

Implementing the third generation research paradigm in policy implementation research: An empirical assessment

Harald Saetren

Department of Administration and Organization Theory,
University of Bergen, Norway

Public Policy and Administration

2014, Vol. 29(2) 84–105

© The Author(s) 2014

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0952076713513487

ppa.sagepub.com



Abstract

A comprehensive literature review a few years ago found, contrary to a commonly held belief, research on public policy implementation to be still alive and have developed further well into the 21st century in *quantitative* terms. We pursue this line of inquiry by asking whether this conclusion also applies to progress in *qualitative* terms. All articles published in *core journals* in political science, public administration/management and public policy under the label implementation or implementing were used to investigate this research question. Key defining features of the plea for a more *rigorous third generation research paradigm* was used as benchmark to measure progress or lack thereof in the policy implementation literature over more than four decades of research. Our investigation basically supports conclusions from more intuitively based earlier state-of-the-art reviews: (1) policy implementation research has reached a relatively mature stage of development, (2) more progress has been achieved on methodological than theoretical fronts and (3) this field of study progressed fairly rapidly through two previous generations of research, coinciding roughly with the 1970s and 1980s with some assumed characteristics while progress since 1990 has been much slower and more incremental. The latter fact can probably best be explained by the very demanding nature of the third generation research paradigm itself and some inherent tensions and contradictions between its defining features. Our contribution in this respect is that we have been able to provide more detailed and nuanced information about exactly where progress has been achieved and when as well as where it is still lagging.

Corresponding author:

Harald Saetren, University of Bergen, Christiesgate 17, Bergen N-5007, Norway.

Email: harald.satren@aorg.uib.no

Keywords

State-of-the-art, literature review, implementation research, core journals, political science, public administration/management, public policy

Introduction

The previous decade saw several state-of-the-art evaluations of the research literature on policy implementation (Barrett, 2004; deLeon and deLeon, 2002; Hill and Hupe, 2009; O'Toole, 2000, 2004; Schofield, 2001; Schofield and Sausman, 2004; Winter, 2012a). The degree of consensus on main conclusions is surprising considering how much implementation scholars otherwise tend to disagree on conceptual, methodological and theoretical issues. Nevertheless, the main message is mixed. While reporting substantial progress on a number of fronts since the 1970s these scholars also point out that much remains to be done. Thus, progress is reported more on methodological than theoretical issues.

This author believes these state-of-the-art reviews to be on the whole quite plausible. Nevertheless, they share at least three shortcomings. The *first* is that they are usually phrased in fairly general terms, though there are some like O'Toole (2000) who is more specific in this respect than others. Secondly, they are all somewhat dated. A *third* and more serious problem is the lack of systematic data to bolster statements and conclusions. Rare exceptions here include the somewhat dated reviews of the research literature by O'Toole (1986) and Sinclair (2001). The review of the research literature by Saetren (2005) was both more up-to-date and quite comprehensive, but intended only to assess it in *quantitative* terms. Nevertheless, the latter literature review demonstrated clearly the difference between *believing* and *knowing* something about the evolution and structure of this research field. It is now time to examine whether there might be similar discrepancies with respect to evaluative statements about the state of the field in *qualitative* terms. This brings us to the purpose of this article.

Our objective is to provide an up-to-date *empirically* based state-of-the-art assessment of policy implementation research. What is the status of this research field today more than a decade since O'Toole (2000) assessed its progress in *qualitative* terms? How accurate and up-to-date is the latter state-of-the-art review? What exactly are the areas in terms of research methodology where some progress assumedly has taken place and where is it possibly still elusive? Likewise, what does it mean when progress on theoretical issues is said to be less evident? Is the implication that there has not been any progress at all in this respect? These are the main research questions we will pursue. Thus, our aim is to provide more nuanced, detailed and reliable information in this respect relative to previous review assessments.

Evaluative benchmark: The third generation research paradigm

All assessments of this kind require some sort of evaluative criteria or a *benchmark*. In this context, *the third generation research paradigm* promulgated by some leading implementation scholars more than two decades ago (Goggin, 1986; Goggin et al. 1990) serve as an obvious and natural *benchmark*. The unique trait of third generation research is its *rigorous research design* (Goggin et al., 1990: 19). Hence, we will examine to what extent the third generation research paradigm has been implemented in research on policy implementation. That is, we will use it as a *benchmark* to assess *degree of progress* towards the scientific standards that the third generation paradigm entails. Thus, whenever we use concepts like progress, advances, developments or synonyms subsequently they are used primarily descriptively and *not* in a normative manner to assess the degree of implementation of some of the key elements of the third generation research paradigm in policy implementation research.

Nevertheless, the overly *ambitious* and *demanding* nature of this research paradigm must also be noted (O'Toole, 2000; Winter, 2012b). Hence, it should be regarded more as an *ideal type* analytical construct rather than something expected to be found fully implemented in single research projects or even the research field as a whole for that matter.

The following defining features of the third generation research paradigm can be gleaned from a careful reading of Goggin (1986: 335–342), Goggin et al. (1990: 15–19) and Lester et al. (1987: 210–213):

- Key variables must be *clearly defined*.
- *Hypotheses* derived from theoretical constructs should guide empirical analysis.
- More use of *statistical analysis* using *quantitative data* to supplement qualitative analysis.
- More *comparison* across different units of analysis *within* and *across policy sectors*.
- More *longitudinal* research design (i.e. research timeframe of at least 5 to 10 years).

Not all of these aspects of the third generation research paradigm have been investigated by this author. This applies particularly to *conceptual* and *definitional issues* that are of crucial importance. The main reason for this omission is that these issues have been discussed in a thorough manner quite recently (Winter, 2012b). Hence, there is no reason for us to waste scarce space reiterating at length what has already been said in this respect. The bottom line here is that scholarly disagreement about what constitute implementation still exists after 40 years of research on this phenomenon. Furthermore, our assessment will focus more on methodological than theoretical issues though we will deal with both. The rationale behind this choice is twofold. First, there is less agreement on what might constitute progress in

theoretical and normative terms compared to research methodology as we will explicate below. Second, theoretical development depends critically on the quality and rigour of research designs, which is the essence of third generation research paradigm.

The third generation research design implies by itself that there has been two previous generations of implementation research. Furthermore, it is generally taken for granted that there has been an orderly and linear development of research in qualitative terms from the first to the second generation assuming the challenge for researchers now being to make the move forward into the third (Goggin et al., 1990). The conventional story has it that implementation research in the 1970s was dominated by explorative a-theoretical single case-studies based primarily on qualitative data. We know that the first analytical and theoretical frameworks were developed from the mid-late 1970s. This allegedly resulted in more comparative and theoretical-deductive oriented implementation research and hypotheses testing on quantitative data during the 1980s (Lester et al., 1987). This change in scholarly orientation and research from the 1970s to the 1980s was later referred to as a transition from a first to second generation of implementation research (Goggin, 1986; Goggin et al., 1990; Hill and Hupe, 2009; Lester et al., 1987). By the late 1980s, the call for the third generation research paradigm emerged on the premise that further theoretical development was contingent on implementing its more scientific rigorous research design.

Research questions and hypotheses

First, we ask whether implementation research during the 1970s and 1980s really developed in the orderly and linear manner suggested by the terms *first* and *second generations of research*. Did first generation of implementation research really start in the early 1970s as universally claimed or perhaps even earlier? Furthermore, is the change in research focus and content as clear cut from the 1970s to the 1980s as suggested by the generation metaphor? Then, we proceed to investigate the main research question of this article: What has happened to policy implementation research after 1990 compared to the decades before in terms of implementing key elements of the *third generation research paradigm*? Has there been any progress at all as often assumed, and if so, on which fronts? And even then, how much or little progress are we talking about? What is the time line of this progress? Has it happened gradually and increasingly over the four last decades, has it perhaps already stagnated or maybe even declined?

We also ask whether significant regional differences exist in the implementation literature along the research dimensions investigated. And if so, on what continent, if any, are scholars leading the way towards the third generation of implementation research? Is there any reason to look to and learn from European scholarship in this respect as suggested by a recent reviewer (O'Toole, 2000)? In short, what is the status today more than a decade into

the 21st century? What have we achieved and what remains to be done? Have we achieved as much progress as can be reasonably expected given the very ambitious and demanding nature of the third generation research paradigm? And finally, how reasonable and realistic is the third generation research paradigm as a bench mark in this respect?

After having summed up and discussed our major findings, we conclude with some comments on the accuracy of previous literature reviews in this respect and the feasibility and prospect of further progress towards the third generation research paradigm.

Data and methodology

Our data set includes all English language articles published in *core journals* with *implementation* or *implementing* as *title word* in the period from 1953 to 2009. Core journals are those journals primarily devoted to general theory development in *political science* or its two sub-disciplines *public administration/management* and *public policy*. The rationale behind this selection of journals is the assumption that contributions from political science and its related above mentioned sub-disciplines are more critical to theory development in the field of policy implementation than those from other academic disciplines. We recognize that other disciplines such as *management* and *organization theory* also play an important role in this respect, though on the whole less so. We therefore classify journals related to these other academic fields as *near-core*.

A *caveat* is in order here with respect to our data source. It should be noted that another major type of publication – *books* – are not included in our investigation of the research literature. How much bias that represents with respect to our findings and conclusion is hard to know. It should, however, be noted that this bias is somewhat ameliorated by the fact that empirical chapters in many books often are published as articles in core journals prior to their book publication.

Some 547 articles were identified given our search criteria explained below. Primary data sources were *the Social Science Citation Index*, *JSTOR* and *Google Scholar*. Most of these articles were downloaded in full text format when that was feasible or copies were ordered through the inter-library system. Furthermore, every article was registered in an electronic data format based on *the Cardbox* software program. Names of authors and journals, titles of articles and their publication year were systematically registered. The same was the case with respect to other information pertaining to our research questions. The latter, however, could only be gleaned from a closer scrutiny of the articles themselves. All this information was simultaneously *indexed* in a retrievable manner for purposes of analysis by the program.

The choice of the title words *implementation* and *implementing* as criteria for selecting articles for analysis is justified on three grounds. First, after publication of the seminal book titled *Implementation* by *Pressman and Wildavsky* (1973) this concept seems to have become institutionalized among policy scholars as a common denominator for the execution stage of policy process. Second, it is

assumed that those who want to make a contribution to the field of policy implementation are more likely to use these as title words than those who do not share this ambition. Third, this concept provides a reasonably well-defined universe of research literature that can be subjected to independent critical evaluation and replication by anybody who might wish to do so.

However, regardless of selection criteria chosen, there are two types of biases operating that should be kept in mind. The first is called *false positive* selection. This refers to publications that use our key word haphazardly in their title without any real focus or interest in implementation phenomena. They should obviously not have been included. This bias is unavoidable when we perform computer searches in digitalized research literature data bases. It can, however, be detected and rectified through careful reading of selected publications of a manageable size. This has been done in our investigation of the literature on policy implementation. The second bias is called *false negative* and is much more difficult to handle. This refers to publications that *should* have been included because of their relevance but are not. These are publications using key words other than *implementation* in their titles or abstracts. The problem here is that we do not know what these other key words are as this often depends on the author's frame of reference. *Compliance*, *enforcement* and *execution* are some that are often used but there are no doubt many other alternative concepts that are used as substitutes as well. We could have tried to amend this type of bias as well by expanding the key words used in literature searches but decided not to do so on grounds already stated above.

Research results and discussion

The research generations metaphor: From a first generation to a second?

The 1970s do indeed stand out in many ways as different from later decades in terms research profile. The pioneering publications during this decade were indeed less theoretically and empirically oriented. The dominant mode of investigation was the *single case-study* based almost exclusively on *qualitative* data (Tables 3 and 4). This made sense then in a relative young, emerging research field where explorative detailed description of single implementation phenomena aiming at generating hypotheses for future research were particularly valuable and hence also appropriate. We must also bear in mind here that the first analytical and theoretical frameworks that caught the attention of policy scholars (Barrett and Fudge, 1981; Sabatier and Mazmanian, 1979, 1980; Van Meter and Van Horn, 1975) were not published until the mid-late 1970s and early 1980s. A couple of other earlier and similar efforts to provide analytical and theoretical roadmaps for implementation researchers (Bunker, 1972; Smith, 1973) were largely ignored in comparison. Another important finding reported earlier by Saetren (2005) in his more comprehensive literature survey is that the first generation of implementation research did not really start with Pressman and Wildavsky's seminal work published in 1973 or earlier in this decade. Thus, no fewer than 19 core journal articles

with the title words implementation or implementing were published before 1970 of which four date back as early as 1956. The equivalent number reported by Saetren (2005) including near-core journals as well were close to 150. This means that the beginning of the first generation of implementation research must be back dated at least a decade to the early 1960s (see Figure 1).

The transition from the 1970s to the 1980s was prominent in many ways, especially during the first half of the latter decade. The increase in *empirical* and *theoretically oriented* articles is substantial. So is the number of *comparative* studies which doubled at the expense of *single case-studies*. A parallel and related development is a substantial weakening of the predominant role of *qualitative* data from the 1970s and more than tripling of studies using *quantitative* data in the 1980s. The latter trends paved the ground for two other important and pronounced developments in the same decade: the large increase of studies where explicitly formulated *hypotheses* were derived from *theoretical constructs* and subjected to *empirical testing* using *statistical* techniques of analysis (see Tables 1 and 2).

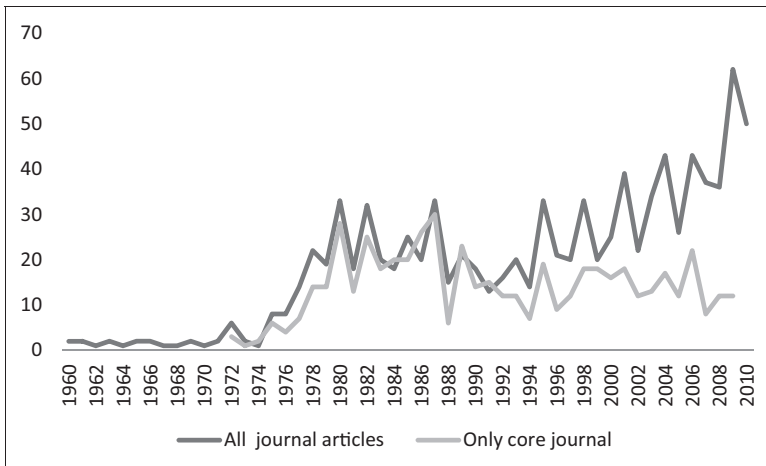


Figure 1. Articles with title words implementation or implementing published in core journals (political science, public policy and public administration/management) by year of publication. Absolute numbers per year.

Table 1. Theory relevant features of articles in core journals by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|---------------------------------|-----------|-----------|-----------|-----------|-------|
| Some theoretical references (%) | 60 | 85 | 88 | 87 | 84 |
| Hypotheses formulation (%) | 7 | 24 | 30 | 40 | 28 |
| Some empirical data (%) | 56 | 69 | 78 | 82 | 74 |
| (N)=100% | (57) | (210) | (138) | (142) | (547) |

Percentage base: all articles. N=547.

Table 2. Level of empirical analysis in core journal articles by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|---------------------------------------|-----------|-----------|-----------|-----------|-------|
| Some form of hypotheses testing (%) | 10 | 33 | 38 | 46 | 36 |
| Some statistical analysis (%) | 3 | 17 | 29 | 24 | 22 |
| Correlation, regression and ANOVA (%) | 3 | 15 | 28 | 24 | 20 |
| (N)=100% | (32) | (144) | (108) | (117) | (401) |

Percentage base: all articles with some empirical data.

To sum up here, the first generation idea must be back dated at least a decade. Furthermore, it seems quite appropriate and correct to refer to what happened approximately before and after 1980 as the transition from a *first* to a *second generation* of implementation research. This can probably best be explained by the sudden growth of various analytical and theoretical frameworks during a relatively short time period from the mid-late 1970s to the early 1980s. Many of these and subsequent books were efforts to offer a first general analytical and theoretical framework for implementation research (cf. Hill and Hupe, 2009). That watershed event offered implementation scholars of the latter decade opportunities to pursue hypotheses in a more deductive manner that required other and more sophisticated research designs than those appropriate for inductive research approaches.

What about the third generation research paradigm?

Let us start with the area of *research methodology* where progress allegedly is most noticeable. There appears to be substantial validity to this claim. While the frequency of theoretical references in journal articles after 1990 is at about the same level as in the 1980 the likelihood to derive hypotheses from theoretical constructs and subject them to some sort of empirical testing has increased substantially from only 10% before 1980 to almost 50% in the 2000s. More important is the fact that the tendency to use of more sophisticated statistical techniques to analyse empirical data has increased even more during the same time period from 3% to 24% (see Tables 1 and 2).

Increased use of *quantitative* data to supplement qualitative data has been an indispensable prerequisite for the positive trend with respect to formulating explicit hypotheses (see Table 7) and applying more sophisticated statistical analytical techniques reported above (Table 2). Thus, not surprisingly, this has in fact happened, even though the change here was most pronounced from the 1970s to the 1980s. There was nevertheless further significant change in the same direction during the 1990s but with a slight reversal during the 2000s (Table 3).

The latter mentioned trend touches on the familiar discussion of the relative merits of *qualitative* versus *quantitative* research methodologies and is unavoidable in the assessment of scientific progress in implementation research. Our comments so far have been premised on the common assumption that certain research design

Table 3. Type of data in core journal articles by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|---------------------------|-----------|-----------|-----------|-----------|-------|
| Only qualitative data (%) | 88 | 65 | 55 | 60 | 62 |
| Only quantitative (%) | 9 | 21 | 28 | 26 | 23 |
| Mixed method (%) | 3 | 14 | 17 | 14 | 14 |
| (N)=100% | (32) | (144) | (108) | (117) | (401) |

Percentage base: empirical articles.

Table 4. Research design of articles in core journal by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|-------------------------|-----------|-----------|-----------|-----------|-------|
| Single case-study (%) | 72 | 42 | 36 | 39 | 41 |
| Comparative-studies (%) | 25 | 50 | 53 | 53 | 50 |
| Longitudinal (%) | 22 | 17 | 22 | 16 | 19 |
| (N)=100% | (32) | (144) | (108) | (117) | (401) |

Percentage base: empirical articles.

features (e.g. comparative case-studies) are more valuable with respect to theory development than others (e.g. single case-studies). We generally believe that to be true, but not absolutely. Thus, the research methodology literature recognizes that the theoretical utility of even single case-studies can be substantially enhanced by careful selection of *crucial* or *critical cases* in terms of testing a particular theory (Eckstein, 1975; Gerring, 2007; King et al., 1994; Stinchcombe, 1968; Yin, 2009). Unfortunately, single case-studies of this kind are rare in implementation studies as elsewhere. For a few rare exceptions see Lundquist (2001) and Lundin (2007).

The ideal, of course, would be for implementation studies to combine the use of qualitative and quantitative methodologies. That is, unfortunately, still not the case. The more even distribution of these two research methodologies is not a feature of individual implementation studies but rather of the research literature as a whole. Nevertheless, we consider this more balanced relationship between the two major research methodologies at the aggregate level a substantial improvement compared to the earlier predominant position of one over the other.

The *comparative* research design is an important defining feature of the third generation research paradigm. *Cross-national comparative* studies are particularly valuable in the sense that they provide the broadest range of cultural and political contexts. This means greater variation in a wider set of independent variables that is crucial to general theory development. Hence, it is disappointing that especially cross-national comparison after 1990 is only marginally more common in implementation studies than it was before. Nevertheless, there are several examples of well-crafted comparative studies of policy implementation (e.g. Falkner et al., 2005,

Table 5. Various types of comparative research design by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|--------------------------------------------|-----------|-----------|-----------|-----------|-------|
| Cross-national comparison (%) | 3 | 9 | 12 | 11 | 10 |
| Cross-state/province comparison (%) | 6 | 22 | 20 | 16 | 20 |
| Cross-local gov. comparison (%) | 6 | 9 | 10 | 10 | 10 |
| Cross policy/program sector comparison (%) | 9 | 13 | 7 | 13 | 11 |
| (N)=100% | (32) | (144) | (108) | (117) | (401) |

Percentage base: all empirical articles.

2008; Goggin, 1987; Knill, 1998; Knill and Lenschow, 1998; May, 1995; Siedentopf and Ziller, 1988). Comparison at the *subsystem* level is more common than cross-national comparison as can be gleaned from Table 5. These studies can also be quite valuable though the *units of analysis* here will vary a great deal depending on whether it takes place in a *federal* or *unitary* national political system. *Cross-state comparison* in this context refers almost exclusively to studies of policy implementation at state level in the US.

This very slow almost non-noticeable progress in application of the comparative research design since the 1980s is disappointing. Nevertheless, it is no small achievement that slightly more than 50% of policy implementation studies nowadays have a comparative approach. Whether this is good enough in terms of the third generation research paradigm is a matter of judgement. Comparison per se is no panacea. It is well known that the usefulness of comparative studies in answering research questions depends critically on careful selection of cases in terms of substantive comparability as well as selection of variables, dependent as well as independent, of theoretical interest (Ragin, 1987). Unfortunately, these requirements often receive insufficient attention in many comparative studies, rendering them less useful than they otherwise could have been.

The *longitudinal* research design is another defining feature of the third generation research design. It was thought of as an antidote to the cross-sectional research instruments like questionnaires and surveys, usually with no inherent time and process dimension in the data gathered. The main trend here is as disappointing as with respect to the comparative research design: no increase in its application after 1990.

Why is this so? We will offer an answer below.

The issue of theoretical progress or rather lack thereof in the implementation literature has been dealt with thoroughly by a number of implementation scholars in recent years (Hill and Hupe, 2009; Matland, 1995; O'Toole, 2000, 2004; Winter, 2012b). There is broad consensus behind the view that no general implementation theory is close at hand. The third generation research design paradigm does not privilege one particular theoretical approach at the expense of others. But it does imply that a theoretical explanatory orientation towards analysis of policy

implementation is preferable to a-theoretical and merely descriptive approaches. Hence, we have limited our assessment here to register what type of theoretical constructs have been in vogue during the last decade or two with a view of commenting on differences and similarities to previous decades.

O'Toole (2000: 273–282) in his assessment identified a handful of broadly defined theoretical approaches as particularly promising for future implementation scholarship: (1) rational choice institutionalism, (2) the governance approach, (3) networks and management, (4) formal rational choice models and (5) policy design and instrument approach. How do these predictions hold up to our empirical results?

We have merged the *institutional* approach, which is the first discussed by O'Toole, with another he does not mention – some forms of political and policy *regime* theory (Stoker, 1989) – though not always called so by their authors. The main reason for this is that both approaches tend to refer to some macro features of the political-administrative units of analysis. Institutional approaches focusing more on the meso-level of the political-administrative system and their constituent organizations we have grouped under the label organization theory as they represent new and old versions of the latter type of theory (March and Olsen, 1984; Ostrom, 1999; Powell and DiMaggio, 1991; Selznick, 1957). We have also merged the networks and governance approach, which O'Toole discuss separately, with a one-sided bottom-up approach, partly because of their emphasis on inter-organizational actor networks and partly because of pragmatic reasons related to relative small n.

Table 15 reports our results with respect to theoretical approaches discussed more generally in e.g. non-empirical articles and Table 16 where theoretical constructs are subject to some sort of empirical testing. The macro-level *institutional* and *regime* approach is the most prominent theoretical framework mentioned more generally and subjected to empirical testing, both overall and especially during the last two decades. The *networks-governance* approach has experienced the greatest increase in popularity since 1990 in the latter respect sharing a second place ranking (Table 16). *Organization theory* was more prominent in the implementation literature during the 1970s and 1980s but has receded substantially as a preferred theoretical framework since then. The *policy design* approach was initially much mentioned in the implementation literature at large (Table 15) but less so after 1990. However, in the literature where efforts are made to confront theoretical constructs with some empirical data it is almost untraceable. Hence, it seems to have suffered the same decline in popularity as the organizational approach (Table 16).

The prominent role of macro-institutional and governance-network theories in the implementation literature after 1990 clearly reflect a broader trend evident in political science. Thus, O'Toole (2000) seems to have been on target with at least three of his five favourite theories but with one important qualification. It is *not* his preferred rational choice type of institutionalism that implementation scholars refer to and use. Rather it is another less rational version expounded by scholars

like e.g. March and Olsen (1984), Hall (1986) and Steinmo et al. (1992). In the same vein the use of more formal, deductive rational choice models, which O'Toole seems to sympathize with as a potential promising parsimonious explanatory framework is scant or almost non-existent in our sample of the implementation literature.

As a supplement to the foregoing analysis we have looked at to what extent one of the most clearly operationalized *top-down* analytical frameworks formulated by Mazmanian and Sabatier in various versions (1981, 1983) have been actively used by implementation scholars in our sample. The change from playing a dominant role before 1990 to being one among several, and not the leading anymore after 1990 is an interesting development. The latter is consistent with another observation made by O'Toole (2000) and Winter (2012b) namely the gradual silencing of the protracted and unproductive debate between top-down and bottom-up scholars that plagued much of the 1980s as the interest of implementation researchers have shifted towards integrating and synthesizing analytical and theoretical frameworks offered by several scholars during the last two decades (Linder and Peters, 1987, 1990; Matland, 1995; Ryan, 1996; Sabatier, 1986; Winter, 1990, 2012a).

Considering the greater interest in integrative and synthesizing frameworks, it is something of a paradox that we a good decade into the 21st century observe more major contending theoretical frameworks rather than fewer (Table 16). There is no agreement on whether this constitutes progress or not. O'Toole (2000: 268) seems to deplore the multitude of different theoretical constructs in implementation research as he observes that 'what has not happened is a careful winnowing of the mass of potential explanatory variables towards parsimonious explanation'. Winter (2012b: 265) on the other hand in his recent review of the literature argues that it is premature to expect any scholarly agreement on a common analytical and theoretical framework and that diversity in this respect is more a strength than a weakness. Thus, he suggests that implementation research can be improved by '(1) accepting theoretical diversity rather than looking for one common theoretical framework and (2) developing and testing partial theories and hypothesis rather than trying to reach for utopia in constructing a general implementation theory'.

Regardless of these different opinions by two leading contemporary implementation scholars we choose to interpret our empirical observations as consistent with the claim, which stated that there has been more progress on methodological than theoretical issues.

The trade-offs that we suspected researchers must make between the various demanding features of the third generation research design paradigm are evident in our empirical findings. Some of these trade-offs may be of a positive nature while others are less so. An example of the former is the relationship between *qualitative* versus *quantitative* data on one hand and *single case-studies* versus *comparative studies* on the other hand (Table 6) where the latter use quantitative data to a much larger extent (51%) than the former where qualitative data dominate even more (92%). This means that the trend from single case to comparative research

Table 6. Type of research design and type of data used

| Type of data | Type of research design | |
|----------------------------|-------------------------|---------------------|
| | Single case-study | Comparative studies |
| Qualitative data only (%) | 94 | 48 |
| Mixed method (%) | 6 | 21 |
| Quantitative data only (%) | – | 31 |
| Longitudinal data (%) | 22 | 14 |
| (N)=100% | (153) | (200) |

Percentage base: all empirical articles of respective type.

Table 7. Hypotheses formulation by type of data used

| | Type of data used in study | | |
|----------------------------|----------------------------|---------------------------|------------------------|
| | Qualitative data only | Quantitative data as well | Only quantitative data |
| Hypotheses formulation (%) | 13 | 34 | 50 |
| (N)=100% | (249) | (58) | (93) |

Percentage base: empirical articles.

designs takes place at the expense of the dominating role of qualitative research methodology. This trend is consistent with the assumption formulated by Hill and Hupe (2009: 156–160) that quantitative data and methodologies are more likely to be used in studies comparing multiple organizations than those only focusing on single organizations. A less positive trade-off is revealed in the fact that the longitudinal research design is used more frequently in single case-studies than comparative studies (see Table 6). Consequently, as the use of the comparative research design increases over time this happens at the expense of the longitudinal research design. This validates the claim by O'Toole (2000: 269) that larger n investigations predominantly use a cross-sectional research design.

The external validity of social science research is always contingent on the context in which it is conducted. This is of course also the case in research on policy implementation and with respect to e.g. the *regional context* in which it is conducted. O'Toole (2000: 271) makes the observation that most implementation research at the turn of this century was still being conducted in the US. As Table 8 shows that is still true but less so during the last decade than earlier. The North-American dominance in implementation scholarship has been decreasing continuously and significantly during every decade since the 1970. Scholars

Table 8. Regional origin of authors or their studies in core journal by time periods

| | 1953–1979 | 1980–1989 | 1990–1999 | 2000–2009 | Total |
|-------------------|-----------|-----------|-----------|-----------|-------|
| North America (%) | 79 | 69 | 63 | 58 | 65 |
| Europe (%) | 18 | 23 | 22 | 35 | 25 |
| Other regions (%) | 9 | 9 | 21 | 13 | 13 |
| (N)=100% | (57) | (210) | (138) | (142) | (557) |

Percentages add to more than 100% since some articles have authors and or data from more than one region. Percentage base: all articles.

Table 9. Research design of articles in core journals by region of focus or author

| | North America | Europe | Other regions |
|-------------------------|---------------|--------|---------------|
| Single case-study (%) | 36 | 43 | 60 |
| Comparative-studies (%) | 52 | 50 | 37 |
| Longitudinal (%) | 18 | 16 | 22 |
| (N)=100% | (257) | (109) | (46) |

Percentage base: all empirical articles.

from other regions and particularly Europe have increased their share correspondingly and equally steadfast. This implies that the regional and cultural context of implementation research has broadened substantially and consistently since its infancy. What are the implications of this development?

O'Toole (2000) criticized his fellow North Americans – who have dominated this research field – for their tendency to neglect the works of European scholars. He also suggested that much could be learned from a trans-Atlantic focus and dialogue. To assess how much merit there is to this claim we have added a regional variable to our previous analysis (see Tables 9 through 14). This brings out some interesting regional differences in scholarly choices of research designs. Most conspicuous in this respect is the relative strong preference among European implementation researchers for *qualitative* data and their lesser concern for developing and especially testing *hypotheses* derived from some theoretical constructs. In contrast, North American scholars are more proficient at formulating explicit theoretical propositions that are tested using *quantitative* data and advanced statistical techniques like ANOVA, factor analysis and multiple regression analysis (Tables 13 and 14, respectively). The single case-study research design is somewhat more commonly used in Europe and even more so in other regions than the US while regional differences with respect to employing a *comparative* and *longitudinal research design* are negligible (Table 9). This regional difference in preferred

Table 10. Different comparative research design in core journal articles by region of focus or author

| | North America | Europe | Other regions |
|--------------------------------------------|---------------|--------|---------------|
| Cross-national comparison (%) | 7 | 23 | 13 |
| Cross-subsystem comparison (%) | 35 | 14 | 15 |
| Cross policy sector/program comparison (%) | 11 | 12 | 9 |
| Other types of comparison (%) | 8 | 6 | 2 |
| (N)=100% | (257) | (109) | (46) |

Percentage base: all empirical articles.

Table 11. Different comparative research design in core journal articles by region of focus or author

| | North America | Europe | Other regions |
|--------------------------------------------|---------------|--------|---------------|
| Cross-national comparison (%) | 13 | 46 | 35 |
| Cross-state/region province comparison (%) | 51 | 7 | 17 |
| Cross-local govt. comparison (%) | 16 | 20 | 24 |
| Cross policy sector/program comparison (%) | 22 | 24 | 24 |
| Other types of comparison (%) | 8 | 5 | 0 |
| (N)=100% | (135) | (55) | (17) |

Percentage base: comparative studies. Since categories are not mutually exclusive percentages add to more than 100.

Table 12. Type of data used in studies by region of focus

| | North America | Europe | Other regions |
|----------------------------|---------------|--------|---------------|
| Only qualitative data (%) | 53 | 80 | 83 |
| Only quantitative data (%) | 31 | 10 | 6 |
| Mixed method (%) | 16 | 12 | 13 |
| (N)=100% | (257) | (109) | (46) |

Percentage base: empirical articles.

research design and methods is consistent with Hill and Hupe (2009: 160) when they marvel at how much bolder American scholars have been in the use of quantitative methods in the study of policy implementation than their European colleagues.

Table 13. Types of data used in comparative studies by region of focus

| | North America | Europe | Other regions |
|----------------------------|---------------|--------|---------------|
| Only qualitative data (%) | 37 | 75 | 77 |
| Only quantitative data (%) | 41 | 9 | 5 |
| Mixed method (%) | 22 | 16 | 18 |
| (N)=100% | (136) | (55) | (17) |

Percentage base: articles with comparative focus.

Table 14. Theory relevant features of core journal article by region of focus or author

| | North America | Europe | Other regions |
|----------------------------------------------|---------------|--------|---------------|
| Some theoretical references ^a (%) | 84 | 86 | 84 |
| Hypotheses formulation ^a (%) | 31 | 23 | 18 |
| Hypotheses testing ^b (%) | 42 | 30 | 22 |
| Some statistical analysis ^b (%) | 30 | 8 | 7 |
| Regression and ANOVA ^b (%) | 28 | 8 | 4 |
| (N)= 100% ^a | (360) | (138) | (62) |
| (N)=100% ^b | (257) | (109) | (46) |

^aPercentage bases of all articles.

^bPercentage bases of only those empirical.

Table 15. Types of theories referenced in core journal articles by time periods

| | 1953–1989 | 1990–2009 | Total |
|-----------------------------------------------------------------|-----------|-----------|-------|
| Political/policy regime and institutional perspective (%) | 16 | 22 | 19 |
| Policy formulation (type, content, design, instrumentation) (%) | 22 | 12 | 16 |
| Organization theory (%) | 12 | 7 | 9 |
| Networks and governance theory (%) | 8 | 10 | 9 |
| Mazmanian and Sabatier 1979, 1980, 1983 (%) | 32 | 13 | 19 |
| Van Meter and Van Horn 1975 (%) | 4 | 2 | 4 |
| (N) = 100% | (213) | (245) | (459) |

Percentage base: articles with some theoretical references.

Another noticeable difference is the much larger share of comparative studies that are cross-national in Europe (46%) compared to those in the US. In the US, however, this is to some extent compensated for by an equally large share (51%) being comparative across states (Table 10). This reflects, of course, the different political organization of the two continents being compared.

Table 16. Types of theories or analytical frameworks most frequently subjected to some form of empirical testing (i.e. both statistical and non-statistical)

| | 1953–1989 | 1990–2009 | Total |
|-----------------------------------------------------------|-----------|-----------|-------|
| Political/policy regime and institutional perspective (%) | 26 | 36 | 32 |
| Organization theory (%) | 18 | 13 | 15 |
| Networks and governance theory (%) | 4 | 13 | 10 |
| Mazmanian and Sabatier 1979, 1980, 1983 (%) | 32 | 13 | 19 |
| Van Meter and Van Horn 1975 (%) | 4 | 2 | 4 |
| (N) = 100% | (51) | (95) | (146) |

Percentage basis: all articles with some form of empirical testing.

Conclusion

O'Toole (2000: 283) concluded that 'the study of policy implementation is in many respects in a relatively *mature* stage of development' (emphasis added). Among the most important indicators in this respect he mentions the fact that inductive approaches have been increasingly supplanted by more deductive hypotheses testing approaches. As already noted above this is the most consistent positive trend in our investigation of the sampled implementation literature towards the requirements of the third generation research paradigm from its beginning that has lasted well into the 2000s. It is also consistent with recommendation by Winter (2012b: 265) about the best way to enhance theory development in implementation research: to test partial theories and related hypotheses. With few exceptions most of these achievements were made already in the 1980s. Progress since then has been much slower and incremental. Thus, our analysis confirms the conventional first- and second generation description of the evolution of implementation research up to the 1990s as reasonably accurate. The slow and incremental advances made since then bolster the impression of several reviewers in later years that many important defining features of the third generation of research are still far away from being implemented. The explanation for this implementation deficit is as suggested by way of introduction to a large extent due to the very demanding nature of the third generation research design and, as we discovered, some inherent dualities and tensions between some of its essential features that make it hard to optimize them all simultaneously. Nevertheless, this has not made the third generation research paradigm useless and irrelevant for analytical purposes as we have used it as a benchmark to measure more or less progress towards it as an ideal typical analytical construct on some of its salient features.

Our main conclusion is on the whole not much different from those of the other most recent reviewers. We have confirmed to a large extent their more or less intuitive based but still quite accurate assessments. What distinguishes ours from theirs lies primarily in documenting and providing more detailed and nuanced

accounts of allegedly mixed results through an *empirical and up-to-date investigation* of the research literature. Thus, we agree with previous reviewers like O'Toole (2000) and Winter (2012b) that much has been achieved but also that a lot remains to be done. Hopefully, our empirical investigation has helped provide more detailed information on both accounts.

References

- Barrett SM (2004) Implementation studies: Time for a revival? Personal reflections on 20 years of implementation studies. *Public Administration* 82(2): 249–262.
- Barrett S and Fudge C (1981) *Policy and Action*. London: Methuen.
- Bunker DR (1972) Policy sciences perspectives on implementation processes. *Policy Sciences* 3(1): 71–80.
- deLeon P and deLeon L (2002) Whatever happened to policy implementation? An alternative approach. *Journal of Public Administration Research and Theory* 12(4): 467–488.
- Eckstein H (1975) Case study and theory in political science. In: Greenstein FI and Polsby NW (eds), *Handbook of Political Science. Political Science: Scope and Theory*, vol. 7. Reading, MA: Addison-Wesley, pp.94–137.
- Falkner G, Treib O, Hartlapp M, et al. (2005) *Complying with Europe. EU Harmonization and Soft Law in the Member States*. Cambridge: Cambridge University Press.
- Falkner G, Treib O and Holzleithner E (2008) *Compliance in the Enlarged European Union. Living Rights or Dead Letters?* Aldershot: Ashgate.
- Gerring J (2007) Is there a (viable) crucial case method? *Comparative Political Studies* 40(3): 231–253.
- Goggin ML (1986) The too few cases too many variables problem in implementation research. *Western Political Quarterly* 39(2): 328–347.
- Goggin ML (1987) *Policy Design and the Politics of Implementation*. Knoxville: University of Tennessee Press.
- Goggin ML, Bowman AOM, Lester JP, et al. (1990) *Implementation Theory and Practice: Toward a Third Generation*. Glenwood, Illinois: Scott Foreman/Little Brown.
- Hall PA (1986) *Governing the Economy: The Politics of State Intervention in Britain and France*. Cambridge: Policy Press.
- Hill M and Hupe P (2009) *Implementing Public Policy: Governance in Theory and in Practice*. London/Thousand Oaks/New Delhi: Sage.
- King G, Keohane RO and Verba S (1994) *Designing Social Inquiry*. Princeton, NJ: Princeton University Press.
- Knill C (1998) European policies. The impact of national administrative traditions. *Journal of Public Policy* 18(1): 1–28.
- Knill C and Lenchow A (1998) Coping with Europe. The impact of British and German administration on the implementation of EU environmental policy. *Journal of European Public Policy* 5(4): 595–614.
- Lester JP, Bowman AOM, Goggin ML, et al. (1987) Public policy implementation: Evolution of the field and agenda for future research. *Policy Studies Review* 7(1): 200–216.
- Linder SH and Peters BG (1987) A design perspective on policy implementation: The fallacies of misplaced prescription. *Policy Studies Review* 6(3): 459–475.
- Linder SH and Peters BG (1990) Research perspective on the design of public policy: Implementation, formulation and design. In: Palumbo DJ and Calista DJ (eds)

- Implementation and the Policy Process: Opening Up the Black Box*. New York: Greenwood Press.
- Lundin M (2007) When does cooperation improve public policy implementation? *Policy Studies Journal* 35(4): 629–652.
- Lundquist LJ (2001) Implementation from above. The ecology of power in Sweden's environmental policy. *Governance* 14(3): 319–337.
- March JG and Olsen JP (1984) The new institutionalism: Organizational factors in political life. *American Political Science Review* 78(3): 734–749.
- Matland RE (1995) Synthesizing the implementation literature: The ambiguity-conflict model. *Journal of Public Administration Research and Theory* 5(2): 145–174.
- May PJ (1995) Can cooperation be mandated? Implementing intergovernmental environmental management in New South-Wales and New Zealand. *Publius* 25(1): 89–113.
- Mazmanian DA and Sabatier P (1981) *Effective Policy Implementation*. Lexington, Mass: Lexington Books.
- Mazmanian DA and Sabatier P (1983) *Implementation and Public Policy*. Glenview, Illinois: Scott, Foresman.
- Ostrom E (1999) Institutional rational choice: An assessment of the institutional analysis and development framework. In: Sabatier PA (ed.) *Theories of the Policy Process*. Boulder, Colorado: Westview. pp.35–71.
- O'Toole LJ Jr (1986) Policy recommendations for multi-actor implementation: An assessment of the field. *Journal of Public Policy* 6(2): 181–210.
- O'Toole LJ Jr (2000) Research on policy implementation: Assessment and prospects. *Journal of Public Administration Research and Theory* 10(2): 263–288.
- O'Toole LJ Jr (2004) The theory-practice issue in policy implementation research. *Public Administration* 82(2): 309–329.
- Powell WW and DiMaggio PJ (1991) *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Pressman JL and Wildavsky AB (1973) *Implementation. How Great Expectations in Washington are Dashed in Oakland*. Berkeley: University of California Press.
- Ragin CC (1987) *The Comparative Method*. Berkeley: University of California Press.
- Ryan N (1996) Some advantages of an integrated approach to implementation analysis: A study of Australian industry policy. *Public Administration* 74: 737–753.
- Sabatier PA and Mazmanian DA (1979) The conditions of effective implementation. A guide to accomplishing policy objectives. *Policy Analysis* 5(4): 481–504.
- Sabatier PA and Mazmanian DA (1980) The implementation of public policy. A framework of analysis. *Policy Studies Journal* 8(4): 538–560.
- Sabatier PA (1986) Top-down and bottom-up approaches to implementation research. A critical analysis and suggested synthesis. *Journal of Public Policy* 6(1): 21–48.
- Saetren H (2005) Facts and myths about research on public policy implementation: Out-of-fashion, allegedly dead, but still alive and relevant. *Policy Studies Journal* 33(4): 559–582.
- Schofield J (2001) Time for a revival? Public policy implementation: A review of the literature and agenda for future research. *International Journal of Management Review* 3(3): 245–263.
- Schofield J and Sausman C (2004) Symposium on implementing public policy: Learning from theory and practice: Introduction. *Public Administration* 82(2): 235–248.
- Selznick P (1957) *Leadership in Administration*. New York: Harper & Row.
- Siedentopf H and Ziller J (1988) *Making European Policies Work: The Implementation of Community Legislation in the Member States*. London: Sage.

- Sinclair TAP (2001) Implementation theory and practice: Uncovering policy and administration linkages in the 1990's. *International Journal of Public Administration* 24(1): 77–94.
- Smith TB (1973) The study of policy making in developing countries. *Policy Studies Journal* 1(4): 244–249.
- Steinmo S, Thelen K and Longstreeth F (1992) *Structuring Politics: Historical Institutionalism in Comparative Analysis*. Cambridge: Cambridge University Press.
- Stinchcombe AL (1968) *Constructing Social Theories*. NY: Harcourt, Brace & World.
- Stoker RP (1989) Regime framework for implementation analysis. *Policy Studies Review* 9(1): 29–49.
- Van Meter DS and Van Horn CE (1975) The policy implementation process. *Administration & Society* 6(4): 445–488.
- Winter S (1990) Integrating implementation research. In: Palumbo DJ and Calista DJ (eds) *Implementation and the Policy Process: Opening Up the Black Box*. New York: Greenwood Press, pp.19–38.
- Winter S (2012a) Implementation perspectives: Status and reconsideration. Chapt.16. In: Peters BG and Pierre J (eds) *Handbook of Public Administration*. London: Sage, pp.255–263.
- Winter S (2012b) Implementation. In: Peters BG and Pierre J (eds) *Handbook of Public Policy*. London: Sage, pp.151–166.
- Yin RK (2009) *Case Study Research: Design and Methods*. London: Sage.

Appendix I. Journals classified as belonging to the Core. Number of articles in each journal in parentheses

Policy Journals:

- American Journal of Evaluation (1)
- Canadian Public Policy (4)
- Cato (3)
- Evaluation Review (16)
- Journal of European Public Policy (8)
- Journal of Policy Analysis and Management (11)
- Journal of Public Policy (14)
- Local Government Policy Making (3)
- Policy Analysis (7)
- Policy Currents (7)
- Policy Evaluation (1)
- Policy & Politics (25)
- Policy Sciences (17)
- Policy Studies (2)
- Policy Studies Annual Review (1)
- Policy Studies Journal / (57)
- Policy Studies Review (43) later renamed:
- Review of Policy Research (4)
- Public Policy (9)

Public Policy and Administration (1)

Public Administration/Management Journals:

Administration and Society (27)

American Review of Public Administration (13)

Australian Journal of Public Administration (12)

Canadian Public Administration (8)

Indian Journal of Public Administration (8)

International Journal of Public Administration (25)

International Public Management Journal (3)

International Review of Administrative Science (11)

Journal of Public Administration Research and Theory (20)

Philippine Journal of Public Administration (3)

Public Administration (23)

Public Administration Quarterly (1)

Public Administration Review (48)

Public Management Review (3)

Sage Public Administration Abstracts (1)

Transylvanian Review of Administrative Sciences (1)

Political Science Journals:

Acta Politica (1)

American Journal of Political Science (3)

American Political Science Review (3)

American Politics Quarterly (2)

Annals of the American Academy of Political Science (2)

British Journal of Politics & International Relations (1)

Canadian Journal of Political Science (1)

Comparative Political Studies (3)

Comparative Politics (4)

European Journal of Political Research (10)

European Political Science (1)

European Union Politics (1)

Governance (4)

International Organization (2)

International Political Science Review (1)

Journal of Commonwealth Comparative Politics (1)

Journal of Political Economy (1)

Journal of Political Science (1)

Journal of Politics (8)

Journal of Theoretical Politics (1)

Political Quarterly (1)

Political Research Quarterly (2)

Political Science Quarterly (3)

Political Studies (2)

Politeia (2)

Polity (2)
Proceedings of the Academy of Political Science (1)
PS – Political Science & Politics (1)
Publius (16)
Scandinavian Political Studies (8)
Schweizerische Zeitschrift für Politikwissenschaft (1)
Statsvetenskapleg Tidsskrift (1)
Tulane Studies in Political Science (1)
West European Politics (3)
Western European Quarterly (9)
World Politics (1)