

Experiments in the Sociology of Education – Promises and Experiences

Expériences en sociologie de l'éducation – promesses et expériences

Experimente in der Bildungssoziologie – Verheissungen und Erfahrungen

Rolf Becker*, Joël Berger**, David Glauser***, and Ben Jann****

1 Introduction

Experimental methods are currently attracting a great deal of attention in some areas of analytical-empirical sociology, such as the sociology of education (Becker 2019). Nonetheless, the use of experimental approaches has not excelled as much in sociology as in other disciplines, such as economics or political sciences (Jackson and Cox 2013; Barone and Solga 2020).¹ Apart from in subfields that are rather distant from the sociological mainstream – such as simulation, mathematical modelling or methods and statistics – experiments have been largely neglected in sociology, and this also applies to the sociology of education (Zangger and Becker 2019). Overall, the field is still far from the “experimental turn” that has occurred in other social-science disciplines.

Although experimental designs have not yet become the new “gold standard” (de Souza Leão and Eyal 2019, 383) in the sociology of education, we observe an increase in longitudinal designs that help understand the processes and mechanisms generating inequalities in educational opportunities (Blossfeld 1996; Blossfeld et al. 2019). These designs provide an in-depth understanding of inequalities in educational performances or access to higher education, as well as of the consequences of education policies and administrative interventions in the education system (Blossfeld et al. 2019). For analysing the aforementioned research problems, sociologists can draw on an increasing number of methodological tools (Zangger and

1 In line with this development, David Card, Joshua Angrist, and Guido Imbens – three prominent advocates of methods based on the experimental ideal – received the Nobel Prize in Economics in 2021.

* University of Bern, Department of Sociology of Education, Fabrikstrasse 8, 3012 Bern, rolf.becker@edu.unibe.ch

** University of Bern, Institute of Sociology, Fabrikstrasse 8, 3012 Bern, joel.berger@unibe.ch; Kalaidos University of Applied Sciences, Department of Business and Management, Jungholzstrasse 43, 8050 Zürich

*** University of Bern, Department of Sociology of Education, Fabrikstrasse 8, 3012 Bern, david.glauser@edu.unibe.ch

**** University of Bern, Institute of Sociology, Fabrikstrasse 8, 3012 Bern, ben.jann@unibe.ch

Becker 2019). However, when it comes to assessing causal effects, one approach has proved to be especially robust to the fallacies of (quantitative) empirical research: the experiment (Angrist and Pischke 2009; Murnane and Willett 2011; Jackson and Cox 2013). Apart from the strength offered in identifying causal effects, conducting experiments and orientating sociological research at the experimental ideal has an additional advantage: experiments and questions of causal inference in general force researchers (in the field of sociology of education) to be explicit about mediating social mechanisms that link cause and effect (Hedström and Ylikoski 2010, 55).

To give an illustrative example, the advantage of combining experimental with longitudinal designs has been demonstrated by Heckman (2006). For gathering empirical evidence on the “Productivity Argument for Investing in Young Children”, highlighting the benefits of early investments into children’s education, Heckman and Masterov (2007) refer to socio-political programmes such as the High/Scope Perry Preschool Project. This project was a longitudinal randomized control trial initiated in the middle of the 1960s in the United States. Extremely disadvantaged children of three years of age with a low IQ, as well as parents of low socioeconomic status, were the target population in this field experiment (for an overview: Baldassari and Abascal 2017; Zangger and Becker 2019). Besides short-term effects on the children’s intelligence (Heckman et al. 2008), there were positive medium- to long-term effects of the treatments on cognitive abilities and skills, as well as on non-cognitive skills such as motivation, self-control, and persistence. Several measurements of the children taken at different ages provide evidence that early investment in the education of disadvantaged children results in positive outcomes across their life course, such as educational and occupational success and a low risk of unemployment and delinquency. The results also reveal important mechanisms of cognitive development, externalization of behaviour, and academic motivation (Heckman et al. 2013). These findings have been supported by Schweinhart (2013), who reports causal long-term effects on the intellectual abilities and verbal performances at the age of 50 among treated individuals.

2 Selected Examples of Experiments in the Sociology of Education

There are several other interesting and prominent field experiments in the sociology of education, such as the “Tennessee Student–Teacher Achievement Ratio project” (the STAR project), which investigated the effect of class size on the students’ achievements (Mosteller et al. 1996). This project is also an illustrative example of the challenges that sociological researchers face when conducting experiments in the education system, with the aim of guaranteeing both internal and external validity (Hanushek 1999; Cook 2002). Both the extent to which a study identifies a reliable causal relationship between a treatment and an outcome (internal validity) and

the generalization of the revealed causality (external validity) may be questioned, although the former is typically easier to evaluate than the latter. In the case of the STAR project, for example, several problems of randomization occurred during the fieldwork period, challenging internal validity. Furthermore, it can be questioned if the findings of the experiment conducted in Tennessee could be generalized for other US states or for countries in Europe, Asia, and Africa. The application of vouchers is another popular example of field experiments conducted in educational research (Friedman 2004; Murnane and Willet 2011). Vouchers are often used in programmes that attempt to reduce inequality in educational opportunity and increase the efficiency of public and private schools. The research of these effects is currently dominated by studies from the US that seek to reveal the effect of vouchers and the related choice of schools on the educational opportunities of economically deprived children (e. g. Howell and Peterson 2002; Chakrabarti 2013). In Chile, however, vouchers were already established across the country in 1981 and their effect has been substantially evaluated in the 2000s (Chang-Tai and Urquiola 2006). These studies also document serious problems in regard to internal and external validity caused by selectivity in the treatment (Clampet-Lundquist and Massey 2008).

Overall, such social experiments are of great importance for sociological research on education. Experimental designs open up the prospect of testing the causal effects of programmes, for example the introduction and repeal of tuition fees in some German federal states (Hübner 2012; see also: Bietenbeck et al. 2020). By employing a difference-in-difference approach, Hübner finds a negative effect on the probability of enrolling in university training for freshmen originating from the concerned federal states (Hübner 2012, 954). In contrast to similar analysis by Helbig et al. (2012) or by Quast et al. (2012), the author considers possible “spill-over effects” on the control group, as well as the selective migration of freshmen into other federal states without tuition fees (“substitution bias”), using a structural model. It is found that the revealed effects have to be interpreted as the lower limit of the true effect size within the population (Hübner 2012, 958). The social experiment reported by Dollmann (2011) is another example of the evaluation of educational reforms through an experimental design. Dollmann investigates the question of how relevant binding or non-binding primary school recommendations are for the structure, direction, and amount of inequality at the first transition stage in the German school system. Dollmann (2011) comes to the conclusion that the binding elementary school recommendations are associated with lower inequality because the effects of the parents’ educational aspirations, as well as the parental strategies for realizing those aspirations at this transition in particular, are lower than in the case of non-binding elementary school recommendations. These “natural” experiments provide tremendous potential to identify causal effects of exogenous variations without carrying out an expensive and laborious field experiment. However, for natural experiments it is not always clear whether there is a valid counterfactual setting due

to processes of self-selection into the programmes (Dunning 2008). To sum up, the field experiments mentioned above not only provide mechanism-based explanations of social phenomena; they also go hand in hand with increasing interest among policy makers and practitioners regarding the implementation of findings from field experiments at scale (Raudenbush 2021; see also: Zangger and Becker 2019).

Compared to field experiments, only few laboratory experiments have been conducted in the area of educational research. One of the reasons for this might be that the external validity of such lab experiments, conducted in highly artificial contexts, is unclear (Jackson and Cox 2013; Pietrzyk and Erdmann 2020), while educational researchers are interested in findings that are generalizable (Becker 2019). A prototype of this issue is the laboratory experiment on stereotype threat (Steele 1997). A large number of methodologically flawless laboratory experiments seem to prove that affected groups, as a result of “stereotype threat”, perform less well than their performance potential would suggest (Nguyen and Ryan 2008). However, as far as we are aware, there is no empirical evidence that teachers systematically behave such that a “stereotype threat” is triggered in migrant children in everyday school life and in upcoming exams (Keller 2007). The external validity of the performance-reducing effects of negative stereotypes found in lab experiments is therefore contestable (Wei 2012). Other laboratory experiments seek to reveal discrimination of lower-class children or children with a migration background by their teachers (e.g. Sprietsma 2019; Wenz and Hoening 2020). Due to minor effect sizes and the rather low number of discriminating teachers, these experiments provide evidence that the teachers’ performance evaluations do not contribute systematically to the social origin-related or ethnicity-related inequality of educational opportunity. In sum, the external validity of these lab experiments on the discrimination of minorities is rather limited. To the best of our knowledge, there is no empirical evidence that these findings could be generalized in the sense that teachers show this kind of behaviour in the real classroom setting (Becker and Zangger 2015, 115).

However, laboratory experiments can be useful to test social theory (Willer and Walker 2007). For example, Berger and Combet (2017) provide a laboratory experiment on the status maintenance motive being an important theoretical element in the explanation of class-related disparities in educational attainment (Breen and Goldthorpe 1997). The theoretical background of the status maintenance motive highlights that the aim of families investing in their children’s education is the intergenerational reproduction of class position and the related living standard (Boudon 1974). The higher a family’s socioeconomic status, the more education is needed for their children to achieve at least a similar/comparable socioeconomic status to that of the parents. This theoretical foundation is based on the social position theory of Keller and Zavalloni (1964), as well as on the prospect theory developed by Kahneman and Tversky (1979). By integrating the motive for status maintenance into a prospect theory framework, the individual’s class origin is the

reference point for any educational decision. Against the background of prospect theory, class origin is associated with whether deciders are in a loss frame or in a gain frame. “Higher classes are in a loss frame – they need to pursue longer school careers to avoid a loss of social status relative to their parents. Lower social classes are in a gain frame, as they gain status if they spend longer in schooling than their parents” (Berger and Combet 2017, 126). It is a tenet of prospect theory that individuals are more averse to uncertainty when in a gain frame (lower classes) than when in a loss frame (upper classes). Early tracking increases uncertainty, for example about one’s future educational performance. As such, individuals in a gain frame should be more reluctant to invest in further education under early tracking. In a computerized laboratory experiment, Berger and Combet (2017) find that high-performing individuals in a gain frame are indeed more likely to continue their education when the timing of the decision is late rather than early. They find no corresponding effect for participants in a loss frame. As such, high-performing individuals in a gain frame profit from later educational decisions, which reduces inequality in educational decision making. These results suggest that the structure of the education system – in particular, the length of schooling up to the first important educational transition – is crucial for the deciders’ framing and decision. In this respect, the findings by Berger and Combet (2017) contribute to the explanation of why the class-related educational disparities are much lower in Sweden, where tracking is late (Erikson and Jonsson 1996), compared to education systems with early tracking (e.g. Germany or Switzerland). Educational decisions to be made at a relatively late stage might therefore be useful for families and individuals from lower social classes, who are more averse to uncertainty.

Combining randomized behavioural experiments or vignette studies (Liebe et al. 2018; Möser et al. 2019) with rigorous theory can provide fruitful insights into the interconnection of individual decision making and institutional settings (Auspurg und Hinz 2015). Unfortunately, choice or vignette experiments have also received little attention in empirical educational research (Auspurg and Liebe 2011; Schulze and Schiener 2011). They offer the advantage of a comparatively economical examination of causal relationships. More significant, however, is that choice or vignette experiments are suitable for modelling decision situations. Respondents in this setting have to opt between different alternatives or indicate how they would act in a given situation. The characteristics (attributes) of alternatives or vignettes can be varied by the researcher, and the assignment of choice sets or vignettes to respondents is randomized. These designs makes it possible to uncover, test, and compare different causal mechanisms in a given decision situation. However, a limitation of vignette experiments is that individuals’ intention is measured, rather than their actual behaviour.

3 Some Conclusions

Overall, the increasing number of experimental studies in sociological research on educational inequalities and the evaluation of educational reforms over the recent years have motivated us to devote this focus issue to experiments in the sociology of education. We would like to inform the scientific community about existing experimental research in the sociology of education, as well as about its potential. While laboratory experiments are well suited to test theories, field and survey experiments can be fruitfully applied to replicate findings from the lab (or from quasi-experimental research) and to test and evaluate educational policies (Guala 2005; Falk and Heckman 2009). Of course, there are several methodological and ethical problems in the way of realizing these goals. As argued by Cook (2001; 2002), however, reservations about experiments among sociologists and other educational researchers are flimsy. The advantages and possibilities that randomized experiments offer for gaining knowledge do not seem to be widely known – not only in educational policy, but also in academic educational research (Diekmann 2008).

Nonetheless, there are still challenges in theoretical respects. As it should in principle apply to all empirical studies, the primacy of theory should also be adhered to in experiments in the sociology of education (see also: Murnane and Willett 2011). It might be tempting to conduct experiments in educational research that are merely driven by *ad hoc* assumptions and educational policy guidelines rather than being grounded in sound theoretical argumentation. In addition, the experiment that is methodologically best cannot counter theoretical weakness and lack of justification for hypotheses, so that it may remain difficult to interpret the empirical findings and classify them for scientific progress. Related to this challenge is the need for replication of experiments as a methodological standard in empirical research in order to establish evidence (Jackson and Cox 2013). This is also valid for educational research (Zangger and Becker 2019).

Another particular problem is the so-called “logic of aggregation” (Coleman 1990; Esser 1999). The “aggregation challenge” (Humphreys and Scacco 2020) is based on the following question: how do we transform experimental findings at the micro level to explain a phenomenon at the macro level (such as inequality of educational opportunity)? According to Hedström and Ylikoski (2010, 60), the simple aggregation of individual outcomes, such as social actions, does not provide an explanation of a phenomenon at the macro level (such as inequality of educational opportunity or inequality of returns to education). Sociologists are used to finding transformational mechanisms by taking institutional settings, structural arrangements, or organizational features into account (see also: Esser 1999). In regard to mainstream topics in the sociology of education, the number of vacancies in the educational systems, the institutionalization of tracking, the degree of stratification of the educational systems, or the definition of entitlements for accessing education

at the various levels are considered as structural constraints shaping individuals' choices (Becker 2019). This final remark is intended to stimulate critical rethinking of current theories in analytical-empirical sociology (of education) and to shape the design of studies regarding the generation of reliable knowledge about causal relationships in sociology in general, and in the sociology of education in particular.

4 Brief Overview on the Contributions in this Focus Issue

The benefits, as well as the theoretical and methodological issues, of experiments in the sociology of education are reflected by the contributions in this focus issue. Using a review of 25 published reports on randomized controlled trials in the Italian sociology of education, *Giovanni Abbiati* and co-authors demonstrate that it is necessary to have rigorously defined standards for sociological experiments as a “checklist” that could be used to evaluate the quality of reports and protocols on randomized controlled trials. These standards should contribute to higher quality in sociological and educational research in regard to experimental design, the rigorous outline of the theoretical background of experiments and report of the experimental outcomes.

Knut Petzold discusses the potentials and pitfalls of factorial survey experiments (FSE) in regard to the sociological explanation and direct test of hypotheses. Issues related to the internal, construct, and external validity of FSE are illustrated for the students' intention to study abroad based on two samples of students at a German and at a Chinese university. The author is addressing important methodological considerations in the implementation of FSE, but also discusses them with reference to laboratory, field, and natural experiments.

Using an FSE, *Sandra Gilgen* and *Milan Stocker* investigate discrimination of pupils by their teachers in their secondary school recommendations. In regard to external validity, it has to be stressed that the target population is limited to teachers in Bern, a Swiss canton. The design of the vignettes is theory-driven and includes information on the pupil's gender, academic potential, motivation and behaviour in class, as well as on their social and ethnic origin and their parents' educational aspirations. In line with results from quasi-experimental studies, the degree of discrimination by social class and ethnic origin seems to be negligible. The study does uncover differential treatment of boys and girls, but a limitation in the randomization design makes interpretation difficult. The study thus also illustrates possible pitfalls in the design of FSE.

The study by *Steffen Hillmert* and co-authors is also an application of an FSE, studying the legitimacy of group-specific support for college access. They show that university students have a preference for positive discrimination, compensating for the underrepresentation of the lower social strata and individuals with a migration background at universities. The students participating in the study support modifi-

cations of the performance principle to support socially disadvantaged individuals. However, this study is also an example of problems associated with external validity, since students are not gatekeepers controlling the access of eligible individuals to university training. Yet the results are promising and might well encourage a replication among gatekeepers.

Finally, *Melinda Erdmann* and her co-authors conduct a field experiment to answer the question of whether intensive guidance programmes might contribute to the reduced social inequality in the transition to higher education. Their study combines an experimental design with panel data, evaluating the effect of a counselling programme with the aim of facilitating college access among socially disadvantaged students. There is no positive effect of participation, but some features of the research design might have undermined internal validity. As such, the null effect does not conclusively demonstrate that the programme is ineffective.

5 References

- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Auspurg, Katrin, and Thomas Hinz. 2015. *Factorial Survey Experiments*. Thousand Oaks, CA: Sage.
- Auspurg, Katrin, and Ulf Liebe. 2011. Choice-Experimente und die Messung von Handlungsentscheidungen in der Soziologie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63: 301–314.
- Baldassari, Deliam, and Maria Abascal. 2017. Field Experiments Across the Social Sciences. *Annual Review of Sociology* 43: 41–73.
- Barone, Carlo, and Heike Solga. 2020. Introduction: Experimental Methods in Social Stratification Research. *Research in Social Stratification and Mobility* 65: 1–2.
- Becker, Rolf. 2019. Key Challenges for the Sociology of Education – Theoretical, Methodological, and Empirical Issues. Pp. 12–16 in *Research Handbook on the Sociology of Education*, edited by Rolf Becker. Cheltenham, UK: Edward Elgar Publishing.
- Becker, Rolf, and Christoph Zangger. 2015. Experimentelle Bildungsforschung. Beiträge experimenteller Forschung zur Bildungssoziologie und -ökonomie. Pp. 100–126 in *Experimente in den Sozialwissenschaften* (Soziale Welt – Special Issue 22), edited by Tobias Wolbring, and Marc Keuschnigg. Baden-Baden: Nomos.
- Berger, Joël, and Benita Comber. 2017. Late School Tracking, Less Class Bias in Educational Decision-Making? The Uncertainty Reduction Mechanism and Its Experimental Test. *European Sociological Review* 33(1): 124–136.
- Bietenbeck, Jan, Jan Marcus, and Felix Weinhardt. 2020. Tuition Fees and Educational Attainment. *IZA Discussion Paper No. 13709*. Bonn: IZA.
- Blossfeld, Hans-Peter. 1996. Macro-Sociology, Rational Choice Theory and Time. A Theoretical Perspective on the Empirical Analysis of Social Processes. *European Sociological Review* 12: 181–206.
- Blossfeld, Gwendolin Josephine, Pia Nicoletta Blossfeld, and Hans-Peter Blossfeld. 2019. A Sociological Perspective on Education as a Lifelong Process. Pp. 18–34 in *Research Handbook on the Sociology of Education*, edited by Rolf Becker. Cheltenham, UK: Edward Elgar Publishing.
- Blossfeld, Hans-Peter, Götz Rohwer, and Thorsten Schneider. 2019. *Event History Analysis with Stata*. New York: Routledge.

- Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality*. New York: Wiley.
- Breen, Richard, and John H. Goldthorpe. 1997. Explaining Educational Differentials: Towards a Formal Rational Action Theory. *Rationality and Society* 9(3): 275–305.
- Chakrabarti, Rajashri. 2013. Do Vouchers Lead to Sorting under Random Private School Selection? Evidence from the Milwaukee Voucher Program. *Economics of Education Review* 34: 191–218.
- Chang-Tai, Hsieh, and Miguel Urquiola. 2006. The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program. *Journal of Public Economics* 90(8–9): 1477–1503.
- Clampet-Lundquist, Susan, and Douglas S. Massey. 2008. Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment. *American Journal of Sociology* 114(1): 107–143.
- Coleman, James. 1990. *Foundations of Social Theory*. Harvard, MA: Harvard University Press.
- Cook, Thomas D. 2001. Sciencephobia. Why Education Researchers Reject Randomized Experiments. *Education Next* 1(3): 63–68.
- Cook, Thomas D. 2002. Randomized Experiments in Educational Policy Research: A Critical Examination of the Reasons the Educational Evaluation Community Has Offered for Not Doing Them. *Educational Evaluation and Policy Analysis* 24(3): 175–199.
- de Souza Leão, Luciana, and Gil Eyal. 2019. The Rise of Randomized Controlled Trials (RCTs) in International Development in Historical Perspective. *Theory and Society* 48: 383–418.
- Diekmann, Andreas. 2008. Soziologie und Ökonomie: Der Beitrag experimenteller Wirtschaftsforschung zur Sozialtheorie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60(4): 528–550.
- Dollmann, Jörg. 2011. Verbindliche und unverbindliche Grundschulpflichtungen und soziale Ungleichheiten am ersten Bildungsübergang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63(4): 595–621.
- Dunning, Thad. 2008. Improving Causal Inference Strengths and Limitations of Natural Experiments. *Political Research Quarterly* 61(2): 282–293.
- Erikson, Robert, and Jan O. Jonsson. 1996. Explaining Class Inequality in Education: The Swedish Test Case. Pp. 1–63, in *Can Education Be Equalized?* edited by Robert Erikson, and Jan O. Jonsson. Boulder, CO: Westview Press.
- Esser, Hartmut. 1999. *Soziologie. Spezielle Grundlagen* (Vol. 1: Situationslogik und Handeln). Frankfurt am Main: Campus.
- Falk, Armin, and James J. Heckman. 2009. Lab Experiments are a Major Source of Knowledge in the Social Sciences. *Science* 326(5952): 535–538.
- Friedman, Milton. 2004. *Kapitalismus und Freiheit*. München: Piper.
- Guala, Francesco. 2005. *The Methodology of Experimental Economics*. Cambridge: Cambridge University Press.
- Hanushek, Eric A. 1999. Some Findings from an Independent Investigation of the Tennessee STAR Experiment and From Other Investigations of Class Size Effects. *Educational Evaluation and Policy Analysis* 21(2): 143–163.
- Heckman, James J. 2006. Skill Formation and the Economics of Investing in Disadvantaged Children. *Science* 312(5782): 1900–1902.
- Heckman, James J., and Dimitrij V. Masterov. 2007. The Productivity Argument for Investing in Young Children. Review of Agricultural Economics. *American Agricultural Economics Association* 29: 446–493.
- Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. Earnings Functions and Rates of Return. *Journal of Human Capital* 2(1): 1–31.

- Heckman, James J., Rodrigo Pinto, and Peter Savelyev. 2013. Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review* 103(6): 2052–2086.
- Hedström, Peter, and Petri Ylikoski. 2010. Causal Mechanism in the Social Sciences. *Annual Review of Sociology* 36: 49–67.
- Helbig, Marcel, Tina Baier, and Anna Kroth. 2012. Die Auswirkung von Studiengebühren auf die Studierneigung in Deutschland. Evidenz aus einem natürlichen Experiment auf Basis der HIS-Studienberechtigtenbefragung. *Zeitschrift für Soziologie* 41(3): 227–246.
- Howell, William G., and Paul E. Peterson. 2002. *The Education Gap. Voucher and Urban Schools*. Washington D.C.
- Hübner, Malte. 2012. Do Tuition Fees Affect Enrollment Behavior? Evidence from a “Natural Experiment” in Germany. *Economics of Education Review* 31(6): 949–960.
- Humphreys, Macartan, and Alexandra Scacco. 2020. The Aggregation Challenge. *World Development* 127: 1–3.
- Jackson, Michele, and David R. Cox. 2013. The Principles of Experimental Design and their Application in Sociology. *Annual Review of Sociology* 19: 27–49.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2): 263–291.
- Keller, Johannes. 2007. Stereotype Threat in Classroom Settings: The Interactive Effect of Domain Identification, Task Difficulty and Stereotype Threat on Female Students’ Maths Performance. *British Journal of Educational Psychology* 77(2): 323–338.
- Keller, Suzanne, and Marisa Zavalloni. 1964. Ambition and Social Class: A Respecification. *Social Forces* 43(1): 58–70.
- Liebe, Ulf, Petr Mariel, Heiko Beyer, and Jürgen Meyerhoff. 2018. Uncovering the Nexus Between Attitudes, Preferences, and Behavior in Sociological Applications of Stated Choice Experiments. *Sociological Methods & Research* 50(1): 310–347.
- Möser, Sara, David Glauser, and Rolf Becker. 2019. Valuation of Labour Market Entrance Positions among (future) Apprentices – Results from two Discrete Choice Experiments. *Journal of Choice Modelling* 33: 1–14.
- Mosteller, Frederik, Richard Light, and Jason Sachs. 1996. Sustained Inquiry in Education: Lessons from Skill Grouping and Class Size. *Harvard Educational Review* 66(4): 797–842.
- Murnane, Richard ., and John B. Willett. 2011. *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. New York: Oxford University Press.
- Nguyen, Hannah-Hanh D., and Ann Marie Ryan. 2008. Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence. *Journal of Applied Psychology* 93(6): 1314–1334.
- Pietrzyk, Irena, and Melinda Erdmann. 2020. Investigating the Impact of Interventions on Educational Disparities: Estimating Average Treatment Effects (ATEs) is Not Sufficient. *Research in Social Stratification and Mobility* 65: 1–14.
- Quast, Heiko, Heike Spangenberg, Bettina Hannover, and Edith Braun. 2012. Determinanten der Studierbereitschaft unter besonderer Berücksichtigung von Studiengebühren. *Zeitschrift für Erziehungswissenschaft* 15(2): 305–326.
- Raudenbush, Stephen W. 2021. Scaling Up Experiments to Reduce Educational Inequality. *Behavioural Public Policy* 5(Special Issue 1): 60–70.
- Schulze, Alexander, and Jürgen Schiener. 2011. Lehrerurteile und Bildungsgerechtigkeit: Schullaufbahneempfehlungen angehender Lehrkräfte im Vignettenexperiment. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie* 31(2): 186–200.

- Schweinhart, Lawrence J. 2013. Long-Term Follow-Up of a Preschool Experiment. *Journal of Experimental Criminology* 9: 389–409.
- Sprietsma, Maresa. 2009. Discrimination in Grading? Experimental Evidence from Primary School. ZEW-Discussion Paper No. 09-074. Mannheim: ZEW.
- Steele, Claude M. 1997. A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. *American Psychologist* 52(6): 613–629
- Wei, Thomas E. 2012. Sticks, Stones, Words, and Broken Bones: New Field and Lab Evidence on Stereotype Threat. *Educational Evaluation and Policy Analysis* 34(4): 465–488.
- Wenz, Sebastian E., and Kerstin Hoenig. 2020. Ethnic and Social Class Discrimination in Education: Experimental Evidence from Germany. *Research in Social Stratification and Mobility* 65: 1–13.
- Willer, David, and Henry A. Walker. 2007. *Building Experiments. Testing Social Theory*. Stanford, CA: Stanford University Press.
- Zanger, Christoph, and Rolf Becker. 2019. Experiments in the Sociology of Education: Causal Inference and Estimating Causal Effects in Sociological Research on Education. Pp. 153–171 in *Research Handbook on the Sociology of Education*, edited by Rolf Becker. Cheltenham, UK: Edward Elgar Publishing.