

THREE WORLDS' TYPOLOGY: MOVING BEYOND NORMAL SCIENCE?

Accepted for publication in Journal of European Social Policy, 2015, 25(1).

Kees van Kersbergen (corresponding author)

Aarhus University
Dept. of Political Science and Government
School of Business and Social Sciences
Building 1331
Bartholins Allé 7
DK-8000 Aarhus C
Denmark
Email: kvk@ps.au.dk
Web: ps.au.dk
Tel: +45 87165727

& Barbara Vis

VU University Amsterdam
Dept. of Political Science and Public Administration
de Boelelaan 1081
1081 HV Amsterdam
The Netherlands
E: b.vis@vu.nl
Web: www.barbaravis.nl / www.highriskpolitics.org
Tel: +31 20 598 8644

Abstract

There is no doubt that Esping-Andersen's three worlds' typology has been extremely valuable. However, the literature inspired by it shows signs of Kuhnian *normal science*, which is impairing empirical and theoretical progress. We explain normal science, demonstrate that it characterizes recent empirical regime studies and ask why this has come about. We show that the welfare regime literature has a tendency to confuse the terms 'typology' and 'ideal-types'. This has prevented the emergence of anomalies that are needed for progress. We argue that normal science is fostered by the combination of researchers' tendency to prefer certainty (e.g. to address solvable research problems) and the environmental pressures they face (especially the 'publish or perish' culture and the need to frame research problems in terms of variation rather than similarity). In the discussion section, we suggest several routes by which the welfare state literature can move beyond normal science.

Key words:

Three worlds, typology, welfare regimes, normal science, ideal types

1. Introduction

The prologue to this issue (Emmenegger et al. 2015 [this issue]) has made clear that it is difficult to overstate the importance of Gøsta Esping-Andersen's landmark study *The Three Worlds of Welfare Capitalism* (1990). We would like to reiterate that two innovations were particularly powerful. First, the concept of welfare *regimes* allowed a broader and better understanding of the different ways in which society's major institutions (state, market and family) interacted to produce specific patterns of work and welfare. This both helped to remove the field's theoretically unsatisfying preoccupation with aggregate social spending as the indicator of welfare state generosity and opened up a whole new area for systematic comparative research. Second, *Three Worlds* proved that the quality of welfare regimes could also productively serve as an independent variable. Regime differences systematically relate to differences in social outcomes that matter, particularly regarding macro-level indicators such as employment structure, labour market behaviour, and income inequality (see Kammer et al. 2012). Since then, much intellectual effort has been invested in further developing, testing, adjusting, criticizing and applying Esping-Andersen's regime typology,¹ which has become a *paradigmatic* one (cf. Arts and Gelissen 2010: 569). Based on a review of empirical studies that use Esping-Andersen's typology, Arts and Gelissen conclude that '– in spite of all kinds of conceptual, operationalization, and data problems that must be solved – (...) his typology is promising enough for work to continue on welfare state models' (581). Similarly, Ferragina and Seeleib-Kaiser's (2011: 598) review confirms the existence of three regimes and concludes that the typology remains 'a fundamental heuristic tool for welfare state scholars' (see also Ferragina et al. 2013).

Generally speaking, we obviously agree with the *communis opinio* about the relevance of the three worlds' typology. We also agree with Arts and Gelissen's (2010) suggestion that progress in this field will very much hinge on the ability to devise a firmer theoretical foundation of contemporary and future welfare state models (understood not as normative ideals but as explanatory categories) (see also Powell and Barrientos 2011). But is the existing literature on this trajectory? We propose that, so far, for most – although not all – existing studies within the field of

welfare state regime research on which we exclusively focus,² this is not the case. Instead, we argue that at least partly because of the typology's success, the recent welfare regime literature that takes *Three Worlds* as its starting point or key point of reference, has started to show signs of what Thomas S. Kuhn (1970) called *normal science*. Normal science is 'research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its current practice' (Kuhn 1970: 10). It 'does not aim at novelties of fact or theory and, when successful, finds none' (idem: 52). Instead, normal science is typically puzzle-solving: it focuses on puzzles for which an assured solution exists (idem: 37). While a scholar working within normal science may 'try out a number of alternative approaches, rejecting those that fail to yield the desired result, he is not testing the *paradigm* when he does so' (idem: 144, emphasis in original). Scholars solve solvable puzzles, possibly with increasingly sophisticated approaches that most likely produce novel empirical insights (that strengthen the paradigm). In the process, however, they may lose sight of theoretical substance. This would mean that the three worlds' typology is at risk of becoming an, what me might call, unproductive intellectual straightjacket standing in the way of, rather than being helpful for, asking theoretically interesting questions about the welfare state.

Why do scholars conduct normal science? One part of our answer comes from the welfare regime literature itself, especially in its tendency to confuse the terms 'typology' and 'ideal-types'. This confusion has prevented the emergence of anomalies, thereby impairing progress. We find another part of our answer in the combination of individuals' (including scholars) tendency to prefer certainty (e.g., to know in advance that a research problem is solvable) and from environmental pressures (the 'publish or perish' culture and the need to frame research problems exclusively in terms of variation rather than similarity).

The structure of the paper is as follows. First, we introduce the Kuhnian account of normal science and specify what problems this potentially causes (section 2). Next, we demonstrate that the three worlds' literature shows signs of normal science by means of examples of recent studies

(section 3). Then we discuss the status and contribution of the three worlds' typology and indicate how the confusion between 'typologies' and 'ideal types' has been conducive to the path of normal science. We also propose how to solve this confusion (section 4). We subsequently address the question of why we as researchers *en masse*, and generally with great enthusiasm, conduct 'normal science', now zooming in on characteristics of the researchers and their environment (section 5). We end by addressing the challenging question of how to move beyond normal science (section 6).

2. The Kuhnian account of normal science

What is the Kuhnian account of normal science? There has always been discussion about this (see e.g., Lakatos and Musgrave 1970). According to Kuhn (1970: 5), normal science is 'the activity in which most scientists inevitably spend almost all their time'. It is 'predicated on the assumption that the scientific community knows what the world is like. Much of the success of the enterprise derives from the community's willingness to defend that assumption, *if necessary at considerable cost*' (idem, emphasis added). Scholars conducting normal science, so argues Kuhn, work on solvable puzzles and do not question the theory (see above). If this were the case, says Popper (1970: 52), 'the "normal" scientist, as Kuhn describes him, is a person one ought to be sorry for. (...) The "normal" scientist (...) has been taught badly'. Instead of being trained and encouraged in critical thinking, the "normal" scientist, 'has been taught in a dogmatic spirit: he is a victim of indoctrination' (idem: 52). For Popper, a researcher who is content to solve "puzzles", is not a *pure* scientist but an *applied* scientist (idem). This applied scientist tries to tackle routine problems, not fundamental ones (idem). While Popper admits 'that this kind of attitude exists', he also believes 'that Kuhn is mistaken when he suggests that what he calls "normal" science is normal' (idem).

The discussion between Popper (and others) and Kuhn about normal science, although interesting, is a philosophy of science debate that we cannot solve here. This is also not our intention. Our intention is to demonstrate that the Kuhnian account of normal science is what we see in large parts of the welfare regime literature. Note that this does not mean that we argue that

this literature is useless; normal science is not necessarily bad. This is also not Kuhn's position. For Kuhn (1970: 52), 'normal science (...) is a highly cumulative enterprise, eminently successful in its aim, the steady extension of the scope and precision of scientific knowledge'. The crux of the argument is this: 'one standard product of the scientific enterprise is missing' in normal science, it 'does not aim at novelties of fact or theory' (idem).

Normal science implies that the core of the regimes approach (that there are three or four regimes) becomes almost impervious to falsification in that all research *must* find, somehow, the three or four way clustering. Here, there is the risk that researchers, perhaps unconsciously, interpret the empirical evidence such that the world will behave according to the welfare regimes approach's core. But it is in our view precisely the extension of the scope and precision of scientific knowledge, without novelties of fact or theory, that has come to characterize large parts of the welfare regime literature. This holds for instance for the literature, discussed below, that attempts to 'force' the regime typology on to various topics and sectors (education, health care, environment, but also happiness and terrorism, etc.) and beyond the world of advanced capitalism with the purpose, or so it seems, of showing that there are three or more worlds of something. This starts to resemble a Lakatosian negative heuristic, which prevents researchers to question the so-called hard core of the regime approach (Lakatos 1970: 133–134). A substantial part of the existing welfare regime literature only specifies and re-specifies welfare regimes, failing to anticipate novel and interesting facts (Lakatos 1970: 191–192).

3. Normal science in the welfare regime literature

Through the mechanisms and habits of normal science, there is the typical tendency to use the paradigmatic welfare typology for formulating questions of which it is plausible that they are answerable within the paradigm: Kuhn's solvable puzzles. This holds for instance for those studies that assess if the regimes also come to the fore in a 'new' area. Examples hereof are West and Nikolai (2013), who examine whether there is regime-clustering in terms of education systems (yes,

four clusters), Bergqvist et al. (2013), who review the findings of 34 studies that examine the relationship between population health and health inequalities and welfare regimes (findings vary across the 34 studies), and Krieger and Meierrieks (2011), who study whether different worlds of welfare capitalism differently affect homeland terrorism (yes, to some extent).

Staying within normal science also applies to the numerous studies that use the three worlds' typology to argue that some country or some region scores so different on the regime features that it merits to get its own type (see also section 4). Examples hereof are Castles and Mitchell's (1992) radical type, Ragin's (1994) 'spare' type, Goodwin's (2001) post-productivist type; Cinalli and Giugni's (2013) youth unemployment regimes, and several authors' hybrid, Eastern European, Confucian or Antipoden types. Staying within normal science also applies to those scholars arguing that the within-cluster variation is so substantial that types do not coincide with regional blocs (Hudson and Kühner 2012), and with attempts to bridge typologies or combine them (e.g. the Varieties of Capitalism and Welfare Regime typologies) into one overarching, unified typology (Schröder 2013). These and like papers and books, in our view, precisely fit Kuhn's description of normal science.

Another consequence of the regime typology's paradigmatic status is that researchers are constantly looking for (and finding) the 'correct' type of variation and clustering. The compilation of datasets and the focus on various variables, for example, seem driven by the conviction that these will demonstrate the existence of the three (or more) worlds or point to some puzzling outlier that will warrant an effort to renormalize it. For example, in a prize-winning article, Stoy (2014) argues that the study of social services has long been neglected but that recent research on this topic has been erratic in the sense that some confirm and others disconfirm three worlds of welfare services (see Jensen 2008). In normal science, this situation needs to be renormalized and this is done by adopting a 'holistic' approach covering both social and health services and by providing new and updated data. Stoy's hierarchical cluster analysis reveals the three traditional worlds with the usual suspects as well as a rest category labelled 'rudimentary'. The conclusion is that the new worlds of

welfare services resemble the old worlds of transfers, with which we – according to the author – also acquire an improved understanding of welfare regimes in general. Exactly what this new understanding is, however, remains implicit.

Similarly, Kuitto (2011: 349) asks a typical normal science question in her study of disaggregated social spending: ‘Do disaggregated welfare spending patterns reveal significant cross-country variation among European welfare states in line with a theoretically meaningful distinction between welfare regimes?’ Even though Kuitto reminds us that ‘(w)elfare state classifications and typologies offer a meaningful point of origin for theory-building and causal analysis’ (348), she loses sight of this in the course of the paper. The conclusion is that ‘divergent welfare spending patterns indeed can be identified along (...) functionally different dimensions of welfare expenditure. The mature European welfare states cluster in a way which coincides largely with the regimes or worlds of welfare identified in previous studies’ (361). To be fair, in the conclusion Kuitto returns to the issue of theoretical relevance and causal mechanisms, but only in a call for further research.

One interesting example of researchers interpreting empirical evidence such that it confirms the three worlds is Willemse and De Beer’s (2012) study with the classic normal science title ‘Three worlds of educational welfare states?’ The researchers correctly observe that education, although a crucial part of the welfare state, is an under-researched topic in the comparative literature. The study then takes two of Esping-Andersen’s main concepts (decommodification and stratification) to test whether there are also three education regimes. Their finding is that the social democratic model fully lives up to expectations, but the other regimes do not.

Willemse and De Beer develop measures and produce a figure (see figure 1), which links the decommodification index to the stratification index, and argue that three clusters can be distinguished. A Nordic, social democratic group with high decommodification scores, a liberal group with low scores on decommodification and the conservative countries with somewhat higher decommodification scores. France, Italy and Portugal are described as conservative countries that appear more liberal. Belgium is seen as an odd outlier, because it is placed in between the groups

and is defined as a hybrid case. However, it seems that these results may have been driven by the 'need' to see the three regimes, not by the data and the positions the various countries occupy in the two-dimensional space. In fact, the authors have drawn the circles by exclusive reference to one dimension only, the one that fitted the three worlds' typology best. We could easily draw the circles differently, for instance around a group with less than 0.2 stratification AND above 0 decommodification: New Zealand, Sweden and Finland; a group with ≤ 0.4 stratification AND negative decommodification scores: France, Canada, Ireland, UK, Portugal, Italy, Australia, and the US; a group of countries scoring high on stratification (≥ 0.5) AND positive on decommodification: Belgium, Austria, and the Netherlands; and finally a group with high stratification and positive negative decommodification: Switzerland, Spain, and Germany. Then there are four worlds of education, which do not correspond to the three worlds of welfare capitalism.

But what was Willemse and De Beer's theoretical argument to look at decommodification and stratification in education in the first place? Willemse and De Beer (2012: 106-107) state that whereas 'Esping-Andersen (1990) applied the concepts of decommodification and stratification to the fields of social security and pension schemes only', they 'examine whether these central concepts of Esping-Andersen's welfare state analysis can also be applied to higher education policies'. Note that we do not argue that purely empirical questions cannot be valuable – they often are. Our point is that to further the comparative literature on welfare regimes, we need to move beyond within-paradigm empirical questions of which the theoretical relevance or goal is not specified.

Figure 1 about here

Might it be that we have become so totally preoccupied with imposing three-way (or four-way) clustering on any type of data, that we in fact are missing potentially interesting variation that does not correspond to the normal science variation we are familiar with? Might it be time to rethink critically the value of the regime typology? Are we, in fact, studying the right phenomena of

theoretical and empirical interest? Might it be that we are greatly exaggerating differences? Are, say, Denmark and Norway really that different from, say, the Netherlands and Austria that we need to put them in different types? Or, alternatively, are we not too exclusively focusing on problems of variation, while the most intriguing empirical puzzles concern similarity (Van Kersbergen 2010), such as the negligible differences between the left and the right when it comes to health care politics (Jensen 2014)? Before coming back to these issues in the paper's conclusion, we first explain why the confusion in the welfare regime literature between typologies and ideal types has been conducive to normal science (section 4) and why scholars conduct normal science, generally with great enthusiasm (section 5).

4. How the confusion between typologies and ideal types has been conducive to normal science³

Why and how has the confusion between typologies and ideal types been conducive to the normal science of the welfare regime literature? We see two, related, reasons for this. First, this confusion leads to a situation in which studies mainly focus on empirical questions that are answerable within the paradigm, while – as we indicated above – their theoretical relevance remains unspecified. Second, and more importantly, the confusion has stood in the way of subjecting the three worlds' typology to a more fundamental scrutiny. In order to elaborate and substantiate these arguments, we first discuss the status and contribution of the three worlds' typology and the intellectual confusion that has surrounded it (see also Emmenegger et al. 2015 [this issue]).

In general, typologies – like Esping-Andersen's welfare regime typology – are analytically useful empirical classificatory devices that reduce observed complexity by cataloguing existing cases as meaningful representatives of some theoretically relevant dimensions (variables) (see e.g., Collier et al. 2012). Cases with similar or identical scores on the variables are classified as belonging to the same type. Although a typology is necessarily constructed on the basis of some implicit or explicit (proto-)theory (to determine the relevant dimensions), it is essentially a classificatory device that helps to arrange the observable empirical 'mess' in a more ordered, transparent and therefore

comprehensible manner. A good typology, in addition to being efficient and reliable in reducing complexity, should be exhaustive and mutually exclusive (Gerring 2012: 144–151). An *exhaustive* typology 1) includes all theoretically relevant dimensions and 2) can assign all existing cases to one of the types. That a typology should be *mutually exclusive* means that an empirical case can be assigned to one type only. Despite this clear definition of a typology and the criteria it should fulfil, the literature has been plagued by an unfortunate mixing-up of the typology tool and the idea of an ideal-type. We think that this been one factor underlying the development towards normal science in the welfare regime literature.

Esping-Andersen's typology meets most of the criteria of a good typology. Regarding *efficiency* in reducing complexity, the typology did a proper job by bringing the 18 welfare states (for which data were available at the time) down to three regimes. Notwithstanding some arbitrary choices (see e.g. Scruggs and Allan 2006; Van Kersbergen and Vis 2014: chapter 4), it did so *reliably*. Moreover, the typology is arguably *exhaustive*. Esping-Andersen's re-specification of the welfare state as a regime – using the dimensions of decommodification, stratification and state-market-family mix – allowed for a classification of roughly all 18 cases in one of the three types. The single biggest missing aspect in the original classification exercise concerned the role of the family in the public-private mix of social provision.

The typology is not unambiguously *mutually exclusive*, though, because there were several cases that could not clearly be assigned to a type. For instance, as Esping-Andersen (1990: 51) honestly reported, Belgium, the Netherlands and Austria had high levels of decommodification and were placed with the social democratic countries. However, the distance between Austria (the lowest of the high scores) and Denmark (the lowest of the social democratic countries) was 7 points (38.1-31.1), exactly the same distance as between Austria and Italy (the lowest in the medium group) (31.1-24.1). It is unclear what criterion Esping-Andersen used for the classification.

Similar classification difficulties emerged regarding the typology's other dimensions, such as the problem of the high degree of benefit universalism in the liberal countries Canada and

Switzerland (see Van der Veen and Van der Brug 2013). Esping-Andersen did point to the problematic issues and cases, but did not always justify well the choices he made to ensure that all cases were assignable to one of the three types, and to one type only.

Esping-Andersen thus clearly struggled with the placement of some of the cases. In his re-examination of *Three Worlds*, he (1999: 73ff) explicitly inculcated that his typology was primarily a classification device and exercise. He furthermore explained that typologies are useful because they allow for parsimony, may highlight underlying patterns and causalities, and help to generate testable hypotheses. He also defended the typological approach against various criticisms. In trying to answer the question whether the typology was robust and valid over time, however, Esping-Andersen (1999: 86) confusingly argued that since typologies 'are, in a sense, ideal types, there are bound to be ambiguous cases'. In the context of a typology, however, cases cannot be labeled as ambiguous, because empirical cases are what they are. The only thing that might be ambiguous is the way the researcher classifies a case as of a certain type. If it is difficult to categorize a specific case, as discussed above, the typology has the problem that it is not mutually exclusive. In that case, it is the typology that is ambiguous, not the case.

The problem is that Esping-Andersen confused the typological approach with the ideal type approach, a misunderstanding that subsequently has been reproduced in much of the literature (e.g. Vrooman 2009: 217; Arts and Gelissen 2010: 572ff; Powell and Barrientos 2011; Aspalter 2013; but see Rice 2013; Becker 2009: 8ff). An ideal type is a theoretical construct to which an empirical case can have no, some, much etc. correspondence, whereas a typology is a classification device in which all empirical cases must find a place as belonging to only one type. In a typology, an empirical case either belongs to a type or not. In an ideal-typical analysis, the question concerns the extent to which an empirical case fits the theoretical ideal. The question of 'goodness of fit' only makes sense when working with ideal types.

The distinction between the three worlds as a typology or as ideal types is important, because it helps to distinguish between constructive and to-the-point criticisms and not very useful

(because ill-conceived) attacks on the three worlds, and between adequate and deficient replies to criticism. For example, a valid criticism is that the chosen dimensions for constructing the typology make it impossible to classify all the existing welfare states (see Castles et al. 2010). As we also discussed in the previous section, some scholars have proposed to broaden the three-way typology by including a wage-earner model that guarantees income through another dimension of empirical variation, namely the wage arbitration system. Others argued that it would be useful to add a Mediterranean type because Southern welfare states are special in (at least) one dimension: they only have residual social assistance programs. Yet others have taken up the task of developing and expanding the regime typology beyond the world of advanced welfare states, exploring to which extent distinct patterns are discernible among welfare states outside Europe, North America and the Antipodes, such as in Asia, Latin American, and Eastern Europe and Russia.⁴

Another valid criticism would be that the regime typology is not exhaustive because it did not include all the theoretically relevant dimensions, including the institution of the family or household, social services (Esping-Andersen 1999: 73; Jensen 2008), and the gender dimension (O'Connor 1996; Orloff 2009; see Hook 2015 [this issue]). Most forcefully, feminist scholarship has demonstrated the analytical incompleteness of the regime approach by showing how assumptions about gender relations underlie welfare arrangements and constitute them. This stimulated criticism of the gender-blindness of the regime approach's central concepts, particularly decommodification, and of the 'falsely universalizing (implicitly masculinist) analytic frames' (Orloff 2009: 319) of the field of comparative welfare state studies. This feminist critique eventually culminated in a radical break with the regime typology. In a literature review, Lewis (2006: 388) argued that the classificatory and explanatory power of the regime approach is minimal in the area of work-family reconciliation policies. She therefore criticized the authors of the reviewed works for uncritically and wrongly using the typology for case selection (and for misguidedly using *Three Worlds* as a source of facts). Lewis underlined that previous feminist scholarship had already amply demonstrated that the typology completely breaks down in the area of work-family relations.⁵

An ill-conceived criticism of the typology concerns the argument that the worlds of welfare are far too complicated to be captured by any typology. This is, for instance, the position taken by Schubert et al. (2009). In fact, they argue that we need to take a step back to a pre-comparative stage to be able to document and highlight the real complexities of social and welfare policies. As a result, studying the diverse European Union systems of welfare necessarily means studying 27 individual countries on as many variables as possible. That, as they realize, is also a step back analytically. After 27 single country chapters, the conclusion is that there is no such thing as *the* welfare state or *the* European social model and that there is no systematic pattern of variation in the key dimensions of spending, financing, actors, and '*Leitmotifs*' (guiding principles). This type of criticism of typological work misses the point of what a typology is all about: a meaningful, efficient and reliable *reduction of complexity* for analytical and comparative purposes. It also misses the point that every typology that reduces complexity does so by necessarily simplifying reality. That is not a disadvantage, but it is what typologies are *meant* to do. Therefore, to criticize a typology that it does not do justice to the full complexity of empirical reality is entirely beside the point. What we end up with the Schubert et al. approach is a complete loss of analytical power without real gains in terms of empirical accuracy (see Arts and Gelissen 2010: 581).⁶

We think that the confusion about typologies versus ideal types has added to the limited progress in terms of developing a stronger theoretical foundation of contemporary and future welfare state models as explanatory categories (Arts and Gelissen 2010; Powell and Barrientos 2011). First, as we showed in the previous section, instead of trying to arrive at novel theoretical insights, the lion share of studies addresses empirical questions that are answerable within the paradigm. The confusion about typology versus ideal types is conducive to this tendency because it helps to create a situation in which there are such typical empirical questions to be asked as how to handle cases considered to be ambiguous (which would be incorrect labelling, as we argued above). Scholars working on such empirical questions typically leave unaddressed *why, theoretically*, these questions are interesting to ask in the first place; or – when mentioned, referred to as a task for

future research. Second, the confusion has made it at times difficult to make a distinction between to-the-point comments and irrelevant criticisms. This rendered the three worlds' typology virtually impervious to a more fundamental scrutiny, because difficult to classify cases – a *typological* problem – could always be reasoned away by arguing that obviously no real type ever fits an *ideal-type*. This is exactly what Esping-Andersen did when he defended his approach with the argument that since typologies are to some extent ideal types there will always be ambiguous cases. If there had not been confusion about typology versus ideal-type, such reasoning away would not have been possible. Under such a situation, scholars would be better able to identify anomalies. In sum, we think that confusing a typology and an ideal-type and defending the one with the other prevented the emergence of real anomalies and thereby helped to establish the regime approach as normal science in the Kuhnian (1970) sense.

5. Why do researchers practice normal science?

At this point, the reader may wonder why it is that researchers (including the present authors!) not only practice normal science, but that they generally do so with much enthusiasm. The confusion in the literature on 'typology' versus 'ideal type' is one part of the answer because, as we explained in the previous section, this confusion is conducive to normal science. Typological anomalies can be reasoned away by reference to the fact that ideal types always generate ambiguous cases. A second reason is found in human psychology and holds that individuals – researchers included – prefer *certainty* to uncertainty. Individuals vary in their need for certainty, but in general, the 'the brain (...) craves certainty' (Rock 2008: 47). This craving is, at least partly, responsible for researchers' tendency to be intolerant of new theories invented by others and reluctant to invent new theories themselves (Kuhn 2012 [1962: 24]). Sticking with an existing theory, or an existing paradigm, brings certainty regarding the results of a research endeavour. Within normal science, if a researcher is 'skilful enough, [s]he will succeed in solving a puzzle that no one before has solved or solved so well' (Kuhn 2012 [1962]: 38). It may be a difficult journey, but you know that the puzzle is solvable.

Factors relating to researchers' preference for certainty are the culture of *publish or perish* and the existence of *publication bias*. In the past decades, the incentives to publish in international refereed journals have increased enormously. This environment induces researchers to work on papers that they assume will be publishable. Since reviewers typically also crave certainty, staying within the paradigm of normal science is the safest route to take. This implies that it is hard if not impossible to avoid the use of the typology, even if it does not really make sense. In the prologue to this special issue, Emmenegger et al. (2015) analyze the 138 articles citing *Three Worlds* published in the *Journal of European Social Policy* (JESP) between 1991 and 2013. Of the 30 per cent of the articles engaging 'deeply' with the book, i.e. whose analysis related directly to the *Three Worlds*, only 12 per cent formulated some form of fundamental criticism. This means that less than four per cent of all JESP articles *Three Worlds* voiced fundamental criticism. Emmenegger et al. (p. X) therefore conclude that 'there is some evidence that *Three Worlds* has turned into a book that needs to be cited, determines the case selection, organises the presentation of the empirical evidence – in short, a book whose findings are taken for granted instead of constantly questioned'. In our view, Emmenegger et al.'s analysis of JESP articles citing *The Three Worlds* demonstrates that incrementally testing and confirming existing theory (i.e. normal science) is precisely what has been going on in a large part of the existing regime literature.

Moreover, based on the empirical welfare regime studies published, it seems that the research problem must be framed as a problem of (regime) *variation*. This may be an effect of both the dominant prescription of (social) science to maximize variation on the dependent variable and because normal science requires the identification of various 'worlds'. Reasoning within the boundaries of normal science, reviewers may conclude that a submitted article perfectly meets the conditions of social science as a craft and yet sense that the paper's contribution is trifling. However, one realizes that advising rejection based on such a hunch is 'not done', thereby bolstering normal science.

The existence of a publication bias reinforces these propensities. The bias is ‘the tendency for statistically significant findings to be published over nonsignificant findings’ (Ferguson and Heene 2012: 555). In general and across all disciplines, the large majority (>80 per cent) of published findings are positive results, i.e. results supporting a hypothesis (Fanelli 2010).⁷ Translated to welfare regime studies, publication bias entails the tendency for results to be published that are largely in line with the three worlds’ typology over results that are not. Or studies that apply the three worlds’ typology irrespective of whether there is a theoretically compelling reason to do so (like some of the ‘non-deeply’ engaging JESP articles citing *Three Worlds*). Of course, since we have no access to all papers ever written on the three worlds’ typology, we cannot be sure that there exists this positive result bias. Still, Emmenegger et al.’s discussion of the studies that get published as well as our own discussion above indicate that also this literature likely knows the so-called file drawer problem, whereby papers that fail to support the welfare regime paradigm’s core remain in the proverbial drawer and thus unpublished.

6. Discussion: Moving beyond normal science?

If the welfare regime literature has indeed turned into Kuhnian normal science are there then any opportunities for innovation? Should we, for instance, as an alternative to the heuristic of the three worlds’ typology, turn to developing theoretically informed ideal types and start examining to what degree welfare states correspond to these ideal types (for examples, see Kvist 2007; Vis 2007; Hudson and Kühner 2012; Ciccio and Bleijenbergh 2014)? While this, in our view, would clearly be a step forward, because it clarifies the typology/ideal-type confusion and because it introduces a more explicit theoretical starting point, it would not move us beyond the normal science of the welfare modelling business.

Or should we develop explanatory typologies, moving beyond the purely descriptive classificatory exercise? Explanatory typologies, i.e. ‘multidimensional conceptual classifications based on an explicitly stated theory’ (Elman 2005: 296), have a long standing in International

Relations' research but are still rare in comparative welfare state research (see Collier et al. 2012; Møller and Skaaning 2013). Elman (296-298) states:

In an explanatory typology, the descriptive function follows the conventional usage, but in a way that is heavily modified by its theoretical purposes. The constituent attributes are extracted from the variables of a preexisting theory. The dimensions of the property space (its rows and columns) reflect alternative values of the theory's independent variables, so each cell in the space is associated with predicted values of the theory's intervening or dependent variables. This association changes the descriptive question being answered from 'What constitutes this type?' to 'If my theory is correct, what do I expect to see?'

The question of what kind of empirical variations and what kind of similarities are relevant for this explanatory purpose are difficult to answer without an understanding of the theoretical foundations of welfare state regimes. For an explanatory typology to work, the existence of pre-existing theory is crucial. In other words, we suggest that the field should rethink the theoretically interesting questions about the welfare state. What is the current and most relevant theoretical substance of the welfare state?

This, of course, presupposes that we actually already know what we wish to explain, but it is precisely here that the normal science character of the field has been standing in the way of a thorough rethinking of the welfare state's theoretical substance. One way to go about is to define or reformulate what it is that, ultimately, we are – or should be – interested in. As we indicated above, feminist scholarship has given one answer and one could argue that gender-based research has moved beyond the normal science of the mainstream. 'Indeed, so fundamental has been the feminist challenge,' writes Orloff (2009: 318), that 'gender studies can arguably be said to represent a paradigmatic change of the Kuhnian variety'. Be that as it may, this literature's emphasis on the study of social outcomes should be taken to heart. In this sense, we suggest that we go back to and revitalize the main message of part II of *Three Worlds* and indeed focus on the *social outcomes* that a

welfare regime produces and that matter. Welfare state outcomes fundamentally affect people's interests, capabilities, life chances, and life cycles, so we should study the welfare state's role in producing or moderating inequality, poverty, stratification, mobility, education, employment, etc.

We can take inspiration from the following, not so often quoted, passage in *Three Worlds*:

(...) be it in contemporary Scandinavia, western Europe, or even North America, the welfare state is becoming deeply embedded in the everyday experience of virtually every citizen. Our personal life is structured by the welfare state, and so is the entire political economy. Given the magnitude and the centrality of the welfare state, it is (...) unlikely that we shall understand much of contemporary society unless it becomes part of our models (Esping-Andersen 1990: 141).

No matter how one evaluates the application of the regime typology in the past 25 years, this key observation of Esping-Andersen easily could – and should – inspire welfare state research for the next 25 years.

Acknowledgement

For their helpful comments, we wish to thank the participants of the Comparative Welfare Seminar (University of Southern Denmark, Odense, 10 October 2012), the participants of the special issue seminar (The Danish National Institute of Social Research, Copenhagen, 24 June 2013), Carsten Jensen, Magnus Rasmussen, the anonymous reviewers and the editors of this special issue.

References

- Arts, W. A. and Gelissen, J. (2010) 'Models of the Welfare State', in F.G. Castles, S. Leibfried, J. Lewis, H. Obinger and C. Pierson, and Christopher (eds), *The Oxford Handbook of the Welfare State*. Oxford: Oxford University Press, pp. 586–583.
- Aspalter, C. (2013) 'Real-typical and Ideal-typical Methods in Comparative Social Policy', in B. Greve, (ed.) *The Routledge Handbook of the Welfare State*. Oxon and New York: Routledge.
- Becker, U. (2009) *Open Varieties of Capitalism. Continuity, Change and Performance*. Basingstoke: Palgrave Macmillan.
- Bergqvist, K., Åberg Yngwe, M. and Lundberg, Olle (2013) 'Understanding the Role of Welfare State Characteristics for Health and Inequalities – An Analytical Review', *BMC Public Health*, 13: 1234.
- Castles, F. G. and Mitchell, D. (1992) 'Identifying Welfare State Regimes: The Links between Politics, Instruments and Outcomes', *Governance: An International Journal of Policy and Administration*, 5 (1): 1–26.
- Castles, F. G, Leibfried, S., Lewis, J., Obinger, H. and Pierson, C. (eds) (2010) *The Oxford Handbook of the Welfare State*. Oxford: Oxford University Press.
- Ciccia, R. and Bleijenbergh, I. (2014) 'After the Male Breadwinner Model? Childcare Services and the Division of Labor in European Countries', *Social Politics*, 21 (1): 50–79.

- Cinalli, M. and Giugni, M. (2013) 'New Challenges for the Welfare State: The Emergence of Youth Unemployment Regimes in Europe?', *International Journal of Social Welfare*, 22 (3): 290–299.
- Collier, D., LaPorte, J. and Seawright, J. (2012) 'Putting Typologies to Work. Concept Formation, Measurement, and Analytic Rigor', *Political Research Quarterly* 65 (1): 217–232.
- Elman, C. (2005) 'Explanatory Typologies in Qualitative Studies of International Politics', *International Organization* 59(2): 293–326.
- Emmenegger, P., Kvist, J. Marx, P. and Petersen, K. (2015) '25 Years and Still Going Strong: *Three Worlds of Welfare Capitalism*', *Journal of European Social Policy*.
- Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Esping-Andersen, G. (1999) *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Fanelli, Daniele (2010) "'Positive" Results Increase Down the Hierarchy of the Sciences', *PLoS ONE*, 5 (4): e10068.
- Ferragina, E. and Seeleib-Kaiser, M. (2011) 'Welfare Regime Debate: Past, Present, Futures?', *Policy & Politics* 39 (4): 583–611.
- Ferregina, E., Seeleib-Kaiser, M. and Tomlinson, M. (2013) 'Unemployment Protection and Family Policy at the Turn of the 21st Century: A Dynamic Approach to Welfare Regime Theory', *Social Policy & Administration*, 47 (7): 783–805.
- Ferguson, C. J. and Heene, M. (2012) 'A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null', *Perspectives of Psychological Science* 7 (6): 555–561.
- Gerring, J. (2012) *Social Science Methodology: A Unified Framework*. Cambridge: Cambridge University Press.
- Goodwin, R. E. (2001) 'Work and Welfare: Towards a Postproductivist Welfare Regime', *British Journal of Political Science* 31 (1): 13–39.

- Hook, J. (2015) 'Incorporating Class into Work-Family Arrangements: Insights from and for *Three Worlds*', *Journal of European Social Policy*.
- Hudson, J. and Kühner, S. (2012) 'Analyzing the Productive and Protective Dimensions of Welfare: Looking Beyond the OECD', *Social Policy & Administration* 46(1): 35–60.
- Jensen, C. (2008) 'Worlds of Welfare Services and Transfers', *Journal of European Social Policy* 18 (2): 151–162.
- Jensen, C. (2014) *The Right and the Welfare State*. Oxford: Oxford University Press.
- Jessop, B. (2002) *The Future of the Capitalist State*. Cambridge: Polity Press.
- Kammer, A., Niehues, J. and Peichl, A. (2012) 'Welfare Regimes and Welfare State Outcomes in Europe', *Journal of European Social Policy* 22 (5): 455–471.
- Krieger, T. and Meierrieks, D. (2011) 'What Causes Terrorism?', *Public Choice* 147 (1-2): 3–27.
- Kuhn, T. S. (2012 [1962]) *The Structure of Scientific Revolutions*. London and Chicago: The University of Chicago Press.
- Kuhn, T. S. (1970) 'Logic of Discovery or Psychology of Research?', in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Kuitto, K. (2011) 'More than Just Money: Patterns of Disaggregated Welfare Expenditure in the Enlarged Europe', *Journal of European Social Policy* 21 (4): 348–364.
- Kvist, J. (2007) 'Fuzzy Set Ideal Type Analysis', *Journal of Business Research* 60 (5), 474–481.
- Lakatos, I. (1970) 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Lakatos, I. and Musgrave, A. (1970) (eds) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Møller, J. and Skaaning, S.-E. (2013) 'Explanatory Typologies as a Nested Strategy of Inquiry: Combining Cross-Case and Within-Case Analysis', manuscript Aarhus University.

- O'Connor, J. (1996) 'From Women in the Welfare State to Gendering Welfare State Regimes', *Current Sociology* 44 (2): 1–124.
- Orloff, A. S. (2009) 'Gendering the Comparative Analysis of Welfare States: An Unfinished Agenda', *Sociological Theory* 27 (3): 317–343.
- Popper, K. (1970) *Normal Science and its Dangers*, in I. Lakatos and A. Musgrave (eds), *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press.
- Powell, M. and Barrientos, A. (2011) 'An Audit of the Welfare Modelling Business', *Social Policy & Administration* 45 (1): 69–84.
- Ragin, C. C. (1994) 'A Qualitative Comparative Analysis of Pension Systems. in T. Janoski and A. M. Hicks (eds), *The Comparative Political Economy of the Welfare State*. Cambridge: Cambridge University Press.
- Rice, D. A. (2013) 'Beyond Welfare Regimes: From Empirical Typology to Conceptual Ideal Types', *Social Policy & Administration* 47 (1): 93–110.
- Rock, D. (2008) 'SCARF: A Brain-Based Model for Collaborating with and Influencing Others', *The NeuroLeadership Journal* 1: 44–52.
- Schröder, M. (2013) *Integrating Varieties of Capitalism and Welfare State Research: A Unified Typology of Capitalisms*. Houndmills: Palgrave Macmillan.
- Schubert, K., Hegelich, S. and Bazant, U. (2009) 'European Welfare Systems: Current State of Research and Some Theoretical Considerations', in idem (eds), *The Handbook of European Welfare Systems*. Oxon: Routledge.
- Scruggs, L. A. and Allan, J. P. (2006) 'The Material Consequences of Welfare States: Benefit Generosity and Absolute Poverty in 16 OECD Countries', *Comparative Political Studies* 39 (7): 880–904.
- Scruggs, L., Jahn, D. and Kuitto, K. (2013) *Comparative Welfare Entitlements Data Set 2, Version 2013-08*. Available at: <http://cwed2.org/>.

- Stoy, V. (2014) 'Worlds of Welfare Services: From Discovery to Exploration', *Social Policy & Administration* 48 (3): 343–360.
- Titmuss, R. M. (1974) *Social Policy: An Introduction*. London: Allen and Unwin.
- Van der Veen, R. J. and Van der Brug, W. (2013) 'Three Worlds of Social Insurance: On the Validity of Esping-Andersen's Welfare Regime Dimensions', *British Journal of Political Science* 43 (2): 323–343.
- Van Kersbergen, K. (2010) 'Comparative Politics: Some Points for Discussion', *European Political Science* 9 (1): 49–61.
- Van Kersbergen, K. and Vis, B. (2014) *Comparative Welfare State Politics: Development, Opportunities and Reform*. Cambridge: Cambridge University Press.
- Vis, B. (2007) 'States of Welfare or States of Workfare? Welfare State Restructuring in 16 Capitalist Democracies, 1985–2002', *Policy and Politics* 35 (1): 105–122.
- Vrooman, J. C. (2009) *Rules of Relief. Institutions of Social Security, and Their Impact*. The Hague: The Netherlands Institute for Social Research/SCP.
- West, A. and Nikolai, R. (2013) 'Welfare Regimes and Education Regimes: Equality of Opportunity and Expenditure in the EU (and US)', *Journal of Social Policy* 42 (3): 469–493.
- Willemse, N. and De Beer, P. (2012) 'Three Worlds of Educational Welfare States? A Comparative Study of Higher Education Systems across Welfare States', *Journal of European Social Policy* 22 (2): 105–117.

Funding

Barbara Vis' research is supported by a Vidi grant from the Netherlands Organisation for Scientific Research (NWO, grant nr.: 452-11-005). Kees van Kersbergen received no specific grant from any funding agency in the public, commercial or not-for-profit sectors for this research.

Notes

¹ As Esping-Andersen (1999: 74, fn. 1) acknowledges, his typology has a close affinity with Titmuss' (1974) earlier one (see Van Kersbergen and Vis, 2014, chapter 4 for a discussion).

² Note that this is a large body of literature. To give an idea of the number of studies to which our argument could apply, we did a Google Scholar search (in July 2014) of Esping-Andersen's *Three Worlds*. Of the >19,000 citations to this book, 4,070 include the term 'welfare regime'. While probably not all of these publications fall in what we label normal science, these are studies that (1) cite Esping-Andersen's *Three Worlds* and (2) use the term 'welfare regime' in their work and that thus could be examples of normal science. Given that over 4,000 publications fit this category, we think it is plausible to assume that our characterization of the field applies to more than, say, a handful of studies.

³ This section draws on Van Kersbergen and Vis (2014).

⁴ Note that this broadening of the three worlds remains within the confines of normal science, as we explained in section 3.

⁵ Feminist scholarship arguably developed a paradigm of its own, in which *Three Worlds*, Esping-Andersen's later work on gender equality and the role of families (2009), and hence the regime typology hardly plays a role anymore. For instance, in the latest issue of *Social Politics* available at the time of writing (April 2014), there is a special section on "Dimensionalizing Care: Children, disability, gender and class". None of the papers that explicitly use the regime concept refers to Esping-Andersen's work, while in the whole special section there is only one single reference to Esping-Andersen (2009), in a footnote. It is because of this development that the otherwise thriving feminist scholarship has remained largely outside the scope of this paper. We exclusively consider the literature that has continued to engage with Esping-Andersen's regime typology.

⁶ An alternative is an ideal-typical approach, as for instance done in the French regulation approach to political economy (e.g. Jessop 2002). This literature's conception of the welfare state is much wider than that of the three worlds' literature, including the entire capitalist system and everything

related to that in terms of governance. The regulation approach scholars theoretically construct two ideal-types: a welfare regime and a workfare regime. The drawback of this literature is that it fails to test empirically the theoretical expectations systematically and, instead, only presents anecdotal empirical illustrations to argue that there has been a development from welfare to workfare (see Vis 2007).

⁷ For example, Fanelli (2010) finds that of the 2,434 articles from 20 disciplines he studied, 2,045 articles (84%) reported a positive or partial support for the tested hypothesis. The percentage of such findings in the social sciences lied around this average of all disciplines, leaving around 20 per cent of all published articles included in Fanelli's analysis presenting negative findings.