Experimental Contributions to Sociological Immigration-Research

Abstract

Experiments have not been common in immigration research. In this article I argue that there is a need for more sociological work utilizing that methodology. I report on a review I conducted of experiments on immigration topics, discuss the advantages of those studies, and propose lines for future sociological experimental research in this area. I also describe and counter the often expressed concerns about experiments in the social sciences – namely, those about (1) artificiality of the designs and (2) limited generalizability of the findings. Full references from the review are presented.

Key words: experiments, immigration, artificiality, generalizability

Introduction

Millions of persons around the world are currently either emigrating from their birthplaces or planning/dreaming of doing so. The reasons are many and they include: the prospect of a better future in economic, educational and/or health-related terms; an escape from political and/or religious persecution and warfare; and family reunification. This situation constitutes a major social issue and has generated a variety of policies on the part of the states involved, ranging from a welcoming accommodation to the imposition of harsh barriers by the intended immigration-countries, and from benevolent attitudes to restrictive measures by the potential emigration-lands. For refugees, travelling to their aimed destination often involves many perils, including grave risks at the hands of traffickers.

Not surprisingly, immigration now also attracts the research interest of an increasing number of sociologists. Examples – to name only a few – of the topics commonly investigated in this discipline in relation to immigrants are: employment, wages, acculturation, new-language acquisition, mother-language maintenance, educational accomplishments of their offspring, and residential segregation. These often relate to rejection by the native-born on many fronts on the basis of prejudices
and stereotypes linked to skin colour and country of origin. A wide diversity of theories are employed in the study of these topics, and the evidence comes from a variety of countries. In terms of methodology, most of the sociological research in immigration uses surveys (including census data) and ethnographies, with experiments being underutilized. On this latter point see, for example, McDermott (2006) about political science; McKenzie and Yang (2012) about immigration.

Experiments can be a powerful instrument for both discovering relationships among variables, as well as for testing theories. Here I argue for more work employing that methodology in immigration research, particularly in sociology. The following is the organization of my article I:

(a) highlight the logic and advantages of experiments,
(b) report on a review that I have carried out of published experiments on immigration topics,
(c) show the variety of experimental designs and techniques that can be used, as well as the topics that can be addressed in this area, and propose lines for future sociological experimental work within it,
(d) suggest improvements to a sample of the reviewed studies, and
(e) describe and counter the often expressed concerns about experiments in the social sciences – namely, those about (1) artificiality of the designs and (2) limited generalizability of the findings.

Definitions

An experimental design investigates the effects of one or more independent (antecedent) variables on one or more dependent (consequent) variables. An example would be to study the impact of the extent of similarity among the members of a group on the strength of the affect ties that they develop among themselves. Levels of the independent variable are the experimental (or treatment) conditions (e.g., either high or low similarity in the group members’ extent of formal education, as reported to them by the experimenter); often a control (or baseline) condition consists of either a neutral level or no information in this respect.

Other factors may also be incorporated, as follows. Scope conditions are clauses specifying the circumstances under which a hypothesis is proposed (Berger, Zelditch 1977, p. 25–28; Cohen 1980; Foschi 1997). For instance, if the group is engaged in completing a valuable task, high motivation to do so could be added as a scope condition in the relationship between similarity and affect. Additionally, that relationship may be elaborated by the inclusion of intervening constructs, namely, factors that mediate the link between the independent and the dependent variables. One example of such construct would be the ease with which group members communicate with each other. In turn, depending on the particular study and its theoretical framework, these constructs may or may not be explicitly associated with observables.

Additional components of an experiment include test limitations and theoretically irrelevant factors. The former are particular characteristics of the study that are often mentioned by the researcher because of their potential relevance to the topic under consideration (e.g., the respondents’ average age). They are not part of the
hypothesis tested. The latter are factors that could reasonably have been expected
to be part of the theoretical argument but that the author explicitly lists as immater-
ial (e.g., if the relationship is initially assumed to hold regardless of task difficulty –
not a trivial assumption) and are allowed to vary. It is left to empirical results to
confirm or disconfirm the assumption – the author may either provide results from
previous studies or suggest direction for future research in this respect. For a de-
tailed discussion and additional examples of the various factors in an experimental
design, see Foschi (2014); Jackson and Cox (2013).

In a true experiment participants are assigned at random to different levels of
each independent variable – e.g., to either high or low solidarity among group mem-
bers. These levels are created (manipulated) by the experimenter – for instance, this
is often done through statements that emphasize either the common or the differen-
tiating traits that the participants supposedly have.\footnote{In this text I use the following pairs of terms as synonyms: ‘participant’ and ‘respond-
ent,’ and ‘variable’ and ‘factor.’ When it is clear from the context, sometimes I also use ‘ex-
periments’ and ‘studies’ interchangeably. I do, however, distinguish between ‘experimental
design’ and ‘experimental technique.’ I use ‘design’ to refer to the variables under investi-
gation and to the random assignment of participants to different conditions, and ‘technique’
to describe the particular means through which the latter are created (e.g., the participants
could be asked to perform a visual perception task, or to choose between job applicants, etc.).
Moreover, for simplicity, I often employ ‘researcher,’ ‘hypothesis’ and ‘variable’ in the singu-
lar, although in many cases it would be more accurate to use the plural form. Also for brevity,
when I discuss the design of an experiment, ‘assigned’ means ‘randomly assigned.’ The term
‘issue’ refers to either a topic or subject matter, or to a part of a journal’s volume (as in, e.g.,
Volume 4, Winter Issue). Finally, in line with common use in this literature, ‘immigration top-
ics, matters, etc.’ involves both immigrants and refugees.}

The aim is to control all other factors, so that it can be claimed that the independent
variable(s) is (are) responsible for changes in the dependent variable(s). This can
be achieved in three ways: (1) by keeping other factors either constant or within
a narrow range, (2) by defining some factors as irrelevant (or at least irrelevant for
the time being), and (3) by using random assignment to experimental and control
conditions. These measures enable the researcher to identify with high certainty how
a set of variables are related – an outcome that is the key theoretical advantage of
experimentation (see Webster, Sell 2014; Willer, Walker 2007, Chapter 4).

There are also other types of designs in which participants can be classified
into two or more levels of an antecedent variable. However, if the latter are inherent
to those persons rather than created and randomly assigned by the researcher,
the design is not an experiment. Because they ‘belong’ to the participants, such
factors are often referred to as ‘organismic’ or ‘quasi-experimental’ variables. Sex-
category and ethnic-group membership are two examples, and such attributes are
not truly experimental variables if they refer to characteristics of the respondents
(on this point, see McBurney 1983, p. 139; Magnusson, Marecek 2012, p. 174). On
the other hand, they are such variables if they describe characteristics of others,
and respondents are assigned at random to different levels of those factors – for
example, if the task is to evaluate a task performer who is introduced as either a man
or a woman.
Sometimes the same variable can, in different situations, be both a truly experimental factor and an organismic one. For example, consider studying two groups of individuals, one consisting of immigrants and the other of native-born, with both groups being asked to complete a questionnaire on their own career preferences. ‘Nativity’ is not an experimental factor in that case, as no random assignment is involved. However, this variable could be truly experimental in other situations, such as having participants assigned at random to assess the suitability of two job applicants who are alike in all key social characteristics except nativity. In this article I include only studies in which either nativity or related variables (or both) are truly experimental factors, as I illustrate later in the section Selected experiments – highlights.

Before discussing specific studies, some additional definitions will be helpful. They are as follows.

(1) Experiments are often classified as either ‘laboratory’ or ‘field.’ That classification is often difficult to make, as there are many grey areas in between. The distinction is useful only if one remembers that, regardless of where a study is carried out, it is an experiment if and only if random assignment to conditions takes place. I exclude the so-called ‘natural experiments,’ or events that result in contrasting situations (e.g. a flood of a specified intensity occurring in one region of a country but not in another, even though the two regions have similar physical and environmental features) because they do not meet the condition of control by an experimenter.

(2) A ‘survey experiment’ (also often referred to as ‘experimental survey’) is one in which the main instrument for data collection is a questionnaire in which various, randomly embedded items represent different levels of the independent variable(s). The design is indeed experimental, but in this case the central interest is on making inferences about a sample (preferably representative) from a particular population defined in terms of time and place (see also section On artificiality and generalizability later in this text).

(3) Finally, experiments may be either ‘exploratory’ or ‘hypothesis testing.’ Although both types can be valuable to immigration research, my focus here is on the latter. That is, I am interested in how experiments can contribute to either testing/developing new hypotheses within established theories, or helping in the design of theory-based interventions, or both.

It is also important to note that I am not advocating experiments as the only or even the most important methodology for immigration research. There are many types of experimental situations that either cannot be created or it would be unethical to do so. Consider, for example, a task group consisting of two persons, A (a confederate of the experimenter) and B (a true participant). Level of A’s evaluation of B’s task performance is the independent variable while B’s reaction to the evaluation is the dependent variable. If A’s judgment is one of extremely high praise, the situation would probably become unbelievable to B. If, on the contrary, A’s assessment is so low that it is meant to humiliate B, ethical considerations would indicate that the experiment should not be carried out. My purpose in this article is to point out that, with a degree of good common-sense, respect for the participants, and a creative yet realistic imagination, many sociological immigration-topics can be studied experimentally.
Literature search

I begin by outlining the limits of my present review. I completed an issue-by-issue online check of nine selected English-language academic periodicals for the years 2000 to 2015, and I also carried out, for the same period, an online search using ‘journal article,’ ‘experiments,’ ‘immigrants,’ ‘refugees’ and ‘immigration’ as key words. This time period is particularly appropriate here, given the increased academic interest in immigration matters that has occurred recently. My objective is to investigate empirically the extent to which experiments have been underutilized in this area.²

The choice of journals reflects my assessment of which ones would be likely to contain articles on immigration-related experiments. It also corresponds with my focus on immigration and social factors. Thus, I have excluded from my issue-by-issue search those journals that specialize in either economics or psychology outside social psychology – on the assumption that although experiments are common in them, immigration topics are not. The periodicals I reviewed were: British Journal of Social Psychology; Canadian Journal of Behavioural Science (this journal publishes work in English, as well as in French; for this review however I have considered only those in English), European Journal of Social Psychology; Group Processes & Intergroup Relations; Journal of Experimental Social Psychology; Journal of International Migration and Integration; Political Psychology; Social Justice Research; and Social Psychology Quarterly. My search identified 68 articles comprising a total of 106 studies on immigration matters, all truly experimental designs with respect to at least one independent variable.³ Overall, their data were obtained in 17 countries (in addition, in two cases the location of the research is not specified but the participants are described as ‘British’). Please refer to Appendix B. Inclusion in this bibliography implies that either the independent or the dependent variables of an article concern immigrants themselves (in some studies, ‘immigrant’ is interpreted in a general sense; in others, the term applies to people from particular

² Note that it is now generally the case that academic experiments have to be approved by the ethical committees of the respective institution(s) where the researcher(s) is (are) located. Here I am not referring to any so-called ‘social experiments’ in which immigrants and members of other vulnerable populations would participate, without their knowledge or consent, in experimentally created situations such as housing assignments.

³ Immigrants do, of course, vary overall in such factors as nationality and ethnic background. I have excluded from this review those studies that concern either or both of those factors but do not investigate nativity and/or related variables. Similarly, I have excluded those works in which the persons to be considered by the respondents are described in terms of ‘majority/minority,’ ‘in-group/out-group’ or ‘foreign/national’ rather than ‘native-born/immigrant (or native-born/refugee).’ I also excluded a study if the out-group is likely to be composed mostly of immigrants but the authors do not identify them as such. In addition, my review does not list (a) those studies in which participants themselves are classified as being either native-born or immigrant (that is, those two levels do not constitute an experimental variable) and (b) random assignment to conditions is not explicitly linked to immigration topics. If participants are either native-born or immigrant, or both groups are represented, those characteristics may be considered test limitations.
geographic areas) and/or immigration topics related to them. I have excluded those studies that have physiological measures as the dependent variables.

The total number of issues I inspected in the nine journals for the 2000–2015 years is 682. Of the 68 articles that I identify in the above paragraph, 62 were published in one of those issues – the remaining 6 were from the online search. This represents 9.09% of the works (62/682) – a small figure given the extent of the search. (If one counts experimental studies rather than articles in those same journals, the number increases to 99. In that case, the proportion would have to be calculated in relation to the total number of studies representing various methodologies appearing in those issues. One should also consider that, in several instances, the various studies presented in a single article are replications of the one that appears first.)

Most of the 62 journal articles originate in the psychology-based social psychology literature⁴, but there are also some from sociology, economics, and political science. Not surprisingly, the periodical that contains the most publications on immigration-related experiments (17) was the Journal of Experimental Social Psychology; the two that followed were Group Processes & Intergroup Relations (9) and Political Psychology (8). Although a more extensive search (e.g., including other journals and earlier years) could have been conducted, I estimate from my reading of the immigration-research literature that it would not have yielded a significantly larger number of studies, and that their subject-matters and designs would not have been considerably different.

The majority of the empirical immigration-studies appearing in these journals are surveys (including census data) and ethnographies. This finding is in line with the view I expressed earlier in the Introduction. It is also consistent with McDermott (2014) and McKenzie and Yang (2012) in their respective assessments that experiments on immigration topics have not been generally common in the social sciences. Indications are, however, that their number is currently growing in the published work I have examined. For reviews that reflect the increased interest in this methodology in two social science fields see, e.g., McDermott (2014) on political science, and Thau et al. (2014) on organizational behaviour.

Overall, the works I identified in my review reveal the different types and levels of difficulties that immigrants and refugees often experience in their new lands.

**Selected experiments – highlights**

In this section I examine in some detail a sample of seven (of the 68) articles that I have selected to highlight their diversity in research topics and techniques. The seven contain nine studies (see Appendix A for the full references). I chose these experiments because they represent useful illustrations of several points I wish to

⁴ In some circles, it is common to distinguish between ‘sociological (or structural) social psychology’ and ‘psychological social psychology.’ It is proposed that the emphasis of the former is on the social context while the latter focuses on the individual. But the difference between the two approaches is neither always sharp nor generally agreed-upon. Thus, my view is that it is useful to think of them in terms of shades of differences.
make – my selection does not imply a comparative judgment of the quality of these studies relative to the others identified here. Those other experiments are listed in Appendix B as additional examples.

The following is a brief description of the nine experiments, with a focus on those independent variables for which there was random assignment. Although limited in size, the set is highly informative, as I discuss next in the Overview section of this article.

1. Caprariello, Cuddy, Fiske 2009. Participants read vignettes (brief accounts of situations) depicting an unfamiliar ethnic group said to be immigrating to the US in the near future. Each person was assigned to a scenario describing the group members in terms of their status (either high or low) and competitiveness (either high or low). The dependent variables assessed warmth-competence stereotypes and emotional prejudices. Source of data: Male and female undergraduates from a private university in the US Northeast.

2. Jackson, Esses 2000 (Study 1). Participants were divided at random to read one of two editorials about immigration (one focusing on economic competition; the other, a control condition that described vague, general immigration trends). The former condition resulted in lower levels of empowerment forms of assistance to immigrants. Source of data: Male and female undergraduates at the University of Western Ontario, Canada.

3. Jetten, Wohl 2012 (Studies 1 and 2). In each case, respondents were assigned to read ‘a travel guide webpage’ that showed either high or low continuity (the former was depicted as ‘high homogeneity’) between England’s past and present history; respondents were also measured regarding their identification with England. Those who revealed a strong degree of such identification indicated more opposition to immigration in the low rather than in the high historical-continuity conditions. Source of data: Participants were undergraduates in England, identified by the authors of the article as English and either men or women.

4. Joona, Nekby 2012. New immigrants were assigned to either a treatment condition (intensive counselling and coaching by Public Employment Service caseworkers) or a control group (regular introduction-programs). The results indicate significant treatment-effects on both actual employment rates and participation in intermediate training programs. Source of data: Men and women, newly-arrived immigrants to Sweden.

5. Oreopoulos 2011. Using a technique often referred to as an ‘audit study,’ thousands of sets containing four résumés each were sent in response to online job-postings across multiple occupations. Sets were constructed to plausibly represent, in random combinations, either recent immigrants from Britain, China, India, and Pakistan, or non-immigrants – in each case either with or without ‘ethnic’ sounding names. Levels of other independent variables, also built at random into the sets were: place of undergraduate degree, whether job experience had been in Canada or abroad, and number of languages in which the applicant was fluent. Ceteris paribus, number of call-backs show employer discrimination against
applicants with ‘ethnic’ names and with only foreign-country work experience. 

Source of data: Male and female employment officers from various companies in the Greater Toronto Area, Canada.

6. Ramos et al. (Studies 2 and 3). In the first of these two experiments, respondents were assigned at random to read one of two versions of a brief paragraph about the relationship between Romanian immigrants in France and the French majority population. In one case, the paragraph described both high discrimination against immigrants by the French and perceived high discrimination by Romanian immigrants; in the other case, both actual and perceived group discrimination were presented as low. Respondents’ minority goals (separation acculturation-strategies) were also assessed. Perceived group discrimination affected identification with other Romanians only when minority goals emphasized seeking distance from the majority. In Study 3, the paragraph was about Polish immigrants in Scotland and their relationship with Scottish people. The findings replicated those of Study 2. Sources of data: Romanian immigrants living in France, and Polish immigrants living in Scotland, respectively. Both samples contained men and women.

7. Vezzali et al. 2012a. Children were divided at random to participate in either a three-session intervention that involved imagining meeting an unknown immigrant-peer in various situations, or a control condition without such an imagined meeting. Those taking part in the intervention, compared to the participants in the control group, revealed more positive behavioural intentions/implicit attitudes towards immigrants. A willingness to disclose a major personal problem or secret with an immigrant child mediated the effect on intentions. Source of data: Fifth-grade boys and girls in Italy.

Overview

All nine selected studies show the rich variety of immigration topics that can be and have been investigated experimentally and, overall, reveal a high degree of creativity in the development of experimental situations. These studies also differ in the designs that they use and in their sample populations (e.g., in terms of age, gender, and country of residence). The theoretical backgrounds represented include, among others, stereotype content (Caprariello, Cuddy, Fiske 2009), inter-group competition and social dominance (Jackson, Esses 2000), and imagined inter-group contact (Vezzali et al. 2012a). The studies vary in the extent to which hypotheses are explicitly stated.

Except for the intervention designed by Joona and Nekby (2012) to help through counselling and coaching, the results from this set reveal the various disadvantages associated with being an immigrant. In general, as I indicate earlier, this is also the case of the present larger set of studies.

The truly independent variables were manipulated in several ways. Thus, there were vignettes, as well as manufactured editorials about: immigrant groups differing in their social status and competitiveness (Caprariello, Cuddy, Fiske 2009; Jackson,
Martha Foschi

(2000), degrees of actual and perceived discrimination (Ramos et al. 2013), and the history of the immigration country as either high or low in the continuity of its constituent groups (Jetten, Wohl 2012). In addition, in one study (Oreopoulos 2011), job applicants were presented as either immigrants or non-immigrants and they also varied in country of origin, languages, and place of both undergraduate degree and job experience; in another (Vezzali et al. 2012a), respondents were asked either to envision meeting an unknown immigrant-peer or to participate in a control group. Joona and Nekby (2012) either did or did not provide a group of immigrants with intensive counselling and coaching by public officials.

Ramos et al. (2013, Studies 2 and 3) present an example of an additional, immigration-related quasi-experimental variable measured through respondents’ answers, namely, support for minority goals that advocate a distanced stance vis-à-vis the majority group.

A few studies list clauses such as the following: an immigrant group whose ethnicity was unfamiliar to the respondents (Caprariello, Cuddy, Fiske. 2009), a strong identification with the participants’ own country (Jetten, Wohl 2012, both studies), and an emphasis on distancing oneself from the host nation’ majority (Ramos et al. 2013). Although the authors do not designate these clauses as scope conditions, they are, in fact, such theoretical limitations of their respective hypotheses. Regarding intervening variables, Vezzali et al. (2012a) explicitly incorporate willingness to self-disclose, while Jetten and Wohl (2012) add confidence in the historical group’s future vitality.

Experimental limitations which are worth-noting concern the country of study and characteristics of the respondents. Here again, even the small set of nine experiments shows useful variation: as can be seen from the summaries in the previous section, participants have included undergraduates, elementary-school students, immigrants, and employment officers. The respondents’ countries are Canada, England, France, Italy, Scotland, and the US. The immigrants are described as from Britain, China, India, Pakistan (Oreopoulos 2011) or Poland and Romania (Ramos et al. 2013); in the remaining cases, they are not identified by country of origin. It would be worth investigating whether or not that variation affects the findings. In addition, all nine studies have male and female participants and show results indicating that this factor, although of potential theoretical relevance, has been for the most part found to be irrelevant to the topic under study – an outcome that deserves attention.

The dependent variables consist of either written responses or actual behaviours concerning immigrants. Overall, these variables deal with the degree of opposition to them in various forms, such as lower degrees of support for empowering ways of help (Jackson, Esses 2000), hiring discrimination (Oreopoulos 2011), warmth-competence stereotypes and emotional prejudices (Caprariello, Cuddy, Fiske 2009), and low probability of both employment and participation in training programs (Joona, Nekby 2012).

I conclude this section with a note on ethical considerations. For example, the work by Oreopoulos (2011) uses a technique by which participants do not know that they are being investigated and that they are part of an experiment. In both respects the work raises questions about several ethical points, but in my reading of
the author’s report I could not find a reference to how the latter have been dealt with. If they had indeed been addressed, it would have been useful to include a sentence to the effect that ‘procedures were in line with the university’s research policies relating to human participants.’ This latter recommendation also applies more generally to the set of nine. However, since in all cases the articles have been written by university-affiliated authors, it is likely that ethical requirements have been met. Only in one of the nine experiments (namely, Joona, Nekby 2012) do the researchers explicitly indicate sensitivity to one aspect of the experimental design that they use. Their work was part of a Swedish program that provided immigrants with extra coaching in the preparation of job applications. The authors point out that random assignment helps most – but not all – of the immigrant participants and suggest steps to remedy this situation. (If one considers the entire set of 68 articles, there is a considerable number of authors who indicate that ‘informed consent has been obtained,’ or ‘the participants were debriefed’ or ‘while mean responses of the overall sample were to be made public, they would remain personally anonymous.’ Two examples are Barreto et al. (2003, Studies 1 and 2) and Beaupré (2003, Study 3).

Suggestions for further research

My comments in the present section refer more widely to all the studies identified in this review, rather than only to the selected nine experiments.

Although my search yielded several experiments in a variety of immigration topics of sociological interest, most of the 68 articles (106 studies in total) originate in psychological approaches that contain individual-level variables only. Thus, in Appendix A, the only two works that concern groups are: Joona and Nekby (2012), which considers person-to-person interactions between an immigrant looking for work and a city officer providing intensive coaching about the application procedure, and Vezzali et al. (2012a), which focuses on native-born children, each of whom is involved in an imagined meeting with an unknown immigrant peer. For examples of Appendix B experiments that include groups (that is, each participant is involved in an interaction, either actual, or computer simulated, or imagined/anticipated with at least one specific other person), see Aydin et al. (2014, Study 1); Harwood et al. (2011); Siem, Lotz-Schmitt, Stürmer (2014, Studies 1, 2 and 3); Vezzali et al. (2015). In both Appendices, I have identified all such articles with an asterisk (*). Since they do not involve group interaction, I am excluding from this set those studies in which respondents are presented with vignettes about an immigrant person and asked about their reactions towards him/her (e.g., Stroessner et al. 2015) or ‘editorials,’ ‘news stories,’ ‘web pages,’ or ‘results from surveys’ on immigration issues and asked about the extent to which they agreed with those communications. It would be worthwhile to expand work in this area by adding other designs that either feature group-level variables or, at least, measure effects from the social context (note, e.g., that Caprariello, Cuddy, Fiske 2009 is the only study among the nine that incorporates
the concept of ‘social structure’). A similar point about situational effects is also made in Berry (2001); Esses et al. (2008, p. 16); Tauber and van Zomeren (2013, p. 155–158); Webster and Rashotte (2009). In Foschi (2013), I illustrate how a sociological approach can be followed to investigate experimentally the practice, common in several contemporary societies, of discounting the professional credentials of immigrants.

I propose that a large part of the experimental designs that have been used in sociology and related disciplines to study biases from gender, ethnicity, and in-group/out-group classifications can be readily adapted to immigration topics (see, for example, the valuable review presented by Goar et al. (2013, p. 57–62). There is a vast literature in those areas, including interventions devised to create equal relations among task performers who differ in status. See Cohen (1994) for her significant research program on how to foster equality among different ethnic groups in task settings.

– It would also be important to develop long-terms plans that incorporate both testing the same ideas with different designs, and carrying out systematic replications. I outline both strategies below, as part of my discussion of artificiality and generalizability. (On the advantages of using different methodologies to study a given subject see, for example, Valentino, Brader, Jardina 2013, p. 152; Verkuyten 2005, p. 236–238). As well, it would be beneficial to carry out longitudinal studies – an approach that is still rare in the social sciences. Through such designs, snapshots obtained at particular times could be combined and result in a more complete picture of the topic under consideration.

– Finally, note that the hypotheses investigated in the nine studies originate in different theories and/or approaches, such as social identity, in-group/out-group, social distance, and social dominance. It would be worthwhile to attempt to find areas of convergence across some of the latter (as well as across the entire set) so that the empirical studies are systematic rather than dispersed. (For a useful discussion of theoretical and meta-theoretical integration, see Wagner 2007).

**On artificiality and generalizability**

Experiments have often been criticized in academic publications (including textbooks) for being artificial and not leading to general conclusions. For the most part, these criticisms have been as persistent as they have been uninformed. Some of the points I make next in response to them have been made before by others, and particularly very eloquently by Berger and Zelditch (1977); Cohen (1989, Chapter 6); Webster and Kervin (1971); Webster and Sell (2014); and Zelditch (2014, 1968). My purpose here is both to highlight those arguments and to present some of my own.

(1) I take ‘artificial’ to mean that participants find themselves in an unusual (out-of-the ordinary) setting, and that they are asked to make decisions about matters with which they are not fully familiar. In response to that criticism, one should consider that (a) a setting that is unusual to one person may not be so to another,
and (b) experiments are not uniform in the extent to which the setting is out-of-the-ordinary for the participants. For example, note the range represented by two experiments: at one end, one in which respondents are asked to memorize nonsensical phrases and, at the other end, one in which they are instructed to act as a team in solving a valuable task. Finally, making decisions on unfamiliar topics may in fact be a legitimate design-requirement in a study in which the author wishes to explore precisely the extent to which such a setting permits prejudices to emerge and be recorded.

Unless *either* levels of familiarity with the setting are part of the hypothesis that is being tested or a specific level interferes with the test (e.g., distracts the participants; precludes them from understanding the instructions), this factor is *not* a requirement of an experimental design.

(2) It is also common to read statements such as ‘the experimental results do not generalize.’ Sometimes this is referred to as a ‘lack of external validity.’ The answer to such pronouncements is not as simple as they suggest that it would be, as the matter of ‘generalize to what?’ is seldom addressed, and the meaning of ‘external’ is often vague. It is also of key importance to distinguish between (i) random *assignment* to the conditions of an experiment, and (ii) random *sampling* from a population defined in particular terms (e.g., time and space). If a person were to conduct an experiment to investigate how one independent variable affects a dependent variable *for a particular group of participants only*, it would of course make sense to take a random sample of that population.

My interest, as well as that of many other researchers doing experiments, is not in that situation. Rather, we intend to test theoretical hypotheses. Since the latter are formulated in abstract, not concrete terms, it is not possible to take a random sample from the population to which these hypotheses apply.

All research data are of course particular in one respect or another. The issue is then how data relate to abstract ideas. They do so through links that tie observables to abstract terms. For example, let us define ‘social status’ as a category with two or more levels, each implying different degrees of respect and, in turn, corresponding performance expectations (Berger, Zelditch 1977, p. 34–36). Sex category may then be seen as an instance of social status. Depending on their perceived applicability, status characteristics vary from specific to diffuse; sex category often has a diffuse dimension.

Let us now assume that the results from an experiment have supported a hypothesis proposing that, under specified scope conditions, sex category is a status factor and indeed has affected both assignment of competence and several behavioural, task-related responses such as interpersonal influence. The next question is how to determine the extent of support that the hypothesis has received.

There is no ready-made procedure that could make research findings more general. (Results from a representative sample apply only to the population from which the sample was drawn, unless theoretical work identifies similarities with other populations.) A wider generalization is a task that requires both theoretical thinking and empirical work, and that is achieved through the systematic, often slow, process of replication (see Cohen 1989, Chapters 13–15; Hendrick 1991; Smith 2008). That process involves the following:
(a) A hypothesis becomes more general if findings from well designed and competently conducted studies indicate that its terms can be rephrased at a higher level of abstraction. For example, if the test of a hypothesis about sex-category effects on competence expectations is replicated successfully with results from ethnicity and formal education, one can reformulate the hypothesis to refer more generally to ‘status effects.’

(b) A hypothesis also becomes more general if empirical results show that some of its factors can be expanded: for example, that a scope condition can be relaxed or even deleted, or that one or more of the independent variables can assume a larger range of values.

(c) Generalization can also occur if, through replications, the empirical base indicates that some factors that had been treated as test limitations (e.g., country in which the experiments have been conducted and the level of education of the participants) are in fact irrelevant.

(d) The empirical base for a hypothesis can become more general (and stronger) if its key concepts are operationalized in different ways, either within or across studies, and/or the designs are varied. Four of the works listed in Appendices A and B exemplify some of those various means of generalizing and contribute to cumulativeness: Shinnaoui, Narchal 2010 replicated Esses et al. 2006’s experiment, while Ward and Masgoret 2007’s design and that of Oreopoulos 2011 are comparable in several key respects.

Summary and conclusions

In order to put my enthusiasm for experiments in a larger context, I should mention that, of course, not all such studies are created equal. A poor experiment is still a poor piece of research regardless of the type of design used. When experiments are carefully designed in the context of a theoretical program, operationalize variables successfully, data are presented in an informative way, instructions for replications are provided, the topic is not trivial, ethical guidelines are followed, and the participants are engaged with their task, this methodology is a powerful instrument.

In this article I present highlights from a review of experimental research on immigration topics. Even the relatively small number of nine selected studies from that review serves to illustrate the variety of factors and designs that have been used in this area. I also discuss and address the commonly raised criticisms about experiments in the social sciences, namely concerns about artificiality and generalizability, and make suggestions for further research. My aim is to promote the use of this methodology in the sociology of immigration. In my view, increasing the use of experimental designs in this area will foster theoretical advances and will guide fair social-interventions in immigration matters. Although only a minority of the works identified in the present review deals with groups, there are plenty of ideas in those studies to extend the experiments to more sociological topics.
Acknowledgments

I would like to thank Emma Dewit and Alina Kosel for their expert library-research assistance with the literature review.

References


**Appendix A: Immigration experiments discussed in this text**


**Appendix B. Other immigration-related experiments identified in this search**


Experimental Contributions to Sociological Immigration-Research


**Co wniosły eksperymenty do socjologicznych badań imigracji**

W badaniach nad imigracją eksperymenty nie było powszechną praktyką. W artykule tym przedstawiam argumenty uzasadniające potrzebę wykorzystania tej metodologii w socjologii, relacjonuję wyniki swojego przeglądu prac poświęconych imigracji, w których metoda ta była stosowana, omawiam pozytyki z tych prac.
oraz wskazuję kierunki przyszłych badań eksperymentalnych w tej dziedzinie. Opisuję także i krytykuję często wyrażane obawy o zasadność stosowania eksperymentów w naukach społecznych, mianowicie przeciwsłowiam się: (1) zarzutowi „sztuczności” założonej w samym planie badawczym; (2) przekonaniu o ograniczonej możliwości uogólniania wyników badań. W artykule podana została także pełna lista pozycji bibliograficznych wykorzystanych w przeglądzie.

Słowa kluczowe: eksperymenty, imigracja, sztuczność, uogólnialność