. We do not want to give up the opportunity to assess the individual researches and acquired findings, although not in the sense of which of them is an objective description of reality, but in the sense of whether the given research provides a meaningful and substantiated image of social reality. Brian Fay gives a good example - a map. A cartographer can create a map of a territory. This map is not the only possible or the only correct map of this territory. There can be many maps: a political map, map of waters, industry, etc. according to the focus of the cartographer. Equally, there can exist many descriptions, stories and analyses of one phenomenon that can differ and in spite of this, they can be an appropriate and faithful description of the studied reality. But it does not mean that all of them are considered equally as good. So as the map can be worse or better because

one can be more precise than the other, it is the same with the socialscientific knowledge acquired through qualitative research.

So we do not give up the assessment of research quality but we apply the criteria of objectivity **on the process rather than on its result.** Here objectivity is related more to the method and can be understood as procedural adequacy, dutifulness of research. We deal with the question of whether the chosen procedures were appropriate and whether the researcher exceeded the subjective inclinations and prejudices by the assessment and interpretation of acquired data, whether he took into consideration also the alternative explanations etc. From the point of view of the evaluator it is then assured credibility of assumed procedures and the knowledge that follows from them. The way to increase the research value of knowledge is to ensure the clashes of different perspectives - to actively search other opinions and to claim the reaction.

# **Bibliography:**

- AJDUKIEWICZ, K. 1953. Zarys logiki. Książka pomocnicza dla nauczyciela. Warszawa: Państwowe Zakłady Wydawnictw Szkolnych
- AUSTIN, J., L. 2004. Ako niečo robiť slovami. Bratislava: Kalligram
- BENEŠ, V. 2008. "Diskurzivní analýza." In Drulák, P. a kol. Kvalitativní metodologie v politologii a mezinárodních vztazích. Praha: Portál: 92-124
- BERGER, P. LUCKMANN, T. 1999. Sociální konstrukce reality. Pojednání o sociologii vědění. Brno: CDK
- CAMPBELL, D. 1998. Writing Security: United States Foreign Policy and the Politics of Identity. Minneapolis/Manchester: University of Minnesota Press/Manchester University Press.
- CHAMBERS. E. 2003. "Applied Etnography." In Denzin N. K. and Lincoln Y.S.. Collecting and Interpreting Qualitative Materials. London: SAGE Publications Ltd.: 389 -418
- DENZIN, N.K., LINCOLN, Y.S. 2012. "Methods of Collecting and Analyzing Empirical Materials." In Denzin, N.K., Lincoln, Y.S. (Eds.) *Collecting and Interpreting Qualitative Materials*. London: SAGE Publications Ltd.: 47-60
- DISMAN, M. 1993. Jak se vyrábí sociologická znalost. Praha: Karolinum
- DISMAN, M. 2002. Jak se vyrábí sociologická znalost. Praha: Karolinum
- DROZDA, V. 1996. Medzi interakciou a mlčaním. Bratislava: IRIS
- DRULÁK, P. 2008. "Epistemologie, ontologie a operacionalizace." In Drulák, P. a kol. *Kvalitativní metodologie v politologii a mezinárodních vztazích*. Praha: Portál: 15-28
- FAIRCLOUGH, N. 1992. "Discourse and Text: Linguistic and Intertextual Analysis within Discourse Analysis", In *Discourse & Society* 3 (2): 193-217
- FAIRCLOUGH, N. 1995. Critical discourse analysis: The Critical Study of Language. London; New
- York: Longman
- FAY, B. 2002. Současná filosofie sociálních věd. Multikulturní přístup. Praha: SLON

- FINDOR, A. 2007. "Sociálna konštrukcia homofóbie." In *Inakosť a diskriminácia*. Odborný seminár organizovaný občianskym združením Iniciatíva Inakosť, ako súčasť projektu SME TU dni inakosti, 19.4.2007
- FOUCAULT, M. 2006. Rád diskurzu. Bratislava: Agora
- GIDDENS, A. 1999. Sociologie. Praha: Argo
- GOFFMAN, E. 1999. *Všichni hrajeme divadlo. Sebeprezentace v každodenním životě*. Praha: Nakladatelství studia Ypsilon
- GRAHAM, P., KEENAN, T., DOWD, A.M. 2004. "A Call to Arms at the End of History: A Discourse–Historical Analysis of George W. Bush's Declaration of War on Terror." In *Discourse & Society* 15 (2-3): 199-221
- HIRNER, A. 1976. Ako sociologicky analyzovať : state k sociologickému spracúvaniu primárnych evidenčných dát. Bratislava: Ústav školských informácií
- JALBERT, P.L. 1999. "Media Studies: Ethnomethodological Approaches." New York, Oxford: University Press of America; International Institute for Ethnomethodology and Conversation Analysis.
- JANDOUREK, J. 2001. Sociologický slovník. Praha: Portál
- JEŘÁBEK, H. 1993. Úvod do sociologického výskumu. Praha: UK
- KABELE, J. 1996. "Sociální konstruktivismus." In Sociologický časopis. 32 (3): 317-337
- KAUFMANN, J.C. 2010. Chápající rozhovor. Praha: SLON
- KELLER, J. 2004. Dějiny klasické sociologie. Praha: SLON
- KOŘAN, M. 2008. "Jednopřípadová studie." In Drulák, P. a kol. Jak zkoumat politiku. Praha: Portál: 29- 61
- KUBÁTOVÁ, H. 2006. *Metodologie sociologie*. Olomouc: Univerzita Palackého
- KUSÝ, I., VAVRINČÍKOVá, L. 2009. Ako písať a ako nepísať záverečnú prácu. Banská Bystrica: Občianske združenie Pedagóg
- KUHN, T.S. 1997. Struktura vědeckých revolucí. Praha: Oikonmenh
- LACLAU, E., MOUFFE, Ch. 2001. *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. London, New York: Verso
- LAMSER, V. 1966. Základy sociologického výzkumu. Praha: Svoboda

- LAZARSFELD, P.F. 1935. "The Art of Asking Why." In *National Marketing Review* 1: 26-28
- LEVIN, W. C. 1991. *Sociological Ideas*. Belmont, California: Wadsworth Publishing Company
- LIGHT, D., KELLER, S., CALHOUN, C., 1989. Sociology. New York
- LIPOVETSKY, G. 1999. Soumrak povinnosti. Bezbolestná etika nových demokratických časů. Praha: Prostor
- NOVOSÁD, F. 1997. Osud a voľba. Bratislava: IRIS
- MILLS, Ch. W. 2002. Sociologická imaginace. SLON: Praha
- ONDREJKOVIČ, P., 1996. Hermeneutické metódy interpretácie v sociologickom výskume., In Sociológia 28(3): 207-218
- PHILLIPS, N., HARDY, C. 2002. *Discourse analysis: investigating processes of social construction*. London: SAGE Publications Ltd.
- SCHENK, J. 2003. "K hlavným princípom Lazarsfeldovej koncepcie metodológie." In Schenk, J. a kol. *Paul Felix Lazarsfeld. Návraty k myšlienkovému dedičstvu*. Bratislava: SOFA
- SILVERMANN, D. 2005. Ako robiť kvalitatívny výskum. Bratislava: Ikar
- SOPÓCI, J., BÚZIK, B. 1995. Základy sociológie. Bratislava: SPN Mladé letá
- SOPÓCI, J. 2007. Teórie sociálnej zmeny. Bratislava: UK
- SURYNEK, A., KOMÁRKOVÁ, R., KAŠPAROVÁ, E. 2001. Základy sociologického výzkumu. Praha: Managament Press
- SWOBODA, H. 1977. Moderní statistika. Praha: Svoboda
- ŠUBRT, J. 2001. Postavy a problémy soudobé teoretické sociologie. Praha: ISV
- TORFING, J. 1999. *New Theories of Discourse. Laclau, Mouffe and Žižek.* Oxford: Blackwell Publishers Ltd.
- TULMETS, E., STŘÍTECKÝ, V. 2008. "Etnografie." In Drulák, P. a kol. Kvalitativní metodologie v politologii a mezinárodních vztazích. Praha: Portál: 149 - 176
- VENNESSON, P. 2008. "Case studies and process tracing: theories and practices." In Della Porta, D., Keating, M. Approacheas and Methodologies in The Social Sciences. A Pluralist Perspective. Cambridge: Cambridge University Press: 223-239

- WALLIS, W. A., ROBERTS, H. V. 1956. *Statistics: A New Approach*. Glencoe: The Free Press
- WELSCH, W. 1993. Estetické myslenie. Bratislava : Archa
- WIATR, J. 1964. *O sociológii (sociologická teória a metódy sociológie)*. Bratislava: Osveta
- WILSON, T.D. 2002. Alfred Schutz, phenomenology and research methodology for information behaviour research. In A paper delivered at ISIC4 - Fourth International Conference on Information Seeking in Context, Universidade Lusiada, Lisbon, Portugal, September 11 to 13, 2002; available at: http://www.informationr.net/tdw/publ/papers/ schutz02.html [30.9.2015]
- WODAK, R., MEYER, M. 2009. "Critical discourse analysis: history, agenda, theory and methodology." Pp. 1-33, In. Wodak, R. Meyer, M. (eds.) Methods of Critical Discourse Analysis. London and Los Angeles: SAGE Publications Ltd.

# Title: Methodology of Social Science Research

Authors:	Ivan Kusý, Jarmila Androvičová
Publisher:	Publishing house of the Matej Bel University in Banská Bystrica - Belianum
Number of pages:	166 pages (9,28 authorial sheets)
Number of copies:	100 copies
Edition:	first
Format:	A5
Published in:	2015
Print:	EQUILIBRIA, s.r.o.

© Matej Bel University in Banská Bystrica, 2015

ISBN: 978-80-557-0971-0 EAN: 9788055709710