FOLIA 296

Annales Universitatis Paedagogicae Cracoviensis

Studia ad Didacticam Biologiae Pertinentia 9 (2019)
ISSN 2083-7276
DOI 10.24917/20837276.9.19

Sławomir Trusz

Do qualitative researchers know what they see and the respondents know what they say? On the quality of qualitative data

Introduction

Analyzing social phenomena, qualitative researchers – also educators – collect empirical material using specific methods and strategies. The most popular are non-standardized participant observation (Lichtman, 2013; Lofland et al., 2005) and in-depth qualitative interviews (Kvale, 1996; Rubin, Rubin, 2012). These methods are characterized by a high flexibility of research procedures (Hammersley, Atkinson, 2007).

Paying attention to contextual factors, e.g. unexpected presence of third parties, clearly affecting the behavior of the observed persons/respondents, the researcher may spontaneously modify the procedures of obtaining information, e.g. by abandoning selected categories of behavior or conversation topics for others – adequate from the point of view of the emotional state and needs of subjects. In other words, advocates of qualitative orientation in social sciences, in contrast to quantitative researchers, adapt their behavior to the ongoing action/events like on the film set (Lofland et al., 2005; Silverman, 2009).

The freedom of action partly results from the postulate of playing the role of a "naive co-participant in the analyzed events". Undertaking the tasks characteristic for the members of the investigated community, the qualitative researcher observes and talks with the interaction partners without making any preliminary assumptions about the character of the analyzed phenomena (Anderson-Levitt, 2006).

Using non-standardized research methods and strategies makes it possible to realize objectives specific for the qualitative orientation, i.e. to discover meanings/senses given by people to personal experiences, and then to present them in cognitive and linguistic categories characteristic for observation/interview participants. Clifford Geertz (1973), an eminent cultural anthropologist, emphasizes: "believing, with Max Weber, that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs, and the analysis of it to be therefore not an experimental science in search of law but an interpretive one in search of meaning" (ibid: 5).

Indeed, in contrast to idiographic studies, in quantitative (nomothetic) ones, the accent is placed on discovering objective – i.e. investigator-independent

– regularities. Experimental research is organized according to strictly defined scenarios, the course of which is determined by the tested causal relations (Coleman, 2018). The basic rule of experimental research is that "their author is not important, but how the research is conducted". The strictly described course of the study allows to repeat it, and the results of the replications can be summarized in a meta-analysis (quantitative synthesis of the results of research on similar topics and with a similar methodology, based on which the average effect size of specific factors for selected outcome variables is quantified) (Glass, 2006; Schoenfeld, 2006).

The problem of data reliability and validity in quantitative vs. qualitative research

The control of measurement error and attempts to minimize it cause that in quantitative studies the reliability of results, and consequently – the credibility of the proposed conclusions is high. On the other hand, as the precision of measurement increases, the ecological validity of the results decreases. It means that the subjects' behaviors in experiments, meticulously recorded in artificial laboratory conditions, may not be repeated in everyday life situations (Nestor, Schutt, 2012). For example, pressing selected computer keyboard buttons in the Implicit Association Test as fast as possible, despite flawless control of response times, often is not related with actual decisions made by the subjects to help others or, on the contrary, their negative treatment (Hofmann et al., 2005).

The problem of reliability and validity of measurement in quantitative research is well recognized and discussed in the professional literature (Coleman, 2018; Weathington et al., 2010). On the contrary, in qualitative research, intentional elimination of standardization of research procedures and mathematizing the variables, and consequently – non-quantitative, often unique nature of the collected empirical material, causes that the discussed problem is still a source of numerous polemics and controversies (Kirk, Miller, 1986; Baumgarten, 2010).

Avoiding, for the reasons given above, statistical interpretation, the reliability of qualitative analyses, can be defined as the precision of the prepared descriptions, and the validity as the accuracy of the interpretations/conclusions proposed by the researcher (Flick, 2008). Together, the validity and reliability of analyses determine the quality of qualitative research.

In this light, the descriptions of high school students' behaviors, which show that Bob was the only teenager drinking alcohol, should be regarded as unreliable (imprecise), even though all the boys observed during the school event drank beer. Similarly, transcriptions of interviews in which selected phrases were omitted, e.g. statements of respondents indicating their negative attitude towards ethnic or sexual minorities, may be regarded as unreliable. On the other hand, the interpretations that drinking alcohol at school events is a marginal phenomenon, concerning only students from working-class families represented by Bob, should be regarded as inaccurate (unjustified). The analysis shows that interviewers are mostly people who are tolerant and open to differences may be considered similarly incorrect. Indeed, qualitative research carried out in this way would be of dubious quality.

Taking into account the given definitions of reliability and validity of the qualitative analyses, it is worth to point out and discuss the significant limitations of empirical material collected using basic methods/strategies of idiographic research, i.e. non-standardized participant observation and free-form interview. To a great extent, they result from cognitive and personality limitations typical for people, which makes them universal. It is worthwhile for researchers to be aware of their potential impact while they discuss the obtained results.

Do qualitative researchers know what they see? The change blindness phenomenon

The abovementioned limitations can be illustrated by the results of intriguing experiments by Daniel Simons and his colleagues on change blindness phenomenon (Chabris, Simons, 2011). Demonstrating its massive impact, the authors have published research materials on the web.

In the first movie (https://youtu.be/vJG698U2Mvo), a group of students, dressed in white or black T-shirts, pass the ball to each other. The task of a person watching the video (an unbiased observer) is to determine the number of passes made by students dressed in white T-shirts. During the recording, a person wearing a gorilla suit passes through the middle of the corridor. The "Gorilla" stops among the players and ostentatiously hits the chest with his fists. About 50% of observers do not notice the gorilla, focusing on counting passes made by people dressed in white T-shirts.

In the second video (https://youtu.be/1nL5ulsWMYc), the observer's task is to determine which element of the scene presented in the movie changes. Most people are not able to point out even the slightest change when they watch the video the first time. In fact, the change occurs – in the lower left corner of the screen, slowly, second by second, a stone emerges from the background. During the second presentation, most observers easily identify the place where the change occurs – perhaps because the stone flickers in the replay.

Indeed, it is extremely difficult for people to register gradual, not rapid changes. A good example is the long-term weight loss process, e.g. 10 kg per 1 year. People who have managed to lose unnecessary weight may not notice these changes. However, the same change will be noticed immediately, if the observer has not had contact with the person for a year, and judged as a clear, but not gradual one.

In the third video (https://youtu.be/ubNF9QNEQLA), the observer's task is to record the changes in the movie set that happen during the film. Noting the behavior of actors – individuals suspected of the murder of Lord Smithe and an officer conducting the investigation, observers usually notice to five changes. In fact, there are 21 modifications – virtually all the elements of the frame are changed, even Lord Smithe's corpse.

The consequences of limitations of non-standardized observation

Omitting relevant data

Considering the results of Simons and his colleagues' experiments, it is worth asking whether a non-standardized observation (free, without a plan, in other words – preliminary assumptions about its course and recorded content) is possible? Undoubtedly yes, but one should remember about the consequences, i.e. omitting the mass of relevant data for the analyzed problems and, as a result, reducing the reliability and validity of qualitative analyses and proposed conclusions.

The observer can easily answer the question whether in the first video, a man dressed as a gorilla walks in the middle of the corridor while students are passing the balls?, or whether in the fourth video, the officer is dressed in a light or rather dark coat?, if he/she knows what details of the movies should be recorded. In other words, it is possible if he/she has some assumptions about the object/content of the observation.

Analogously, in the case of the second video, the purposeless observation does not allow recording of gradual changes typical for processes analyzed in social sciences, e.g. developmental ones. This implies that only planned observation – with a set goal and course – allows focusing on these elements of the dynamic reality that indicate a change within the tested process (Chabris, Simons, 2011).

Generating pseudo-data

Moreover, the lack of an observation plan, arising from the goals or research problems, may encourage observers to construct non-existent facts (pseudo-data). The source of such pseudo-data is cognitive schemas, e.g. researcher's social stereotypes, scripts and feature schemas (Darley, Fazio, 1980; Neuberg, 1994).

Research shows that after inducing a cognitive schema, e.g. an aggressive Negro, a cunning Jew, a wise four-eyes, etc., the observers are sensitive to information consistent with the previously aroused stereotype (Fiske, Taylor, 2009). Moreover, it happens that in fact, the neutral behavior of individuals is recorded, remembered, and then interpreted according to their cognitive schema (Fiske, Taylor, 2009).

Non-standardized interview – limitations and consequences of using the method

Using a non-directive interview (Kvale, 1996), the researcher establishes topics regularly, i.e. during the conversation, he/she, with great freedom, asks the respondents detailed questions. However, analyzed topics should arise from previously defined objectives or research problems. It is impossible to conduct reasonable conversations with the partners without even a preliminary outline of the main topic of the interview. Therefore, despite the enormous freedom in choosing topics and asking more or less detailed questions to respondents, one can claim about the so-called hidden structure of the interview or the interview plan hidden in the researcher's head.

Some authors argue that "a free-form interview also has a structure, but it is defined at an abstract (theoretical and/or methodological), not a specific level (i.e. questions and answers applied in the interview). Therefore, describing it

as 'unstructured', although applied in literature, is misleading" (Stemplewska-Żakowicz, 2009: 55). Therefore, if the researcher "does not have an interview plan and conducts a conversation 'following the subject' without any goal or idea, and if his/her questions arise from what just came to his mind, and after the conversation he/she does not know what to do with the obtained data and bases conclusions on what made a special impression on him/her during the interview or what he/she remembers from it – this is a truly unstructured interview. It can be a very pleasant experience for both sides (…) one can enjoy such a conversation in various ways and probably derive various benefits from it, except for one thing – research benefit" (ibid: 55).

This implies that an interview without structure and plan (i.e. assumptions) is malpractice. The extreme freedom of research procedures can transform an interview into a noncommittal conversation between a researcher and respondents about "everything and nothing". In extreme cases, the conversation may resemble a game of free associations led by the subject – a social exchange of opinions between alleged friends over coffee or cake, not a conversation of a well-qualified researcher with selected respondents, under a defined goal or research problems.

Furthermore, when a researcher intentionally resigns from managing of the interview, the direction of conversation may be defined by respondents, according to their own needs or goals of self-presentation ("see the researcher, how smart, happy, miserable, etc. I am") or other, more or less clear for the subjects – and even more unclear for the researcher – motives (Cialdini, 2009). In conclusion, the assumed lack of control over the course of the conversation may transform the interview into a caricature of the scientific research method.

Regardless of the type of conducted research – quantitative or qualitative – the researcher should respect the universal GIGO rule (garbage in, garbage out, cf. Dunbar, 1996). Using low-quality data, e.g. collected during conversations without any structure (garbage in), the researcher, willy-nilly, has to formulate worthless conclusions about the analyzed phenomena (garbage out). Briefly, the plan of conversation hidden in the researcher's head is also recommended for qualitative interviews.

Questionable quality of interview data. Do the subjects know what they are saying?

The basic source of data collected in interviews is the respondents' declarations (Oppenheim, 2000). Interlocutors can answer questions honestly or intentionally/unintentionally depart from the truth. This means that the reliability and validity of data gathered during the interview may be lower than the information collected under observation (Figure 1).

As the number of paths (mediators) increases, information can turn into gossip, losing credibility. A researcher using an interview does not reach the facts (observable data), but rather their subjective interpretations, given by more or less reliable informants.

Figure 1. Assessment of the reliability of empirical data collected under observation and survey studies

1. The data collected under observation (zero number of mediators)

BEHAVIOR/EVENT (B/E) ← **observation**: the researcher notices B/E.
E.g. Mary hits Joseph in the head ← **observation**: the researcher notices that Mary hits Joseph in the head

2. The data collected under interview (1., 2., n mediators)

BEHAVIOR/EVENT (B/E) ← **observation**: someone (informant) notices B/E ← **interview**: the researcher hears from the informant that B/E occurred
E.g. Mary hits Joseph in the head ← **observation**: someone (informant) notices that Mary hits Joseph in the head ← **interview**: the researcher learns from the informant that "Mary hits Joseph" (reliable informant)

...and even...

BEHAVIOR/EVENT (B/E) ← **observation**: someone (informant 1.) notices B/E ← story: informant # 1. tells someone (informant # 2., 3., n.) about B/E ← **interview**: the researcher learns from informant # 2., 3. that B/E occurred
E.g. Mary hits Joseph in the head ← **observation**: someone (informant # 1.) notices that Mary hits him in the head ← story: informant # 1. tells someone (informant # 2., 3., n.) that "Mary hits Joseph in the head" (reliable informant) ← **interview**: researcher learns from informant # 2, 3, n., that "someone said that Mary beats Joseph" or "Joseph beats Mary" (informant # 2, 3, n. unbelievable)

It should be explicitly stated that declarations about behavior are not the same as the presented behavior. For example, spectacularly obese people may declare their care for body weight, while pushing another hamburger into their mouths. Similarly, members of neo-Nazi organizations can declare a positive attitude towards racial or sexual minorities, etc., at the same time burning a rainbow at Savior Square in Warsaw and screaming slogans promoting violence against "freaks".

It turns out that people often say something different about their beliefs and preferences, and then, just after the survey (e.g. an interview), they do something different (Baumaister, 2010).

The indicated inconsistency can be illustrated by the results of the experiment conducted by this paper's author, where the relationship between the declared readiness to provide selfless help and the provided help was assessed. In the first part of the study the participants (students of the first year of health education) were informed that for taking part in a simple physical exercise, the so-called "chair", they could earn a certain amount of money, i.e. PLN 2 for every 10 seconds of participation in the task. The experimenter measured the time and then informed

the researched person about the prize. The earned money could be assigned to the participants at their discretion: (1) to oneself, (2) to a close relative or (3) to a public benefit organization.

In the second stage, i.e. one week after the control of altruistic behavior (sharing the earned money), the same participants were asked to complete a questionnaire on their attitudes towards helping others. Examples of items of the tool were as follows: "If I have the opportunity, I try to help others", "Financial support for charity institutions is commendable". The statements were assessed on a scale from 1 (strongly disagree) to 6 (strongly agree).

The subjects earned PLN 19.85 on average, while they attributed PLN 19.15 to themselves, gave PLN 0.69 to their relatives and zero PLN to organizations. On the other hand, the same participants declared a high level of readiness to help others (average rating of the claim: "If I have the opportunity, I try to help others", was 5.31) and they stressed that supporting charity institutions is a noble activity (the average score for the statement "Financial support for charity institutions is commendable", was 5.38), transferring a few days earlier as much as PLN zero for their functioning.

An insightful Reader may provide an alternative interpretation of the obtained results. People may agree with the statement that supporting charities is commendable and at the same time not to do so, e.g. due to lack of sufficient funds. In this light, the result under discussion would indicate that there is a clear discrepancy between the assessment of the respondents' self and the assessment of their obligations towards others. The crucial point is that the participants had at their disposal – modest but nevertheless – financial resources that could be transferred to a selected public benefit organization. However, any amount was PLN zero. This means that if we want to collect reliable data on people's behaviors, we should not ask about them, but rather observe subjects in various types of situations over a longer time (Jussim, 2012).

The only reasonable exception is when the interesting behavior for the researcher is declarations, i.e. respondents' reactions to the asked questions. However, they should not be confused or identified with behaviors treated as real manifestations of the analyzed attitudes/beliefs of the respondents. Why? It is worth remembering – strong acceptance of the statement (and thus improvement of self-esteem) – "If I have the opportunity, I try to help others", cost the experiment's participants about PLN 0.69. The aim of scientific research, however, is not to therapize respondents at an affordable price, but to make sensible conclusions about the analyzed phenomena based on reliable empirical data.

Selected personal factors of the low quality of qualitative data collected in the interview

There is a group of factors that negatively affect the reliability of qualitative data collected under the interview. The most important factors include: (1) intentional lies, (2) gaps in memory and false memories, (3) interpretational bias resulting from beliefs, (4) low motivation, fatigue and, more broadly, bad psychophysical condition

of respondents and (5) the way of conducting a conversation, especially the use of tendentious questions.

Intentional lies

Respondents may cheat, especially when they are asked about uncomfortable, e.g. sensitive, embarrassing, intimate, etc. matters. Briefly, some topics are a social acceptance factor tinged. Questions: "Do you betray your husband?", "Do you share with others the secrets entrusted to you by your friends?" etc. may cause fear of judgment and/or a tendency for positive self-presentation. Therefore, respondents may be afraid of negative opinions from others, but they may miss the truth or intentionally conceal it.

To access the sensitive data, the researcher should not ask about them directly, but rather talk about them in a camouflaged way, e.g. using the so-called projective or hypothetical questions (Ustjan, 2009a, 2009b). For example, instead of asking: "Do you abuse alcohol?", the following introduction can be used during the conversation: "Sometimes people drink alcohol, e.g. in company at parties, and then start to behave characteristically, e.g. their tongue gets tangled, their movements slow down, they speak louder than usual. When was the last time you experienced this state? How often have you experienced this state in the last week/month/year?".

Gaps in memory and false memories

Besides intentional cheating, respondents, describing events from their own lives, may unconsciously miss out on the truth. Autobiographical memory, i.e. the part of memory specifically relates to facts, events, and experiences from the life of people, e.g. poisoning with mayonnaise during the wedding of a beloved aunt, balloon flight with parents, etc. (Bluck, 2003) – is a creative and selective process. Some information is removed, e.g. as a result of childhood or dissociative amnesia, others are added, although they are not related to the experience of individuals. This means that some of the data stored in the autobiographical memory should be considered as confabulations, i.e. false memories of events in the life of the examined person.

It turns out that approx. 20–30% of respondents are susceptible to developing realistic confabulations, although not related to life experiences. For example, do you remember when you fell and hurt your knee painfully in your childhood? Yes? But are you sure that it happened to you, not to your siblings or a close mate? People have great difficulty in recognizing which autobiographical events occurred in their lives and which of them were creatively developed for the needs of the life story (narration) presented in a given moment.

In a spectacular experiment of Loftus and Pickrell (1995), the participants were presented with 4 realistic stories from their early childhood (4–6 years of age). Three of them were true, one was false. The false story concerned the alleged loss in a hypermarket, crying of a child, searching for parents and happy-end, i.e. finding the child. Respondents were asked to reproduce all stories and visualize them as

accurately as possible in their memories if the participant had problems with remembering the events.

After 2 weeks, the experimenter returned to the subjects and again asked them to recall four events and describe them. About 30% of the participants remembered a false memory. Briefly, simple manipulation allowed for the implantation of foreign (!) memories, so-called "memory implants". One of the participants described the "implant" as follows: "I mean this is very vague, remember the lady helping me and Tim and my mom doing something else, but I don't remember crying (...). I remember being with the lady. I remember going shopping" (ibid: 723)..

In another, equally intriguing experiment (Wade et al., 2002), participants were presented with 4 photographs from their childhood, and then asked to recall the events that were documented, and describe them as well as possible. One of the photos was prepared and related to a balloon flight, an event that never took place in the lives of the subjects. Then, at weekly intervals, three in-depth interviews were conducted with the participants on the events shown in the photo.

In the last interview, as a result of active recalling of a fictitious event, 50% (!) of the participants declared that they remembered the balloon flight. One of the participants described the experience as follows: "I'm still pretty certain it occurred when I was in form one (6th grade) at um the local school there (…) Um basically for $10 or something you could go up in a hot air balloon and go up about 20 odd meters …it would have been a Saturday and I think we went with, yeah, parents and, no it wasn't, not my grandmother …not certain who any of the other people are there. Um, and I'm pretty certain that mum is down on the ground taking a photo" (ibid: 600).

During properly conducted interviews, respondents can recall everything – even very painful, but fictional experiences. The results of research on recovered-memory movement (Travis, Aronson, 2014), dynamically developed in the 1980s, provide interesting examples. Using cognitive interviews[1], analytical therapists helped clients to reconstruct traumatic events allegedly repressed from memory – especially incestuous rapes experienced in childhood.

Loftus (1997) describes the story of Beth Rutherford, a woman who, as a result of the active reconstruction of memories (in fact: implanting false memories), remembered that her father, a pastor, together with her mother, raped her regularly for 7 years (from 7 to 14 years of age). During the therapy, Beth developed a memory of her father getting her pregnant twice and having a hanger abortion (sic!). The court found the parents guilty and sent them to prison for many years. Also, the "victim" received $1 million in compensation for her moral damages.

Most vulnerable to generating false memories are suggestible people with vivid, plastic imagination. Therefore, children very often can confabulate, confusing memories of real events with fiction. Moreover, it is easiest to implant memories

1 i.e. an interview in which the researcher directs the respondent's attention to small details of the event and its context. Moreover, the interviewer encourages the respondent to describe the events from different points of view, to reproduce them in the mind many times. In this way, the "untruth" repeated many times is transformed in autobiographical memory into "misty truth", "truth almost clear and at hand" and finally – "subjectively true truth".

that come from close, important and similar people, such as relatives, and memories concerning family and peer relationships (which is particularly evident in twins whose biographies can strongly intertwine and complement each other. In such conditions, after some time, it is difficult for twins to identify which of the events belong to them in contrast to sibling experiences). Memory implants must also be probable. It seems completely impossible to tell a person with both legs that he/she lost the right one in the mouth of a shark in his/her childhood.

Interpretative bias arising from beliefs

Beliefs about the physical and social world, e.g. the attributes of representatives of various social, racial, age, etc. groups, can be a source of prejudice or inadequate – too high or too low – expectations (Jussim, 2012; Trusz, Bąbel, 2016).

Therefore, judgments about the same object (events, human behavior) made by two independent observers with distinct opinions may differ radically. For example, beating black people on their faces in some circles is seen as a sign of heroism in the fight for a better homeland, while in others it is seen as human baseness and barbarity (Nelson, 2009; Whitley, Kite, 2010).

Low motivation/bad psychophysical condition of the subjects

Respondents may not need to answer interview questions reliably. Low motivation is particularly possible when the respondents' attitude towards the researcher or interview issues is negative, e.g. when the respondents are forced to participate in the survey or when the structure of the asked questions is wrong.

Constructional errors may concern e.g. double or even multiple negations in the content of statements addressed to the respondents ("I never thought it would not happen"), too long statements ("My teachers often repeated during classes that the division of household duties into typically masculine, e.g. car repair or work in a mine, and typically feminine, e.g. preparing meals or feeding children, is natural and should be accepted, even if egalitarian values dominate in society"). Such statements (especially when there are many) may discourage or irritate respondents. This increases the chance of obtaining perfunctory, avoiding, etc. answers.

Moreover, the source of low-quality empirical material may be a bad psychophysical condition of the respondents. Too long or complicated interviewing points may discourage informed, consistent and in-depth speaking on specific topics. Therefore, interviews should be conducted with respondents who are refreshed and interested in the issues discussed in the survey.

Interview conduction – biased questions

The source of low-quality subjects' responses could be inadequately asked questions. Example #1: The question: "Do you still abuse alcohol?", cannot be answered: "I have not had any problems with alcohol so far". The negative response indicates that the respondent has been abusing alcohol in the past, whereas the positive one

suggests that he or she is still systematically drinking. This is due to biased questions (alternative and suggestive), which the researcher should avoid during the interview.

Example #2: In the Loftus and Palmer (1974) experiment, the participants were shown a film concerning a car accident, and then some people were asked: "About how fast were the cars going when they smashed into each other?" The others were asked: "How fast were the cars going when the collision occurred?" (or bump, hit and contact, depending on the condition). Minor manipulation of the question content had a significant impact on the estimation of the cars speed – smashed cars, according to the participants, were driven at an average speed of 40.5 km/h. The average ratings of people in the "collision", "bump", "hit" and "contact" condition were lower: 39.3; 38.1; 34.0 and 31.8 km/h, respectively.

People think the way of using language, i.e. which words were used in the interview, often affects the verbal subjects' responses. It is worth noting that in the Loftus and Palmer (1974) experiment, the participants did not lie – they answered according to the beliefs concerning the speed of the observed cars. Differences in assessments were imposed by arousing different cognitive patterns related to the term "smash" vs. "collide", "contact", etc.

In summary, each of the discussed factors, i.e.: (1) intentional lies, (2) false memories and confabulations, (3) interpretational bias resulting from beliefs, (4) low motivation and bad psychophysical condition of respondents, and (5) inappropriate way of conducting a conversation, negatively affects the quality of empirical material collected during the interview, undermining the reliability (precision) and validity (sense) of the proposed interpretations, and consequently – the conclusions formulated by the researcher.

How to maintain high quality of qualitative data – practical proposals

For high quality of the qualitative data collected under observation and interview methods/strategies, the following procedures or rules should be recommended to researchers: (1) ensuring research internal validity, (2) triangulation of data sources, methods and investigators, and (3) supportive skepticism.

Research internal validity

Ensuring the internal validity of qualitative research is connected with the rule of the dominance of the subject over the research method. According to this rule, the way of collecting and processing empirical data should be adjusted to the content of analyzed phenomena.

Therefore, in the case of external (observable) verbal and non-verbal behaviors, e.g. kisses and the pace of speaking – considered as the symptoms of falling in love and nervousness, respectively – the more optimal method of data collection is observation rather than interview. Therefore, in the case of external (observable) verbal and non-verbal behaviors, e.g. kisses and the pace of speaking – considered as the symptoms of falling in love and nervousness, respectively – the more optimal method of data collection is observation rather than interview. In contrast, if

the subject of the research is difficult to identify, exceeding the possibility of direct observation, phenomena, e.g. thoughts, personality, etc., or about it, e.g. low self-esteem, could be deduced from the statements of the investigated subjects, such as: "I am nobody… I am useless…", etc., the interview seems to be a more suitable method of data collection than observation (Oppenheim, 2000).

Another important criterion for the selection of the method is the purpose of the research. Analysis of cause-effect relations (explaining and predicting phenomena) is possible if the researcher uses the experimental procedure. For example, assessing the extent to which a self-efficacy affect the self-esteem of people, the researcher may manipulate the level of perceived control throughout events, and then check whether the loss or not of control significantly affects the level of self-esteem of the subjects.

On the other hand, the assessment of the co-occurrence of the phenomena (correlations between variables), with no indication of the direction of dependence, is possible when the investigator uses the method of experiment, survey, and observation. For example, analyzing the relationship between openness to experience and time spent reading travel books, a researcher may assess both variables, using e.g. personality and time spent reading books questionnaires, or measuring time spent by individuals on their favorite activities. In the first case, the researcher would use the survey, and in the second case – the observation method.

Finally, a thick description of the phenomena, specific to qualitative research, is possible when the researcher has data collected under non-standardised participant observation or free-form interview, but not under standardized observation/interview or experiment. For example, analyzing the motives of using drugs by students, the researcher may ask about them, e.g. during an interview, or observe in what circumstances, and with what emotions the subject uses a cigarette, alcohol or other drugs.

In summary, the selection of an appropriate research method is crucial for assessing the reliability of the collected empirical data. Inappropriate methods (not adjusted to the research subject), in comparison to adequate ones, are a source of information of lower quality. A hit on the jaw of a black person is a more accurate and reliable manifestation of racism than a declaration of love and friendship expressed by the skinhead against racial differences. Briefly speaking, the data collected based on direct observation, in comparison to information from self-descriptions, are usually more reliable (i.e.: observation > survey).

However, on the other hand, people may behave in a non-authentic way, e.g. knowing that they are being observed, but speaking honestly and clearly on various topics during the interview (i.e.: interview > observation). Of course, the surveyed person may "seem" to be truthful, although he/she can skillfully manipulate information, e.g. presenting himself or herself in the best light or forgetting about certain events (i.e.: observation > survey).

Triangulation

The reliability and validity of qualitative data and analyses based on them can be assessed using the triangulation procedure (comparison, compilation), and within its framework:

(1) triangulation of data sources. Analyzing a specific phenomenon, e.g. domestic violence, the researcher may analyze the behavior of spouses by asking only their wives, only their husbands, wives, and husbands, or wives, husbands, children, and neighbors. Conclusions drawn from various sources (e.g. wife, husband, children, and neighbors) are usually more reliable than conclusions based solely on information from a single (perhaps biased/prejudiced) source. The similarity of information from different sources (e.g. wife and child) confirms their reliability. Descriptions based on reliable data, their interpretation and, consequently, the proposed theory, are characterized by higher reliability (precision) and validity (reasonable) than those based on unreliable data.

(2) triangulation of methods. The researcher may describe or interpret phenomena using data from the interview only, observation only or both methods at once. For example, analyzing the circumstances in which people most often and willingly lie, one can ask them about it or observe in various situations, asking from time to time questions about selected details. Reliable conclusions about the investigated phenomenon can be formulated using data collected via more than one method (it is more difficult for people to communicate untruth through several independent channels, e.g. words and body language, c.f. Babad et al., 1989).

(3) triangulation of the researchers. Qualitative data can be collected and then processed independently by several researchers. It is worth noting that as the level of compliance of paraphrases, categorization, and coding of a text material of several independent researchers increases, the reliability and validity of proposed descriptions and interpretations increases.

Skepticism

People are susceptible to various types of bias and cognitive limitations. For example, interviewers can create false memories. On the other hand, observers, due to e.g. change blindness phenomenon, may tend to register, remember and then recall some information, omitting other important issues.

Moreover, the researcher in specific for his/her way could interpret statements and behaviors of the subjects, consciously or unconsciously using cognitive schemes, e.g. stereotype of a smart vs. dumb pupil, hard-working vs. lazy employee, responsible vs. aggressive “euro-orphan”, etc. Therefore, when analyzing the collected material and then formulating conclusions about the examined phenomena, the researcher should maintain a supportive distance towards himself/herself and his/her own analytical and interpretation skills and data from various sources, remembering about cognitive limitations, personalities, etc., which are characteristic for people.

Conclusions

The author hopes that the presented issues, i.e. limitations of non-standardized research methods in social sciences, as well as strategies for improving the quality of qualitative data, will help readers in planning and conducting their research. Undoubtedly, qualitative research, conducted reflectively, is an equal toward quantitative approach. The combination of the two methodological approaches makes it possible to present the social world in a more complete and therefore more real way than when research is monopolized by only one approach - quantitative or qualitative (Trusz, 2017).

References

Anderson-Levitt K.T., (2006), *Etnography*, [In:] J.L. Green, G. Camilli, P.B. Elmore (eds.), Handbook of complementary methods in education research, Mahwah: Lawrence Erlbaum Associates Publishers, 279–295.

Babad E., Bernieri F., Rosenthal R., (1989), *Nonverbal communication and leakage in the behavior of biased and unbiased teachers*, Journal of Personality and Social Psychology, 56, 89–94.

Baumaister R.F., (2010), The cultural animal: Human nature, meaning, and social life. New York: Oxford University Press.

Baumgarten M., (2010), Paradigm wars – validity and reliability in qualitative research, Auflage: GRIN Verlag.

Bluck S., (2003), *Autobiographical memory: Exploring its functions in everyday life*, Memory, https://doi.org/10.1080/741938206.

Chabris C., Simons D., (2011), The invisible gorilla: How our intuitions deceive us, New York: Broadway Paperback.

Cialdini R., (2009), Influence: The Psychology of persuasion, New York: Harper Business.

Coleman R., (2018), Designing experiments for the social sciences: How to plan, create, and execute research using experiments, London: SAGE Publications, Inc.

Darley J.M., Fazio R.H., (1980), *Expectancy confirmation processes arising in the social interaction sequence*, American Psychologist, 35, 867–881.

Dunbar R., (1996), The trouble with science, New York: Harvard University Press.

Fiske S.T., Taylor S.E., (2009), Social cognition. From brain to culture, Boston: McGraw Hill Higher.

Flick U., (2008), Managing quality in qualitative research, Los Angeles: SAGE Publications.

Geertz C., (1973), The interpretation of cultures, New York: Basic Books.

Glass G.V., (2006), *Meta-analysis: The quantitative synthesis of research findings*, [In:] J.L. Green, G. Camilli, P.B. Elmore (eds.), *Handbook of complementary methods in education research*, Mahwah: Lawrence Erlbaum Associates, Publishers, 427–438.

Hammersley M., Atkinson P., (2007), Ethnography: Principles in practice. 3rd edition, New York: Routledge.

Hofmann W., Gawronski B., Gschwendner T., Le H., Schmitt M., (2005), *A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures*, Personality and Social Psychology Bulletin, 31, 1369–1385.

Jussim J., (2012), Social perception and social Reality: Why accuracy dominates bias and self-fulfilling prophecy, New York: Oxford University Press.

Kirk J., Miller M.L., (1986), Reliability and validity in qualitative research, Newbury Park: Sage Publication Inc.

Kvale S., (1996), InterViews: An introduction to qualitative research interviewing, London: Sage Publication Inc.

Lichtman M., (2013), Qualitative research in education. A user's guide (2nd edition), London: Sage Publications Inc.

Lofland J., Snow D.A., Anderson L., Lofland L.H., (2005), Analyzing social settings: A guide to qualitative observation and analysis (4th edition), Belmont CA: Wadsworth Publishing

Loftus E.F., Palmer J.C., (1974), *Reconstruction of automobile destruction: An example of the interaction between language and memory*, Journal of Verbal Learning and Verbal Behavior, 13, 585–589.

Loftus E.F., Pickrell J.E., (1995), *The formation of false memories*, Psychiatric Annales, 25, 720–725.

Loftus E.F., (1997), Creating false memories, Scientific American, 277, 70–75.

Neuberg S.L., (1994), *Expectancy-confirmation processes in stereotype-tinged social encounters: The moderating role of social goals*, [In:] M.P. Zanna, J.M. Olson (eds.), The psychology of prejudice: The Ontario Symposium, 7, London: Lawrence Erlbaum Associates.

Nelson T.D., (2009), Handbook of prejudice, stereotyping, and discrimination, New York: Psychological Press.

Nestor P.G., Schutt R.K., (2012), Research methods in psychology. Investigating human behavior, Los Angeles: Sage Publication Inc.

Oppenheim A.N., (2000), Questionnaire design, interviewing and attitude measurement. The 2nd edition, London: Continuum International.

Rubin H.J., Rubin I.S., (2012), Qualitative interviewing. The art of hearing data (3rd edition), Los Angeles: Sage Publication Inc.

Schoenfeld A.H., (2006), *Design experiments*, [In:] J.L. Green, G. Camilli, P.B. Elmore (eds.), Handbook of complementary methods in education research, Mahwah: Lawrence Erlbaum Associates, Publishers, 193–205.

Silverman D., (2009), Doing qualitative research, Los Angeles: Sage Publications Ltd.

Stemplewska-Żakowicz K., (2009), *Metoda wywiadu w psychologii*, [In:] K. Stemplewska-Żakowicz, K. Krejtz (eds.), Wywiad psychologiczny. Tom 1. Wywiad jako postępowanie badawcze, Warszawa: Pracownia Testów Psychologicznych, 31–89.

Travis C., Aronson E., (2014), Mistakes were made (but not by me), Harvest Book, Orlando.

Trusz S., (2017), *O znaczeniu badań ilościowo-jakościowych w edukacji. Próba łączenia wody z ogniem*, Studia Edukacyjne, 44, 93–104.

Trusz S., Bąbel P., (eds.), (2016), Intrapersonal and interpersonal expectancies, London: Routledge.

Ustjan D., (2009a), *Instrumenty konwersacyjne. Rozróżnienia, klasyfikacje i badania*, [In:] K. Stemplewska-Żakowicz, K. Krejtz (eds.), Wywiad psychologiczny Tom 2. Wywiad jako spotkanie z człowiekiem, Warszawa: Pracownia Testów Psychologicznych, 131–151.

Ustjan D., (2009b), *Pytania i inne instrumenty konwersacyjne. Zastosowania praktyczne podczas wywiadu diagnostycznego*, [In:] K. Stemplewska-Żakowicz, K. Krejtz (eds.), Wywiad psychologiczny Tom 2. Wywiad jako spotkanie z człowiekiem, Warszawa: Pracownia Testów Psychologicznych, 132–187.

Wade K.A., Garry M., Read J.D., Lindsay D.S., (2002), *A picture is worth a thousand lies: Using false photographs to create false childhood memories*, Psychonomic Bulletin and Review, 9, 597–603.

Weathington B.L., Cunningham C.J., Pittenger D.J., (2010), Research methods for the behavioral and social sciences, New Jersey: John Wiley & Sons.

Whitley B.E., Kite M.E., (2010), The psychology of prejudice and discrimination, Belmont: Wadsworth.

Do qualitative researchers know what they see and the respondents know what they say? On the quality of qualitative data

Abstract

Empirical data collected under qualitative methods and strategies may be more or less reliable and valid. In the article were discussed cognitive and social factors affecting the credibility of information obtained using non-standardized observation and interview The abovementioned issue was illustrated by the results of the studies on the change blindness and false memories phenomena. Finally, the procedures for improving the quality of qualitative data and, consequently, the reasonableness of conclusions regarding the analyzed social phenomena were discussed.

Keywords: reliability, validity of qualitative data, non-standardized observation and interview

dr hab. prof. UP Sławomir Trusz
Educational Microprocesses Laboratory, Institute of Educational Sciences,
Pedagogical University of Krakow
email: trusz@up.krakow.pl