

This is the version of the article that was accepted for publication by the Journal of Comparative Policy Analysis. All rights reserved.

Brans, M. & Pattyn, V. (2017): Validating Methods for Comparing Public Policy: Perspectives from Academics and “Pracademics”. Introduction to the Special Issue, 19(4), 303-312.

## **The second generation of comparative policy analysis**

The launch of the Journal of Comparative Policy Analysis almost twenty years ago marked the emergence of a new generation of explicit comparative policy analytic studies (Geva-May, 2013). The birth of the journal was an indicator of the revitalizing interest in the field, or what promised to be the “renaissance” of comparative policy analysis (Deleon and Resnick-Terry, 1998: 13). Although the history of comparative policy analysis (CPA) goes back decades earlier, with a boom especially in the 1970s and 1980s, CPA was not perceived to be the best remedy to the problems faced at that time. Among other reasons, the recognition of the inherent contextuality of policy problems and of their wider politics was considered to jeopardize the possibility of making reliable comparisons across nations and other jurisdictions. In addition, and related to this, because of the very peculiarities of different macrosocial units (Ragin, 1987, 2014), the ambition to practice policy sciences from one mega-policy umbrella appeared unrealistic. How to aggregate national policies that are marked by extensive heterogeneity? These complexities, combined with the absence of demand outside the academic community for genuine comparative policy studies, resulted in CPA falling largely into disuse in the 1980s (Deleon and Resnick-Terry, 1998).

Importantly though, the decrease in CPA did not imply that the need for comparative policy analysis disappeared. Quite on the contrary, from the early 1990s onwards, the necessity for comparative policy was greater than ever before. With the removal of the ideological barriers between many nations, and the technological advances made in telecommunications, the adage that the ‘world became a global village’ was strongly manifested (Deleon and Resnick-Terry, 1998). In addition, the development of theoretical frameworks such as neoinstitutionalism involved new potential for bridging national differences. Against the backdrop of these developments, the demand for CPA reached a new momentum. Via an explicit comparison over time and/or macrosocial units CPA held the promise to “gain a better understanding of the causes, factors and institutional or actor constellations that bring

about different kinds of policy decisions”. Learning from policy successes and failures in different social units can moreover help to learn how to maximize output, in a certain polity context (Schmitt, 2012:29).

With this so-called second generation of comparative policy analysis, a substantial leap forward was made both at the demand side and at the supply side for CPA. On the supply side, a more global scholarly community engaged in explicit policy comparisons. On the demand side, “clients from both the public and private sectors are asking for high quality comparative research because their futures are relying upon the analyses” (Deleon and Resnick-Terry, 1999: 19).

Two decades after the start of the second generation, it was time to ask where comparative policy analysis stands, in terms of supply and demand, and especially in terms of the methodological match (or mismatch) between the two. Addressing this question is challenging, and complicated by a lack of consensus on the definition of CPA. Not only are there many different methodological approaches to CPA in academic circles (see Schmitt, 2012 for an excellent taxonomy of most common approaches). But also the demand side shows heterogeneity in terms of methodological expectations placed on academic research.

This Special Issue addresses this challenge. The idea for it originated at the Celebratory Comparative Policy Analysis Conference organized at KU Leuven in November 2013 at the occasion of the 15th Anniversary of the Journal of Comparative Policy Analysis and the 10th Anniversary of the International Comparative Policy Analysis Forum. The conference addressed the question of how academics practice and teach methods of comparison, while it also sought to generate an advanced understanding of comparative methods used by national governments and international institutions. The Conference goals were achieved in two days through fifteen panel sessions, two plenary sessions and two key note speeches. Theodore Marmor’s keynote exposed ‘The Unwritten Rules of Comparative Policy Analysis’ and is revisited in Marmor’s commentary in this Special Issue. Michael Keating gave a keynote on practices and challenges of “Policy Learning and Diffusion across Regions in Europe”. He addressed the conceptual confusion of regions as a particular kind of social unit or jurisdiction and laid out the opportunities and pitfalls of policy transfer between them.

## **Towards a methodology of comparative policy analysis**

In the first roundtable of the conference, Frank Fischer of Rutgers and Kassel Universities, Michael Keating of the University of Aberdeen and the University of Edinburgh, Christoph Knill of the Ludwig Maximilian University and Benoît Rihoux of Université Catholique de Louvain discussed the issue of how to move towards a methodology or methodologies of Comparative Policy Analysis. Their discussion started with a consideration of the most important contributions of comparative policy analysis to then appreciate the lessons learned from evaluating different research methods over time. It was also debated whether there are separate worlds of comparative policy analysis or bridges between different methods. The field is naturally diverse, with policies being analysed as material subjects in many guises, as causes and effects, in interplay with different contexts and temporal dimensions. There is agreement that overarching theory is absent and hesitation on whether it is even possible to develop such theory. Grand theory may even be undesirable, as the grander a theory gets, the less it might be able to explain, because of the high level of abstraction of causal mechanisms. Most advances have taken place at the middle range level, and theoretical progress has been aided by heuristic frameworks, such as the stages and substages models, by the development of policy typologies and taxonomies (see e.g. Steinberger, 1980), and by conceptual sophistication on policy change and policy content. Even if strictly empirically tested theories are few, some degree of indeterminacy of concepts has been countered by the rich body of knowledge and efforts to define the most basic of dependent variables in comparative policy analysis, both substantively and procedural. Yet, the participants to the debate agreed there is more ground to cover by developing meaningful concepts that can travel across different contexts and time, and by further building typologies that reveal the features that cases share or do not share. One may also expect progress in the direction of theory development if the scholarly communities in whose camps analytical frameworks have advanced the most, spoke to each other more and developed communities around particular problem-based research puzzles and questions.

The second question of the debate focused on methodological advances and challenges. There is broad agreement that the field of comparative policy analysis has witnessed great progress in the sophistication and specialization of methods, both in quantitative and qualitative policy analysis. Qualitative analysis in particular increasingly uses transparent and replicable methods, ranging from systematic reviews, content analysis, to ethnographic research.

Advances in qualitative comparative analysis deserve especially to be noted, even if the language of the method is still considered rather cryptic by some, including Marmor in this Issue. Qualitative Comparative Analysis (QCA) is a means of analysing the causal contribution of different conditions to an outcome of interest. QCA starts with the documentation of the different configurations of conditions associated with each case of an observed outcome. These are then subjected to a minimisation procedure to identify the simplest set of conditions that can account for all the observed outcomes, as well as their absence. According to Rihoux (see also Rihoux et al., 2011; and Hudson and Kühner, 2013), there are at least five reasons why QCA is well matched with comparative policy analysis. Some of these are developed at length in Fischer and Maggetti's contribution to this issue. First, QCA is suited for typical comparative policy analytical designs of a cross-national, cross-regional and cross-sectoral nature. Second, it is useful for intermediate *n*- designs where researchers ideally meet the conflicting demands of case knowledge, interpretation, and some level of generalization. Third, QCA enables policy analysts to test alternative policy intervention models leading to contrasted policy outputs or outcomes and can help to identify multiple paths leading to the same outcome. Fourth, since QCA does not present a "push-button methodology" and involves the researcher as an active player, QCA also lends itself to interpretative comparative analysis. And fifth, with QCA one can perform many types of meta-analyses of existing multiple case studies, monographs and reports, which also makes it an attractive methodology to national and international government institutions. At the very least, QCA can be used for typology building, mapping and synthesis.

Next to advances in qualitative methods, the panel also lauded the move away from single method designs as a positive development for the validity of research findings. We witness this development in the combination of experimental designs with case studies; or QCA with process tracing. The conference also heard pleas from quantitative analysts to combine quantitative survey methodology with qualitative case studies in order to reveal causal mechanisms, as well as to increase the validity of statistical patterns. Jacques Billiet, one of the leading authorities on measurement errors and validity of cross-national research, not only asked for better qualitatively grounded concepts to start with, but also for qualitative case studies to complement the results of statistical analyses (Davidov et al. 2014).

The validity of inferences may thus be increased by using multi-method strategies, combining statistical analysis with intensive case study analysis to test the real world value of statistical

models. Also, Frank Vandenbroucke, another participant at the ICPA-JCPA Leuven conference, who has been engaged in quantitative social welfare research for years, called for combining survey analysis and panel data with qualitative designs, to really understand the conditions under which some policy instruments work or not in combatting poverty. This call is repeated in his contribution to this issue. These are of course routes where qualitative research follows quantitative analysis, but the sequence can also be reversed. There is still much ground to cover through descriptive studies, which can help codify policy problems, solutions and outcomes, the results of which can later be used for constructing classifications of cases and typologies. As Christophe Knill said in the plenary debate, typologies have the advantage of not going out of fashion too quickly and are great for inspiring case selection for further comparative research, or for helping to design surveys for quantitative research. There is one important caveat though. As Michael Keating put it, we have to be wary about hyper-conceptualisation, and not lose “sight of analysis for policy, in the way we like to do analysis of policy.” This warning brings us to the subject of the second plenary debate, that of matches and mismatches of comparative policy analysis in academia on the one hand, and in governments and international institutions on the other.

### **Academics and government in dialogue**

In this second roundtable discussion another group of distinguished speakers, namely Geert Bouckaert of the KU Leuven Public Governance Institute and President of the International Institute of Administrative Sciences, Marco Cangiano, then Assistant Director of the IMF Fiscal Affairs Department, Alessandro Colombo of Eupolis Lombardia-Institute for Research, Statistics & Training, and Frank Vandenbroucke of the University of Amsterdam discussed among them the match or mismatch between comparative policy analysis and policy making, problems of validity and reliability of governments engagements in comparisons, and the ways forward to matching supply and demand with the aim of improving public policy in practical terms. Applying a supply and demand approach, Geert Bouckaert spoke of several “frustration zones”, where policy analytical supply lacks demand from policy-makers, or where the demand from government is unmatched by academic supply. Academics’ work may not be considered relevant in the eyes of practitioners, because it is not problem-based, too remote from the nuts and bolts of policy-making, too locked in paradigmatic camps, or too prospective to fit short term policy agendas. Matchmaking can proceed by science communication, research brokerage, and mutual engagements in formulating and re-

formulating research questions in official calls for research. Also worth mentioning are policy seminars, such as those organized by the OECD with individual experts and research networks.

A particular mismatch was mentioned by Frank Vandenbroucke on reflection of the statistical infrastructure in support of the European Union's method of management by objectives and mutual learning. While academic publications on the policy process of this open coordination have piled up over the years, little academic interest was shown in researching the quality of the statistical systems itself. The vast amounts of data are underutilized in terms of validity and quality of datasets, and to the extent that existing surveys and datasets contradict each other in terms of understanding the size and causes of poverty, the policy implications of this lack of attention are rather alarming. Research agendas had better focus on data-collection, data-analysis and cross-linkages between datasets, but this is complicated in an era where much research and publishing is driven by the ready use of available datasets without questioning the meaning and validity of data, and where research foundations find the mere collection of reliable comparative data inferior to the sophistication of theoretical frameworks. Also, in the context of performance budgeting and the use of indicators by international institutions such as the IMF and Worldbank, Marco Cangiano sees too few incentives for keeping databases alive, and to work on metadata. What is lacking is a kind of rigorous validation of the information process, and government and international institutions rely too much on self-assessment and peer evaluation. Here governments and academia are challenged to join forces in creating independent assessments of statistical data and indicators.

### **Contributions to the Special Issue**

The conference panel sessions of papers were organized around methods (Qualitative Methods, Quantitative Methods, QCA, Mixed Methods and Small N and Case Design), and around the boundaries of academia and practice (Policy learning across Boundaries, International Comparators, Policy Dialogue and Diffusion, Welfare and Poverty Indicators, Predictive Methods and Governance Mapping). They promoted dialogue between academics and practitioners and the discussion of advances in comparative research design, policy analysis, construction and validation of comparative knowledge of manipulable policy, program or institutional variables. Four of the papers presented in the panels appear in this Special Issue (Adam, Hurka and Knill; Fischer and Magetti; Hoffmann and Van Dooren;

Cangiano), complemented by one revised plenary contribution (Vandenbroucke), and one commentary derived from one of the conference's keynotes (Marmor).

The articles in this issue address the question of how academics practice methods of comparison while it also seeks to contribute to an understanding of the methods of comparison used by national governments and international institutions. Although only a few of the articles presented in the issue explicitly address the supply-demand nexus, all can be read as a contribution to a better fit between supply of comparative research by academics on the one hand, and demand for information on comparative practices by international organizations or governments on the other hand. The issue combines articles that are representative of the two worlds of academia and policy makers. Half of the contributions deal with academic practices of comparison. The other half focuses on practices of comparison in national governments and international institutions. In addition, the Special Issue combines the points of view of academics, and of persons who belong to the category of so-called "pracademics". The latter denominator is often used to refer to people who have careers rooted in both academic and public service camps (Posner, 2009: 14). More than most, pracademics are best positioned to identify the causes of possible mismatches between the "two communities" (Caplan, 1979), and to provide methodological suggestions to close this gap.

As for the academic practices of comparison, the Special Issue brings together two articles, one on conceptual development and one on qualitative comparative analysis.

The contribution of *Adam, Hurka and Knill* addresses one particular prerequisite for engaging in comparison: the development of meaningful descriptive concepts capturing the essence of empirical phenomena. While they note that much progress has been made on the conceptual development of policy content and process, they attempt to fill a gap on defining patterns of regulatory outputs. Regulatory outputs are ultimately about balancing individual rights with collective intervention, the variations of which have traditionally been understood as manifestations of varying traditions of state-society traditions. Instead of focusing on politics and policy process dimensions that informed the dominant research tradition, the authors specify regulatory outputs in terms of constraints on individual behavior and of sanctions for deviations. They identify four ideal-typical styles of regulatory outputs: authority, lenient authority, permissiveness and punitive permissiveness. After clarifying their method for measuring the different dimensions, the authors use their typology to compare handgun

regulation in 12 European countries. In addition, examples of regulatory outputs controlling pollution, drug abuse and sports gambling illustrate how the typology can fruitfully inform empirical research in other regulatory domains.

The article by *Fischer and Maggetti*, for its part, critically examines the potential of QCA for the study of policy making processes. QCA, as an approach and an array of techniques, has become increasingly widespread. The authors investigate whether the analytical features inherent to QCA are adequate to capture variations in policy making processes, which are inherently complex and dynamic over time. They argue that the study of multiple configurational causal relations helps researchers to deal with the extremely complex set of factors that interact in the policy process. While QCA's designs can carefully balance inductive and deductive logics, and are well suited to deal with the complex causation of policy processes, QCA is less apt to deal with problems of temporality, which feature recurrently in policy process theories. In their article, the authors offer a number of solutions to reconcile the dynamic nature of policy processes as moving targets that evolve over time, with the static character of QCA frameworks that allow mostly for synchronic comparisons. Their methodological solutions are relevant for influential theories and models of the policy process to better capture the timing of policies, policy punctuations and the identification of patterns over time.

As for the practitioner approaches to comparison, conducted by national governments and international institutions, the Special Issue includes three diverse but complementary contributions. They all deal with the challenges of cross-country comparisons and the development of suitable indicators and measurements.

Based on the examination of the case of sick leave statistics, *Hoffmann and Van Dooren* identify various concrete pitfalls of international comparative measurement, both as definitional issues and measurement issues. The authors present several coping strategies to overcome these difficulties. Their article is anchored in the debate on international comparative governance indicators. Using examples from different EU countries and regions, the authors discuss one specific type of indicator, sick leave, and address the broader implications for international performance management, linking it with the work of such organisations as the OECD and Eurostat. The authors' critical analysis points at multiple problems of validity and reliability in international governance comparisons. In addition to

identifying conceptual deficiencies, they criticize the opacity of indicators, which are often arrived at through shortcuts via perceptual data and small size sampling. Measurement validity is compromised by lack of measurement and data transparency. A particularly thorny issue is the absence of disaggregated data. Yet, despite criticisms, the analysis ends on a positive note, speaking to both worlds of practice and academia. The authors expect progress in reducing problems of reliability and validity, provided international organizations find general acceptance of definitions of indicators, and national and international statistical offices open up their datasets. Researchers in turn are reminded to critically evaluate available datasets before using them, and to look for ways to combine quantitative measurement with qualitative analysis of context.

*Marco Cangiano*, in his article ‘On the Limits of Rating and Ranking When Comparing Fiscal Institutions,’ revisits the debate on the validity of using metrics in international comparisons. The article is a warning against excessive reliance on the outcomes of indices in designing reform agendas. Comparisons of governments’ performance is fraught with definitional and methodological problems. They are often based on synthetic indicators that confuse normative aspects (what a government should do) with efficiency/effectiveness arguments (how governments perform vis-à-vis stated policy objectives). Indicators tend to be based on so-called best practices; measure what can be measured rather than measuring what is relevant; and synthesize complex institutional aspects in numerical ratings without due consideration to contextual factors. Finally, as also pointed out by Hoffman and Van Dooren, most indicators rely on perception based measures of phenomena that are often not directly observable. The problems identified by Cangiano are particularly relevant in the wake of the financial crisis when comparing fiscal institutions across countries. Central in determining governments’ capacity to deliver efficiently and effectively goods and services while maintaining macro-economic stability, comparative public policy analysis requires a sounder theoretical framework than the one on offer by existing synthetic indicators. Cangiano illustrates how two indices of fiscal rules developed by the European Commission (EC) and the International Monetary Fund (IMF) are prone to a normative or best practice bias, thus ignoring essential contextual factors. He makes a case for measuring what is relevant and for not shying away from descriptive analysis. “Comparing governments’ relative performance is too important and complex to be left to sound bites, texting, or a single numerical ranking”.

The article by *Vandenbroucke* presents the case of comparing European social policy. As a unique kind of polity, the EU developed particular governance tools such as the Open Method of Coordination (OMC), a method which seeks to achieve common objectives through peer learning and comparison among member states in Europe. The implementation of OMC implies a particular role for CPA. The article addresses several challenges of a political, scientific and technical nature that complicate CPA, and that hence also impede successful OMC and the assessment of the comparative efficiency of EU welfare states in general. Each of these challenges is richly illustrated with examples that stem from the author's previous background as a Belgian Minister of Labour and of Education, and as one of the founders of OMC. Vandenbroucke's contribution is a strong call for an investment in the development of reliable and comparable data, as a prerequisite for mature comparative social policy analysis. Three of his points stand out. First, the EU has much improved its statistical capacity, but is still wanting in analytical capacity, particularly in linking administrative data with outcomes, or inputs with outcomes. Second, cross-national comparisons in the EU do not produce valid results for policy-making if not combined with cross-time comparisons. Countries that seem leaders in combatting poverty are doing rather badly from a cross-time perspective. Third, statistical analyses are useful, but need to be supplemented by qualitative case research to uncover the causal mechanisms of the effectiveness of policy instruments. Statistical analyses alone do not give directions to designing policy instruments.

All the above articles are preceded by an introductory essay by *Theodore Marmor* on 'Comparative Studies and the Drawing of Policy Lessons'. In this introduction, Marmor first revisits the rules of the game of international comparisons, which he presented in his keynote at the ICPA-JCPA Leuven Conference. Second, from his long experience in researching health care in the US, Canada and Europe, he identifies some of the deepest sources of conceptual confusion in cross-national health policy analysis. Third, his essay presents a commentary on the other five articles in this issue and an appreciation of the question of cumulative advance in comparative policy analysis. Of the five rules of international comparisons Marmor lays out, three are guiding and two are cautionary. The three guiding rules highlight that cross-national comparison should serve the purposes of illumination, of defense against 'explanatory provincialism', and of increasing the validity of inferences from treating cross-national experience as quasi-experiments. The two cautionary rules warn cross-national inquiry against the risks of "naïve transplantation" and "intellectual nihilism", the latter springing from the "fallacy of comparative difference". Following these rules implies

that policy analysts should clarify the very purposes of their comparative policy work, reflect deeply on the substance of the policy sector under study, and invest in operational definitions. And before qualifying the success or failure of policy reforms, they should first seek a thorough understanding of the dominant values, institutions, and organized interests in the studies policy domain. Marmor shows the conceptual confusion of the vocabulary in cross-national health care studies, which in his view often leads to “cycles of naïve enthusiasm and regrettable neglect.” He richly illustrates muddled language and insufficient appreciation of political context, as well as the risks of unwarranted generalisations and misleading policy recommendations. In the final part of his introduction, Marmor critically assesses the contributions to this issue. In his evaluation, he rather skeptically notes a lack of cumulative advance. He does not see the contributions converging towards a consistent comparative method, but rather as a representation of the state of CPA by various separate intellectual communities. More optimistically, he sees this variety as a possible source of eventual analytic improvement.

## **Conclusion**

The papers in this issue show more agreement than disagreement on how comparative policy analysis, both in academia and government, should proceed to increase the validity of cross-national findings. The increased specialization in the last 20 years has greatly contributed to greater sophistication and formalization of methods, and has engendered many conceptual frameworks to guide the comparative study of policy processes and manipulable policy variables. Yet with specialization has come data-driven and method-driven research, and a tendency for partisans of particular approaches to stay within the comfort of their silos. The prevention of data, method, and theoretical parochialism calls for academic debate and dialogue, the essence of any science. The Special Issue contains many suggestions for advancing comparative policy analysis. The validity of comparative policy analysis will be enhanced by having members of the scholarly communities join approaches in addressing problem based research puzzles, combining their methods in designs that do justice to complexity and context. Other processes of validation can take place in dialogues between scholars and practice. Scholars need sufficient access to problematize and help improve the scientific validity of institutionally produced data. At the same time, scholarly interactions with policy-makers can contribute to another kind of validation, call it real-world or practical validation, that often lies beyond the scope of research or research funders. Comparative policy analysis should not fall into the trap of building overly abstract models and testing

them with hyper-sophisticated methods, to the extent that actual policy-makers no longer recognize what the discipline is on about: to paraphrase Michael Keating's (and Lasswell's (1971) for that matter) expression: scholars should not lose sight of comparative analysis *for* policy in their way of doing comparative analysis *of* policy.

## References

'Towards a methodology of comparative policy analysis'. Roundtable Transcript. ICPA/JCPA Celebratory Comparative Policy Analysis Conference: Validating Methods for Comparing Public Policy: Academia and Government in Dialogue. KU Leuven Public Governance Institute, 28 November 2013.

'What do institutions want from comparative research and what does comparative research want from institutions? / Towards a better match between comparative Policy Analysis and Policy-Making?' Roundtable Transcript. ICPA/JCPA Celebratory Comparative Policy Analysis Conference: Validating Methods for Comparing Public Policy: Academia and Government in Dialogue. KU Leuven Public Governance Institute, 29 November 2013.

Caplan, N. (1979). The two-communities theory and knowledge utilization. *The American Behavioral Scientist*, 22(3), 459.

Davidov et al. (2014) E., Meuleman B., Cieciuch J., Schmidt P., Billiet J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology*, 40, art.nr. 10.1146/annurev-soc-071913-043137, 55-75.

Deleon, P., & Resnick-Terry, P. (1998). Comparative policy analysis: Déjà vu all over again?. *Journal of Comparative Policy Analysis: Research and Practice*, 1(1), 9-22.

Geva-May, I. (2013). Foreword. Conference Book. Celebratory Comparative Policy Analysis Conference. Validating methods for comparing public policy: Academia and Government in Dialogue. KU Leuven. November 2013.

Hudson, J. R., & Kühner, S. (2013). Innovative methods for policy analysis: QCA and fuzzy sets. *Policy and Society*, 32(4), 279-356

Lasswell, H.D. (1971). *A Preview of Policy Sciences*. New York, American Elsevier.

Posner, P. L. (2009). The Pracademic: An Agenda for Re-Engaging Practitioners and Academics. *Public Budgeting & Finance*, 29(1), 12-26.

Ragin, C. C. (1987, 2014) *The Comparative Method. Moving beyond Qualitative and Quantitative Methods*. Oakland, University of California Press.

Rihoux, B., I. Rezsöházy and D. Bol (2011). Qualitative Comparative Analysis (QCA) in Public Policy Analysis: an Extensive Review. *German Policy Studies* 7(3), 9–82.

Schmitt, S. (2012). Comparative approaches to the study of public policymaking. Chapter 3. Routledge Handbook of Public Policy Routledge. Accessed on: 05 Jul 2016. <https://www.routledgehandbooks.com/doi/10.4324/9780203097571.ch3>

Steinberger, P. (1980) Typologies of public policy: Meaning construction and the policy process. *Social Science Quarterly*, 61(2),185–197