

Chapter 8

Categorising What We Study and What We Analyse, and the Exercise of Interpretation



Dirk Jacobs

8.1 Introduction

A lot of qualitative researchers have a healthy wariness about straightforward categorisation and modelling endeavours undertaken by quantitative researchers. Too often, variables and measurements are too rigid in quantitative analysis to take stock of all the complexity and context-dependency of human behaviour, attitudes and identities. In the worst-case scenario for migration studies, this leads to oversimplification, essentialisation and culturalism. In line with King et al. (1994), I would, however, in this chapter, like to plead for qualitative researchers to take into account that, in terms of challenges of validity and reliability, we have a lot to learn from each other. Acknowledging that qualitative research has its distinctive advantages (Brady and Collier 2004), I will argue that choices in categorisation, case selection and research design are of crucial importance, perhaps even more in qualitative studies than in quantitative studies, even if in both methodological traditions we are confronted with similar challenges. Being transparent and reflecting on the consequences of our choices of categorisation, analysis and interpretation is of crucial importance. It is too easy to think that qualitative research would, by definition, be better equipped in doing justice to the phenomena we wish to study in the field of migration, especially if our research focusses on migrants.

I will first discuss the general challenge of categorisation, then reflect on the issue of working with essentially contested concepts and shed light on the debate whether making ethnic distinctions is legitimate or not. Consequently, I discuss the importance of including a comparative approach in the research design as a good practice.

D. Jacobs (✉)
Université Libre de Bruxelles, Brussels, Belgium
e-mail: dirk.jacobs@ulb.ac.be

8.2 The Challenge of Categorisation

In one of his typical long multi-layered sentences, French sociologist Pierre Bourdieu has well described the challenge of categorisation. It merits to be carefully read (and re-read):

Every science which pretends to propose criteria which are in the best way anchored in reality should not forget that it does not do anything else than registering a particular state of the struggle of classification, that is to say, a particular state of material and symbolic relations of power between those who have an interest in this or that particular way of classifying and who, just as itself, call upon scientific authority to establish in reality and in reason an arbitrary division which it hopes to impose". (Bourdieu 1980, p. 66)

The double hermeneutics – the two-way relationship between ‘lay’ concepts and social science terminology (Giddens 1987) – that are inherent to social scientific activity does not allow us to imagine the constitution of scientific categories that are truly autonomous. Products of a social and political context, they are not immutable. They can be redefined when the context changes or they can lose their relevance when they have been instrumentally used – for instance, when being used more as means of declassification than as means of classification. Categories that want to distinguish social groups and individuals – such as ‘migrants’ – should thus be treated with prudence and large reservations. Caution is particularly relevant in the field of migration studies, in which scholarly choices can quickly become performative acts with political consequences given the high political salience of the topic at hand. In the worst-case scenario, one becomes complicit to forms of elite racism (see the Chap. 13 by van Dijk).

Quantitative researchers sometimes insufficiently take this cautionary advice at heart, qualitative researchers sometimes exaggerate in deconstructing existing categorisations and entrench themselves in critical discursive and self-reflexive analysis. Nevertheless, at the same time, we do need categorisation in our scholarly work. Indeed, one should be able to identify particular groups in order to be able and study them, we need to be clear on what we actually wish to examine and we need to name problems in order to resolve them. Serendipity is never to be totally excluded, but often vague delimitations and exploratory inductive work run the risk to create more interpretation problems than they actually solve.

Clear definitions, replicable categorisations and a targeted research design, in our opinion, is also of considerable importance in qualitative research, particularly when treating highly political topics which are often central to migration studies. It is true that quantitative research often imposes a too rigid research design, often lacks sensitivity to political dimensions, reflexivity and/or individual agency and often runs the risk of oversimplifying reality in its models following a mathematical logic. However, one can also only hope that a qualitative researcher no longer acts as if the methodological and epistemological wars of the 1980s in social sciences are still waging and all insights and approaches from their quantitative colleagues would be useless. If one would refrain from daring to make choices in classifying pupils according to ethnic background for statistical analysis of educational

attainment (as any classification will do some injustice to the complexity of ethnic identification processes), it becomes, for instance, almost impossible to objectify structural discrimination stemming from patterns of segregation in education. That's too big a price to pay.

The linguistic turn in social sciences has been of crucial importance, but as a field, we should also go further than only critically deconstructing (academic) discourse. It should rather push us to be more aware of decision making processes in conceptualisation, categorisation and operationalisation and be clear and open about them. Some post-modern approaches seem to have done our field more harm than good in advocating the stance 'anything goes'. There is still 'science' in 'social science'. Being aware of the political implications and responsibilities of our work is important, but we are not in the business of 'slow journalism' or 'essayism' either. The 'science' in 'social science' is not about having the naïve objectivist ambition to reveal the 'truth', but it is all about being entirely clear about methodological choices and procedures in order to produce the best possible 'knowledge'. Most qualitative researchers at some point struggle with finding the right balance between not opting for a postmodern 'anything goes' – stance on the one hand and not being too limited in possibilities for research and interpretation by following a too formatted cook-book of methodological procedures. The same holds, by the way, for quantitative researchers who have to strike a balance between not following a too rigid research design on the one hand and not going on a fishing expedition for statistically significant effects on the other hand either.

We, of course, have to reflect on the impact of our work. If we only critically deconstruct categories or refrain from using clear categorisation out of fear of essentialising, we run the risk of falling in the trap of conceptual relativism or vagueness, rendering our role as academics a seemingly pointless ivory tower exercise for policy makers and the general public. If we uncritically adopt state categorisations, political or day-to-day discursive categories, we risk reinforcing stereotypes, particular power relations and/or (elite) racism, racialisation and essentialist forms of culturalism. Patrick Simon has nicely formulated this dilemma with which researchers and policymakers are confronted when it comes to the controversial issue of ethnic categorisations:

(...) is it preferable to defend the invisibilisation of ethnic differences in the observational apparatus, while at the same time risking allowing hidden discriminatory practices to prosper, or should one construct categories which, by their simple existence, can potentially reinforce a stigmatizing designation of particular populations?. (Simon 1997, p. 9)

In the post-migration context, especially in countries with liberal nationality legislations, it is clear that the legal category of foreigner will often not be sufficient as a selection criterion when wanting to evaluate, say, the integration of groups of foreign origin. Social scientists (and policymakers) need new categories to be able to count and classify people according to their ethnic origin in order to be able to examine their integration and measure the racial discrimination or processes of social exclusion of which they are victim (or not). The classification of ethnic groups most probably constitutes a necessary tool in the construction of an efficient

policy aiming at equal opportunities and in the struggle against racism. The hesitations with regard to the performative effects of ethnic categorisations, especially in their statistical form, should invite us to epistemological vigilance, but should not frighten us in a way leading to retreat. At the same time, claiming to be studying ‘migrants’ or ‘ethnic minorities’, without being able to clearly define who we are talking about (and why), is highly problematic.

8.3 Essentially Contested Concepts

In the field of migration studies, most – if not all – of our concepts can be considered to be ‘essentially contested concepts’ (Gallie 1956). Gallie, in his seminal article, attributes seven characteristics to essentially contested concepts. They are (1) appraisive (they have a moral component); (2) internally complex (have multiple dimensions); (3) describable in multiple ways; (4) open (subject to revision in particular situations); (5) reciprocally recognised as being contested among contending parties; (6) they have an original exemplar that anchors conceptual meaning and (7) they are subject to progressive competition. Examples are notions like ‘democracy’ or the ‘rule of law’. Collier et al. (2006) provided a very interesting appraisal of Gallie’s approach to the often-unavoidable intersection of normative and empirical concerns in social science. Some have criticized that it leads to self-defeating skeptical nihilism and leads to moral and conceptual relativism (Gray 1977). Others (Waldron 2002) have pointed out that the use of the term ‘has run wild’ when implying that all concepts are contestable. Collier et al. (2006) stress that one probably should not expect all criteria put forward by Gallie to always apply and rather has to use them as a framework for the analysis of conceptualisations. They also pinpoint that Gallie’s framework is useful to study conceptual issues both in scholarship as in ‘real’ politics, which can, of course, also be interrelated. Obviously, for all concepts used in scholarly work a clear definition and operationalisation is needed. Not in all cases the issue of ‘appraisiveness’ arises. With appraisiveness, Gallie refers to particular concepts being ‘achievement’ words, ‘signifying or accrediting some kind of value achievement’. Paradigmatic examples include democracy and social justice. In the field of migration studies, one could think of ‘integration’, ‘inclusion’ or ‘equality’ as being such concepts. Collier et al. (2006) stress that the normative appraisal could also be imagined to be of a negative nature. In our field, we can think of notions like ‘discrimination’, ‘racism’, ‘xenophobia’ and ‘ethnocentrism’ as phenomena the bulk of researchers will wish to combat. The normative and moral dimension is not – or less – outspoken when we use concepts like ‘migration’, ‘migrants’, ‘origin’, ‘culture’, ‘segregation’ or ‘ethnicity’. In those cases, however, the criteria of internal complexity and diverse options for description certainly apply. Internal complexity makes it plausible that different views or definitions can be relevant. The scope and relevance of particular concepts like ‘migration’ and ‘ethnicity’ will largely depend on the (often geographical and cultural) points of reference we adopt and can hence, from context to context, differ. Without

necessarily raising issues of appraisiveness, they are often also intrinsically tied to power relations between groups and/or the functioning of the nation-state and hence historically and politically situated. We should be aware of this and explicitly discuss it, however, without at the same time going into retreat from using clear definitions and delimitations of our concepts all together.

In the field of migration studies, race and ethnicity are central examples of contested or contestable concepts being embedded in particular political, cultural and linguistic traditions. It will have struck any newcomer to the field that there is a strongly embedded tradition to talk about race and race relations in the Anglo-Saxon world – even if there is a wide consensus that race is a social construct without a genuine clearly delineated underlying genetic or biological basis –, while continental European academics will routinely avoid using the word ‘race’, unless they are communicating in English at international venues in presence of Anglo-Saxon colleagues.

Often Anglo-Saxon researchers simply mimic official discourses when making classifications of groups and, for instance, routinely speak about ‘race’. The continental European elephant in the room – going back to the memory of the systematic persecution of Jews and the Holocaust (Jacobs et al. 2009) – is that using the word race in the own language actually signals that you are either a racist or at least someone who believes in a biological basis for racialized differences. As an alternative, one will most often use the concept of ethnicity or ethnic group.

In some national spheres, France being the clearest example, even the use of the notion ethnicity is seen to be controversial, in light of its idiom of nationhood and citizenship (Brubaker 1992) and philosophy of integration (Favell 1998). This is also reflected in the (European) debate on ethnic categorisations and ethnic statistics. The political passions that feed the scientific debate strongly demonstrate that the definition of statistical categories on ethnicity and race is not merely a technical matter. The construction of these categories is influenced by ideologies, visions about nations and visions about interrelations between social groups. An additional element that further complicates the debate is that they are also performative: the use of ethnic categories reinforces the ethnicisation of society.

Once they are socially constructed, these categories gain their own life. Some people are, hence, convinced that the negative performative effects of talking in terms of ethnic categories precludes any legitimate use and one should hence refrain from doing so. I do believe that this is a legitimate position to take by those who fear that academic research actually reinforces racialisation and ethnicisation of social phenomena from the moment one uses the words ‘race’ and ‘ethnic group’, i.e. that it feeds declassification rather than classification of individuals and has an essentialising effect. However, if these same people want to study individuals or groups with a migration history, migration background or part of an ethnic majority or minority group, they will have to resort to alternative designations.

This can either be done by using hyphenated identity markers (for instance “African-Americans” or “Turkish-Germans”), descriptors (“people of migration background” or “people of Moroccan origin”) or euphemisms (“youngsters in deprived areas”). These solutions do not always deliver more precision. Let us give

just one example: in the case of Turkish migrants, for instance, one can quickly embark on politically laden debates on, for example, who is to be classified as a ‘Turk’ or as a ‘Kurd’, and how long one actually remains a person of Turkish (or Kurdish) origin, if one, for instance, was born in the territory of the receiving nation. For how many generations does ‘origin’ remain relevant and how are people of mixed origin to be classified? When asking an auditorium of students whether Barack Obama is white or black, one can get a huge variety of answers, especially after having pinpointed that the former president of the US has a ‘black’ father and a ‘white’ mother, and confronting people with their conscious or unconscious decision criteria.

8.4 To Ethnically Categorise or Not to Ethnically Categorise, That Is the Question

Data on immigrants and ethnic minorities of different European countries are, today, hardly comparable. A number of countries can produce very detailed distinctions with regard to the foreign origin and composition of their population, while other countries feel the production of such data is inappropriate and dangerous. As a result, we have quantitative data on apples and pears and proper comparative social scientific work is often being frustrated. If we want to do serious (quantitative) comparative work with regard to foreign origin groups across Europe, we need comparable operationalization systems which go beyond the simple distinction between nationals and non-nationals. The latter system is biased given the important variation in nationality legislations across Europe. The same challenge holds for qualitative analysis: a foreigner in Sweden or the Netherlands is not quite the same kind of sociological category as a foreigner in Austria or Greece, as it is much more difficult in some countries than others to obtain state citizenship.

The Regulation on harmonized statistics proposed by the European Commission is a step forward to sensitize on this issue, but does not resolve the issue of identifying and quantifying second generation immigrants and longer established ethnic minority groups. In the Netherlands and Nordic countries, a formalized criterion (birth place of parents) has been introduced in the 1990s to pinpoint ethnic minority and foreign origin groups. It has proven to be a useful instrument in documenting discriminatory practices and social exclusion of ethnic groups. At the same time, however, the differentiation between western allochthones or immigrants and non-western allochthones or immigrants (see Jacobs and Rea 2012), has added to the process of racialisation of society. Even worse, in public discourse these statistical notions sometimes function as (dis)qualifying social categories. As a result, the categories have incited quite some resistance among those being classified against their will. One way out would be to allow people to classify themselves as is done for instance in the UK. From a methodological point of view, there are also problems. Country of birth of parents (or grandparents) can only function as a proxy for

immigrant background and ethnicity for a limited time span, especially because people with mixed origins are difficult to classify in a coherent and sensible manner. Furthermore, there is no 'objectively' fixed transition point – like being of the third or fourth generation of immigrants – between being part of an ethnic minority group and no longer being part of an ethnic minority group. Self-identification shifts the burden of this problem to the people we want to classify. That does not entirely resolve a number of fundamental challenges. People might legitimately want to classify themselves as part of the dominant ethnic group or as part of no ethnic group at all, but still be faced with discrimination (or ethnic disadvantage) if they are judged to be part of a visible minority or negatively racialised group.

When we confront the ethnic self-categorisation questions with the theoretical framework on ethnic groups, we notice that these do not overlap; also, opinions and behaviours of others are crucial in the formation of ethnic groups (Jenkins 1994, 1997). Furthermore, objective approaches, as well as self-categorisation questions, threaten to erase the necessary logical connection between categorisation by others and disadvantageous treatment. A person is not discriminated because he or she is black, but because other people believe this person is black (Sabbagh and Morning 2004, p. 50). In this sense, an important distinction is to be made between self-categorisation (how does one define or see oneself), other-identification (how is one defined or seen by others) and perceived identification (how does one think others define or see you). Therefore, for conceptual as well as for policy reasons, it might be useful to let people classify themselves in the way they believe most other people perceive them. When ethnic group membership is operationalized in this manner, it comes close to the Canadian concept of visible minorities, in which visibility (often in terms of skin colour) has a central place. By going beyond self-categorisation and combining it with other-identification and perceived identification, a balance can be struck: on the one hand, one can give justice to identification choices by individuals, while at the other hand one can embrace the empirical challenge to grasp the prevalence of discriminatory treatment among different analytical groups (independent of self-classification choices by individuals).

In the end, every system of ethnic categorisation holds the risk of essentialism: it reifies ethnic groups. Furthermore, ethnic categorisations reflect (dominant) opinions about who is 'in' and who is 'out', which are embedded in a specific time and place. Scientific classifications, and their statistical formalisation – even if informed by self-classification – are not immune to this. They are equally subordinate to the societal context and power relations as other social products. This can be exemplified by the various manners in which self-classification questions in the US, UK and Canada are posed.

Ethnicity, race, skin colour, cultural background, ancestry and geographical areas or countries are used as operationalisations in various forms and in various combinations, across as well as within countries. Social scientists (and policy-makers) need analytical categories that allow counting and classifying people according to their foreign origin or ethnic background in order to be able to examine their integration into mainstream society. We need reliable data to measure racial discrimination or processes of social exclusion of which visible minorities are

victim. General public support to such registration is rather high: in 2006, 75% of the EU25 population said to be in favour of providing information about their ethnic origin if that could help to combat discrimination in their country (Eurobarometer 2007, p. 169). However, we should be conscious (and remain vigilant) with regard to the performative effects of ethnic categorisations, especially in their statistical form. As Keith puts it

empirical academic studies potentially reify minority presence through ascribed ethnicities that are monitored, counted, and measured in terms of demographic penetration of political systems, employment profiles, and attempts to promote equal opportunities. Such measurement may be pragmatically progressive and politically defensible but inevitably it highlights the 'border problems' of definitions of demographic fixity that reveal the absurdity of racial languages enshrined in politics of affirmative action and census monitoring. (Keith 2005, pp. 258–259)

The classification of ethnic groups in our view, however, constitutes a necessary evil in the construction of an efficient policy aiming at equal opportunities and in the struggle against racism. Furthermore, if we want to promote the quality of international comparative work on the issue, it is essential that classification systems of foreign origin and ethnic background are as similar as possible. For the time, being this is not (often) the case. Researchers should bare this in mind and reflect on the consequences. Policy-makers want to know whether their immigrant integration policies are effective and in a European context increasingly seek to compare the outcomes of their policy choices to those of other national models (for instance, assimilationism vs. multiculturalism). Academics are urged to provide answers on questions like what the best model for immigrant inclusion might look like and some colleagues (Koopmans 2008) are trying to do so even though they have to rely on limited amounts of genuinely comparable data. Immigrant integration policy is high on the political agenda of quite a number of European Union member states. The debate on what the best immigrant inclusion policy might be is at the political centre stage in several countries. The media and policy makers expect that academics working in the field of ethnic and migration studies help and come up with some reliable analysis on which these evaluations can be based. Even in perfect circumstances this is a risky business, but a lack of qualified comparable data to base claims on, makes things even worse.

I cannot propose a 'perfect' system for classification here. I do, however, think that comparative research will profit from the availability of reliable data on country of birth of parents of the population across Europe to be able to investigate recent immigrant groups of first and second generation. Possibly, a two speed Europe could develop, in which those countries that today already have a good registration system agree on some basic definitions so that at least these countries produce comparable data in a short time frame. Later, other countries may adjust their definitions and data collection methods. In the long term, and for those countries which have already long established ethnic minority groups, such a formalized classification system should, in our opinion, be combined with a self-identification procedure. It is not a matter of one or the other system being better. For (comparative) research on immigrant origin groups and ethnic minority groups, both strategies have their

advantages and disadvantages. External classification (with proxies as country of birth of parents) is better suited for statistical comparative work on people linked to recent immigration waves. Self-identification allows to (somewhat) remediate imposition effects and is better equipped to deal with ‘historic’ ethnic minority groups, but is more difficult to organise and more difficult to compare across countries. In qualitative research, taking into account ‘other identification’ (how you think others see you), is a useful addition. We do need to take into account, however, that also these perceptions can be influenced by class or national context.

Once the issue of classification is (temporarily) resolved, we should also reflect on the research design and the consequences of the (lack of a) comparative outlook. What needs to be avoided as much as possible in case selection is ‘selection on the dependent variable’. Let us take the example of a study on homeless people. If we would undertake a biographical analysis among homeless people, an inductive analysis might lead to the conclusion that most homeless have in common a history of drug abuse, relational disruption and professional failure. If, however, we have only interviewed homeless people, it might be misleading to identify these as explanatory factors for homelessness. Quite likely other people who also lost their job and relationship at a particular moment in time and responded with drug abuse as a coping mechanism, did not become homeless. Case selection should be done in assuring there is variance on the independent (explanatory) factors, if we want to put ourselves in the position of doing any inference. Often the mistake is made to select on the dependent variable (see King et al. 1994). Not everyone agrees on such a strong rejection, however, as authors like Alexander George and Andrew Bennett (2005) point out:

Practitioners and analysts of case study methods have argued that selection on the dependent variable should not be rejected out of hand. Selection of cases on the basis of the value of their dependent variables is appropriate for some purposes, but not for others. Cases selected on the dependent variable, including single-case studies, can help identify which variables are not necessary or sufficient conditions for the selected outcome. (George and Bennett 2005, p. 23)

8.5 The Perils and Necessity of Comparison

If one wants to investigate the behaviour, attitudes or participation in a social field of an immigrant group (having determined how to identify them), one needs a comparative group of non-immigrants (or other immigrants) in order to be able to evaluate the specificities of the immigrant group, if that is the aim. Indeed, by focusing the analysis solely on a group of immigrants, one faces the risk of overemphasizing ethnic or cultural characteristics: one should make sure that a characteristic is only present among the immigrant group and thus not shared by the non-immigrant population (and/or other immigrant groups) before grasping at ethnic or cultural explanations.

Empirical research on immigrant issues most often, therefore, needs a control group of non-immigrants in order to investigate characteristics that could be due to the origin of respondents (i.e. ethnic characteristics or the migration experience). The comparison of attitudes, behaviours or participation of an immigrant group with a non-immigrant group enables, indeed, the researcher to evaluate and measure specificities of the immigrant group relatively to the non-immigrant group. In other words, such a comparison allows the researcher to detect attitudes, behaviour or participation of the immigrant group that effectively differ (or not!) from the non-immigrant population. However, the comparison should not be simply done at face value, but should check for the impact of a number of intervening variables.

Let us try and elaborate this by means of an example. A researcher could be interested in evaluating the participation of immigrants in the labour market of the receiving society. If we take the situation of Brussels, the unemployment rate of non-EU immigrants was 35.7% in 1999 (Thys 2000). That clearly is a high figure, but at the same time, this percentage does not tell us whether the non-EU population in Brussels suffer from high unemployment risks compared to other groups. Indeed, the unemployment rate of the non-immigrant Brussels population might be as high as the unemployment rate of non-EU immigrants. Thus, in order to evaluate immigrant participation in the labour market, one needs to compare the unemployment rate of the immigrant group with the unemployment rate of the non-immigrant group.

The proportion of unemployed Belgians in 1999 in Brussels was 13.0%. The comparison between the unemployment rates of both Belgians and non-EU immigrants tells us that people from non-EU countries are almost three times more likely to be unemployed than Belgians. Nevertheless, this picture of immigrant participation in the labour market is still incomplete. Indeed, these unemployment rates of non-EU citizens and Belgians can give us an idea about the average unemployment risks between these two groups. However, a closer look at the unemployment risks of non-EU immigrants in comparison to Belgians according to their educational profiles gives another picture of immigrant participation in the labour market.

Brussels labour market is constituted mainly of high-qualified jobs in the tertiary sector, so that low skilled inhabitants suffer to a larger extent from high unemployment rates. In addition, the education profile of the non-EU population differs largely from the education profile of Belgians: the non-EU population in Brussels is on average lower educated than Belgians. Because Brussels labour market is largely constituted of high skilled jobs and because the average level of education largely differs between non-EU citizens and Belgians, one should also take into account the education profile of both the immigrant population and the non-immigrant population in the evaluation of immigrant participation in Brussels labour market.

Differentiating the participation in the labour market of non-EU immigrants and Belgians along their education profiles gives indeed a more precise idea of the unemployment risks of the immigrant population, all other things being equal. What do the figures tell us then? Among the Brussels population with at most a secondary education degree, non-EU inhabitants are twice more likely than the Belgians to be unemployed. Among the Brussels population possessing a tertiary education degree,

people from a non-EU country are even five times more likely than Belgians to be unemployed (Thys 2000). In other words, it is not that non-EU immigrants, who are often low skilled in Belgium, simply have the bad luck to find themselves in an economy which does not need low skilled labour and that this explains differences in participation rates. As it turns out, participation levels of highly educated immigrants are actually even more problematic. Situating testing according to the procedures established by the International Labour Organisation helps us to understand why: ethnic discrimination is a huge problem on the Belgian labour market (Jacobs et al. 2004).

This example shows that immigrant participation in Brussels labour market varies largely across the education profiles of the population, which, in turn, affects the likelihood of unemployment (in this particular example in a rather unexpected way). Thus, if one wants to evaluate accurately the integration of immigrants in Brussels labour market, it is not sufficient to compare the average participation of the immigrant population with the average participation of the non-immigrant population. One needs also to differentiate these two groups according to their education profiles, since their participation varies largely across education profiles. Having done this exercise, we actually discovered a bit of a surprise in our Belgian data: highly educated immigrant groups are even more the victim of ethnic discrimination than lower educated immigrant groups.

In the remainder of this chapter, we will emphasize the necessity of using a control group in migrant studies. We will, furthermore, try and show that potential intervening factors need to be introduced in the comparison of the immigrant group with the non-immigrant group. One of the conclusions we can draw from our Brussels' example is that, in order to shed light on the specificities of immigrants' attitudes, behaviours or participation, one needs a comparison group of non-immigrants that has the same socio-economic profile (and, for that matter, the same gender and age profile) as the immigrant group. In other words, an immigrant group should be compared to a comparable group of non-immigrants. Indeed, if the research interest lies in investigating immigrant attitudes, behaviour or participation in a social field, focusing on immigrant status, one should make sure that observed differences between an immigrant group and a non-immigrant group are not due to other socio-demographic factors than the ethnic origin. In multivariate statistical models for quantitative research, these matters are relatively 'easy' to verify, in qualitative research we are only able to do so if our initial research design gave us the means to do so. Often, by not following a hypothetico-deductive approach, qualitative researchers fall into the trap of missing this point.

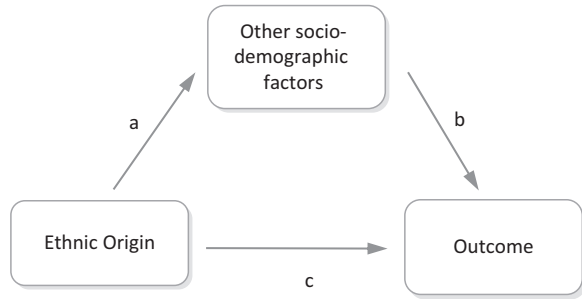
Ethnic disparities observed at first glance might not be due to effective ethnic differences, but to socio-demographic disparities between the two groups. Ideally, thus, research on immigrant integration should strive to compare an immigrant group with a non-immigrant group that has exactly the same socio-demographic profile as the immigrant group. In such a research design, the only socio-demographic characteristic that differentiates the two groups would be the ethnic origin. Thus, such a design would enable the researcher to link without doubt any observed difference between the two groups to the immigrant background of the immigrant

group. For instance, a sample that approximates such a research design would be composed of a group of immigrants and a group of non-immigrant respondents who live in the same neighbourhood as the immigrants (see, for example, Crul and Hering 2008; Modood et al. 1998; Swyngedouw et al. 1999). Indeed, sampling immigrants and non-immigrants living in the same areas would increase the likelihood that the immigrant and the non-immigrant groups have the same socio-economic profile. Admittedly, such a reasoning takes as a point of departure the additive (and distinguishable) nature of explanatory factors. In light of reflections on intersectionality, such an assumption may be untenable.

However, even if such a sample improves the comparability of the two groups, it would not completely ensure that immigrant and non-immigrant respondents have exactly the same socio-demographic profile: it is virtually impossible to construct a subsample of non-immigrants who have exactly the same socio-demographic characteristics as the respondents composing the immigrant subsample. Indeed, migrant surveys with a control group are not experimental studies performed in a laboratory in which the researcher is able to control for every stimulus of the experiment. This is even less the case in qualitative research. Rather, migrant studies aim at investigating respondents' attitudes, behaviours or participation that are shaped by their individual trajectories, experiences and situations; sampling two (or more) groups of respondents in a society that are completely equivalent in their socio-demographic profiles is thus illusory. Therefore, even if a control subsample is drawn by maximizing the socio-demographic similarities with the immigrant group, socio-demographic disparities between the two subsamples are still likely to arise and, consequently, to hamper the measurement of ethnic specificities. Thus, in the (statistical) analysis of integration of immigrants, one should pay attention to other socio-demographic characteristics that might account for ethnic disparities observed in a face-value (bivariate) analysis. One should, therefore, analyse the variation of an outcome between the two groups with a multifactorial perspective. The use of multivariate techniques enables researchers embarking on quantitative analysis to isolate the effect of ethnic origin on an outcome by statistically controlling the effects of potential alternative explanatory variables (Cohen et al. 2003), such as socio-economic variables. Thus, before concluding to significant ethnic differences on an outcome (i.e. significant differences between the immigrant group and the non-immigrant group), one should investigate with a multivariate analysis whether observed ethnic differences are not due to other socio-demographic characteristics. In quantitative studies, this can be included in the statistical model, in qualitative studies all will depend on the researcher taking multiple explanatory factors in consideration while doing interpretations.

This procedure (or way of reasoning) can best be explained with the graph displayed in Fig. 8.1. We give an example from a quantitative approach. The multivariate model is adapted from the mediation model of Baron and Kenny (1986). In a bivariate analysis, the path between ethnic origin and the outcome (i.e. path 'c' on the graph) is the only parameter to be estimated. By contrast, a multivariate analysis allows the researcher to isolate the association of ethnic origin with the outcome by controlling for the effects of other socio-demographic characteristics in the parameter estimation.

Fig. 8.1 Multivariate model for analysing ethnic differences on an outcome. (Adapted from Baron and Kenny 1986, p. 1176)



The situation in which such socio-demographic characteristics may account for the association of ethnic origin with the outcome is depicted on the graph: the association of ethnic origin with the outcome is (partly) explained by other socio-demographic factors that are associated with both ethnic origin (i.e. path ‘a’) and the outcome (i.e. path ‘b’). In our example with the unemployment rate, if one is interested in the relationship of ethnic origin with unemployment risk, one should also introduce the education level of respondents in the analysis, since the education level of Brussels population varies according to the origin of the population and, at the same time, unemployment risks varies as a function of education level. Thus, if the research interest lies in the association of unemployment risks with the ethnic origin of the population, one needs to isolate the effect of ethnic origin on unemployment by holding constant the other socio-demographic variables (such as education) that might affect the association between ethnic origin and unemployment risks. In quantitative approaches, one can statistically model this; in a qualitative approach, one has to deliberately build it in during the case selection procedure and make sure sufficient cases of subgroups are present in the study. On the graph, the ‘other socio-demographic factors’ are mediating the association of ethnic origin with the outcome. Baron and Kenny (1986, p. 1176) stated that ‘in general, a given variable may be said to function as a mediator to the extent that it accounts for the relation between the predictor and the criterion’. In other words, ‘a mediator variable is one that explains how or why another variable affects the outcome’ (Kraemer et al. 2001, p. 852). The ‘other socio-demographic factors’ displayed on the graph are also called confounders in quantitative research (Cohen et al. 2003), because they might confound the relationship between ethnic origin and the outcome. For instance, observed ethnic differences on an outcome in a bivariate analysis might be solely due to a different distribution of socio-demographic characteristics (such as SES) across the origin categories. If some of the socio-demographic characteristics are associated with the outcome, observed ethnic differences on the outcome are (at least partly) spurious. In this case, by introducing socio-demographic factors in a multivariate analysis, ethnic differences on the outcome will not be significant any longer. Thus, if a bivariate analysis results in significant ethnic differences on an outcome, a multivariate analysis might show that the association of ethnic origin with an outcome is spurious and can be totally explained by other socio-demographic factors that function as confounders. The procedure of controlling for possible con-

founders in the investigation of the association of a variable on an outcome is not specific to empirical research on immigrants, but is recommended for any empirical research in social sciences (for instance, see the methodological book of Cohen et al. 2003). However, the failure of following adequately this procedure in migrant surveys can have particularly severe social and political implications. The same holds for qualitative research, where case selection procedures are of crucial importance (in order to avoid selection on the dependent variable).

Think, for instance, of the so-called link between ethnic origin and criminal behaviour. Van San and Leerkes (2001) published the results of a study ordered by the Belgian Justice Minister on the criminal behaviour of immigrant youth. In their analysis of crime data, they showed that the proportion of immigrant youngsters with a criminal record was much larger than the proportion of their Belgian peers. However, their results were based on a bivariate analysis of ethnic origin with data on registered criminal behaviours. They did not control for socio-demographic factors that have shown to be largely associated with criminal behaviours such as age, neighbourhood or socio-economic status. However, the population with an immigrant background in Belgium is characterized by a large proportion of young people and of people with low socio-economic status. Thus, if one wants to analyse the association of ethnic origin with criminal behaviour, one should take into account intervening variables that may confound this association (such as age or SES, without even mentioning different profiling practices by the police). Indeed, when analysing crime data, it will often be the case that, when controlling for socio-economic status and age and police procedures, there is no difference at all in crime rates between immigrants and non-immigrants. Nevertheless, the results of the study of van San and Leerkes (2001) with their shortcomings were largely echoed in the Belgian media with the consequence of strengthening immigrant stereotypes as criminals among the broader population. Thus, while multivariate procedure is recommended in many social sciences fields, it is of particular importance in migrant surveys. Indeed, failing to control for possible confounders in migrant surveys can result in ethnicising or culturalising phenomena that are in reality perhaps not associated with ethnicity or culture. This, in turn, can have severe implications for the academic arena and the broader society. One of the same authors also undertook qualitative research among young Antillian offenders in the Netherlands and their mothers (Van San 1998). One of the outcomes of the research was that mothers seemed to be (too) empathic and understanding towards their sons showing violent behaviour, triggering the author to conclude there was an anthropological dimension to Antillian forms of crime in the Netherlands. This is an interesting insight. Without wanting to totally discard the value of the study, not having interviewed non-Antillian mothers of offenders undertaking the same type of crimes can, however, be argued to be quite problematic. It is very plausible that mothers, irrespective of their ethnic background, but stemming from a similar socio-economic group, find some ways in trying to justify or empathize with the problematic behaviour of their children. When this has not been studied explicitly, we simply do not know. Obviously, this does not mean either that cultural explanations are to be ruled out automatically. However, giving a culturalist explanation, while not having variation on the cultural component in the research design, is a risky endeavour.

8.6 Conclusion

Categorisation is a central feature and challenge for social scientific research. Particularly in the field of migration studies, critical reflection on potential bias triggered by categorisation, case study selection, comparative designs (or an absence of a comparative design) and analysis methods is of essential importance. Given that political stakes and societal fears with regard to the topic of migration issues are very high, precision and clarity on choices and limitations of choices made in categorisation are very important. When dealing with categorisation of migrants, for instance, it is useful to make a distinction between self-identification (how does one define or see oneself), other-identification (how is one defined or seen by others) and perceived identification (how does one think others define or see you). These different types of categorisation do not necessarily overlap and qualitative research should take different combinations serious, as they have differential consequences. Often, there is no easy way out. However, choices need to be made and instead of aiming for a ‘perfect’ solution for categorisation, one should focus on clearly explaining the advantages and disadvantages of one’s choice in categorising. Furthermore, particular attention should be given to the changing nature of categorisations according to the context or particular types of social interactions. Qualitative research is often much better equipped to shed light on this complexity than quantitative research. However, researchers need to be conscious about pitfalls and choices. Qualitative researchers tend to be very strong in deconstructing categorisations and highlighting complexity of identification processes. At the same time, they should remain sufficiently pragmatic and operational when doing deconstruction of socially constructed categories, at the risk of otherwise getting stuck in a swamp of non-decision making and forsaking the task of social scientists to “reduce the complexity of social reality”. Concluding that everything is complex in itself is something we already know. Furthermore, by obscuring or stalling on clear decision making in research designs, researchers might make unwished and uncritical implicit choices, while fleeing from explicit choices out of (legitimate) fear of inherent limitations and performative discursive acts. In this chapter, largely inspired by the approach taken by King et al. (1994), who claim that quantitative and qualitative approaches fundamentally need to tackle very similar methodological challenges, we have provided some examples of pitfalls and choices with regard to categorisation and analysis in the field of migrant studies. I have stressed that one should avoid selecting on the dependent variable. Furthermore, I have argued that the introduction of possible confounders in the analysis is necessary for a correct estimation of ethnic differences on the outcome. If a bivariate effect is estimated in the absence of potential confounders, the analysis may produce spurious relationships between ethnic origin and the outcome (i.e. relationships between ethnic origin and the outcome that are not due to an effective effect of ethnic origin on the outcome, but to other socio-demographic or attitudinal factors). Ideally, a researcher should take into account in his or her analysis all the potential confounders (or mediators) that have shown in the literature to be associated with the outcome. This is perhaps easier done in a quantitative approach (as one can do it post-fact data collection), but

actually should just as much apply in a qualitative approach. In that case, however, the research design has to have foreseen this from the start. In practical terms, this means that an element of comparison is always necessary (be it with non-migrant groups or other migrant groups) in almost any study in the field of migration studies: one should always leave open the possible outcome that, in fact, there is nothing 'special' to the particular group one wishes to focus one's research on.

Acknowledgments The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement 28360, for the EQUOP-project 'Equal opportunities for migrant youth in educational systems with high levels of social and ethnic segregation - assessing the impact of school team resources'. I wish to thank Celine Teney (Universität Bremen), who contributed extensively to an earlier version of this chapter.

References

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.
- Bourdieu, P. (1980). L'identité et la représentation. *Actes de la recherche en sciences sociales*, 35, 63–72.
- Brady, H., & Collier, D. (2004). *Rethinking social inquiry: Diverse tools, shared standards*. Lanham: Rowman & Littlefield Publishers.
- Brubaker, R. (1992). *Citizenship and nationhood in France and Germany*. Cambridge: Cambridge University Press.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah: LEA.
- Collier, D., Hidalgo, F., & Maciuceanu, A. (2006). Essentially contested concepts: Debates and applications. *Journal of Political Ideologies*, 11(3), 211–246.
- Crul, M., & Heering, L. (2008). *The position of the Turkish and Moroccan second generation in Amsterdam and Rotterdam. The TIES study in the Netherlands* (IMISCOE research). Amsterdam: Amsterdam University Press.
- Eurobarometer. (2007). *Special Eurobarometer 263, Discrimination in the EU*. Brussels: European Commission.
- Favell, A. (1998). *Philosophies of integration immigration and the Idea of citizenship in France and Britain*. London: Palgrave Macmillan.
- Gallie, W. (1956). Essentially contested concepts. *Proceedings of the Aristotelian Society*, 56, 167–198.
- George, A., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Boston: MIT Press.
- Giddens, A. (1987). *Social theory and modern sociology*. Cambridge: Polity Press.
- Gray, J. (1977). On the contestability of social and political concepts. *Political Theory*, 5, 331–349.
- Jacobs, D., & Rea, A. (2012). Allochthones in the Netherlands and Belgium. *International Migration*, 50(6), 42–57.
- Jacobs, D., Van der Straeten, T., Brans, M., & Swyngedouw, M. (2004). De impact van het ILO-onderzoek naar discriminatie bij aanwerving op het Vlaams beleid m.b.t. tewerkstelling van allochtonen. Een gevalstudie. In M. Brans, D. Jacobs, et al. (Eds.), *Recherche et politiques publiques: le cas de l'immigration en Belgique – Onderzoek en beleid: de gevalstudie van immigratie in België* (pp. 286–317). Gent: Academia Press.

- Jacobs, D., Swyngedouw, M., Hanquinet, L., Vandezande, V., Andersson, R., Beja Horta, A.-P., Berger, M., Diani, M., Gonzalez-Ferrer, A., Giugni, M., Morariu, M., Pilati, K., & Statham, P. (2009). The challenge of measuring immigrant origin and immigration related ethnicity in Europe. *Journal of International Migration and Integration*, 10(1), 67–88.
- Jenkins, R. (1994). Rethinking ethnicity: Identity, categorization and power. *Ethnic and Racial Studies*, 17(2), 197–223.
- Jenkins, R. (1997). *Rethinking ethnicity: Arguments and explorations*. London: Sage.
- Keith, M. (2005). Racialization and the public spaces of the multicultural city. In K. Murji & J. Solomos (Eds.), *Racialization. Studies in theory and practice* (pp. 249–270). Oxford: Oxford University Press.
- King, G., Keohane, R., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Cambridge: Princeton University Press.
- Koopmans, R. (2008). *Tradeoffs between equality and difference. Immigrant integration, multiculturalism, and the welfare state in cross-national perspective* (Discussion Paper). Berlin: Wissenschaftszentrum Berlin für Sozialforschung.
- Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry*, 158(6), 848–856.
- Modood, T., Berthoud, J., Lakey, J., Nazroo, J., Smith, P., Virdee, S., & Beishon, S. (1998). *Ethnic minorities in Britain: Diversity and disadvantage*. London: Policy Studies Institute.
- Sabbagh, D., & Morning, A. (2004). *Comparative study on the collection of data to measure the extent and impact of discrimination in a selection of countries – Medis Project – Final report on The United States*. Luxembourg: European Commission, Employment and Social Affairs DG.
- Simon, P. (1997). La représentation statistique de l’immigration. Peut-on comptabiliser l’ethnicité ? In J. Rallu, Y. Courbage, & V. Piche (Eds.), *Anciennes et nouvelles minorités* (pp. 1–30). Paris: INED.
- Swyngedouw, M., Phalet, K., & Deschouwer, K. (1999). *Minderheden in Brussel. Sociopolitieke houdingen en gedragingen*. Brussels: VUB press.
- Thys, S. (2000). *Populations issues de l’immigration, marché de l’emploi et discrimination: Situation en Région de Bruxelles-Capitale*. Brussels: Observatoire Bruxellois du Marché du Travail et des Qualifications.
- Van San, M. (1998). *Stelen en steken: Delinquent gedrag van Curaçaose jongens in Nederland*. Amsterdam: Het Spinhuis.
- Van San, M., & Leerkes, A. (2001). *Criminaliteit en criminalisering. Allochtone jongeren in België*. Amsterdam: Amsterdam University Press.
- Waldron, J. (2002). Is the rule of law an essentially contested concept (in Florida)? *Law and Philosophy*, 21(2), 137–164.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

