CROSS-NATIONALLY COMPARATIVE, EVIDENCE-BASED EDUCATIONAL POLICYMAKING AND REFORM

INTERNATIONAL PERSPECTIVES ON EDUCATION AND SOCIETY

Series Editor: Alexander W. Wiseman

Recent Volumes:

Volume 11: Educational Leadership: Global Contexts and International Comparisons
Volume 12: International Educational Governance
Volume 13: The Impact of International Achievement Studies on National Education Policymaking
Volume 14: Post-Socialism Is Not Dead: (Re)Reading The Global in Comparative Education
Volume 15: The Impact and Transformation of Education Policy in China
Volume 16: Education Strategy in the Developing World: Revising the World Bank's Education Policy
Volume 17: Community Colleges Worldwide: Investigating the Global Phenomenon
Volume 18: The Impact of HIV/AIDS on Education Worldwide
Volume 19: Teacher Reforms around the World: Implementations and Outcomes
Volume 20: Annual Review of Comparative and International Education 2013
Volume 21: The Development of Higher Education in Africa: Prospects and Challenges
Volume 22: Out of the Shadows: The Global Intensification of Supplementary Education
Volume 23: International Educational Innovation and Public Sector Entrepreneurship
Volume 24: Education for a Knowledge Society in Arabian Gulf Countries
Volume 25: Annual Review of Comparative and International Education 2014
Volume 26: Comparative Sciences: Interdisciplinary Approaches
Volume 27: Promoting and Sustaining a Quality Teacher Workforce Worldwide
Volume 28: Annual Review of Comparative and International Education 2015
Volume 29: Post-Education-for-All and Sustainable Development Paradigm: Structural Changes with Diversifying Actors and Norms
Volume 30: Annual Review of Comparative and International Education 2016
Volume 31: The Impact of the OECD on Education Worldwide
Volume 32: Work-Integrated Learning in the 21st Century: Global Perspectives on the Future
Volume 33: The Century of Science: The Global Triumph of the Research University
Volume 34: Annual Review of Comparative and International Education 2017

INTERNATIONAL PERSPECTIVES ON EDUCATION AND SOCIETY VOLUME 35

CROSS-NATIONALLY COMPARATIVE, EVIDENCE-BASED EDUCATIONAL POLICYMAKING AND REFORM

EDITED BY

ALEXANDER W. WISEMAN
Lehigh University, Bethlehem, PA, USA

PETRINA M. DAVIDSON
Lehigh University, Bethlehem, PA, USA

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2018

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78743-768-5 (Print)
ISBN: 978-1-78743-767-8 (Online)
ISBN: 978-1-78743-972-6 (Epub)

ISSN: 1479-3679 (Series)

ISOQAR certified Management System, awarded to Emerald for adherence to Environmental standard ISO 14001:2004.

Certificate Number 1985
ISO 14001

INVESTOR IN PEOPLE

CONTENTS

PART II
IMPLEMENTING NATIONAL EDUCATION REFORM

PART III
EVALUATING NATIONAL EDUCATION REFORM

This page intentionally left blank

LIST OF CONTRIBUTORS

Abdullah Ahmed Al-Ghamdi	King Abdulaziz University, Jeddah, Saudi Arabia
Asila Al-Ma'awali	Sultan Qaboos University, Muscat, Oman
Munira Al-Wahaibi	Sultan Qaboos University, Muscat, Oman
Emily Anderson	Centenary University, Hackettstown, NJ, USA
Petrina M. Davidson	Lehigh University, Bethlehem, PA, USA
Nino Dzotsenidze	Lehigh University, Bethlehem, PA, USA
Tariq Elyas	King Abdulaziz University, Jeddah, Saudi Arabia
Mobarak Hossain	Paris School of International Affairs, Sciences Po Paris, Paris, France
Johanna Kallo	University of Turku, Turku, Finland
Jaakko Kauko	University of Tampere, Tampere, Finland
Daniel Kratochvil	Abu Dhabi School of Management, Abu Dhabi, UAE
Mariam Orkodashvili	Pennsylvania State University, University Park, PA, USA
Risto Rinne	University of Turku, Turku, Finland
Azza Shafei	Université de Genève, Geneva, Switzerland
Sajida Hasan Shroff	Altamont Group, Dubai, UAE
Tuomas Takala	University of Tampere, Tampere, Finland
Tshegofatso D. Thulare	Lehigh University, Bethlehem, PA, USA
Alexander W. Wiseman	Lehigh University, Bethlehem, PA, USA
Mike Zapp	University of Luxembourg, Luxembourg

This page intentionally left blank

ABOUT THE VOLUME EDITORS

Alexander W. Wiseman is Professor of Comparative and International Education at Lehigh University, USA. Dr. Wiseman holds a dual-degree Ph.D. in Comparative & International Education and Educational Theory & Policy from Pennsylvania State University and M.A. in International Comparative Education from Stanford University. Dr. Wiseman conducts internationally comparative educational research using large-scale education datasets on math and science education, information and communication technology, teacher preparation, professional development and curriculum as well as school principal's instructional leadership activity.

Petrina M. Davidson holds a Ph.D. in Comparative and International Education from Lehigh University, a M.S. in Teaching, Learning, and Leadership with an emphasis on Curriculum and Instruction from Oklahoma State University, and B.A. in English and Education from The University of Tulsa. Dr. Davidson taught high school English and served as the secondary English curriculum coordinator in one of the largest school districts in Oklahoma. Her research interests include the institutionalization of education, curriculum in post-conflict societies, education for refugees and marginalized populations, and measures of teacher quality.

This page intentionally left blank

PREFACE

Given the importance of evidence-based educational policymaking and reform, and the contribution that large-scale datasets make to evidence-based decisions in national educational systems worldwide, the goal of this volume in the *International Perspectives on Education and Society* (IPES) series is to describe, synthesize, and forecast how large-scale assessments and quantitative data impact evidence-based policymaking derived in part from policymaking examples from national educational systems and international organizations. The volume provides a forum for scholars and policymakers to identify how large-scale assessments and quantitative data can be used to inform policymaking at all levels of education, and how these data can be used to better understand specific country- and regional-level educational challenges.

The main question this volume addresses asks how and why large-scale educational datasets impact evidence-based educational policymaking, and how that is empirically observed in national educational systems in several Middle East and North African countries as well as others worldwide. There are many kinds of large-scale datasets that are relevant to evidence-based educational policymaking. For example, many Arabian Gulf countries participate in several of the internationally comparative education assessments, including the Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading and Literacy Study, and the Programme for International Student Assessment. Of these large-scale assessments, the TIMSS includes representatives from all of the Gulf Cooperation Council countries as well as many from the Arab community worldwide, but Arab countries represent only a fraction of the systems that participate in large-scale international assessments, like TIMSS.

The scope of international large-scale assessments is a bit overwhelming. For example, in 2015, the number of countries participating in TIMSS was 57 countries and 7 benchmarking communities. From these 64 countries and benchmarking communities, more than 580,000 students worldwide participated by completing academic assessments and background questionnaires. In addition, each of the students' classroom teachers in the tested subject and

their school principals or head teachers responded to background questionnaires as well. The amount of data just from TIMSS 2015 alone is enormous, but we also have to remember that TIMSS has been administered every four years since 1995. In other words, 2015 marked the sixth data collection cycle for TIMSS, and TIMSS is just one of many large-scale international assessments like it. There are even more nation- or state-specific educational assessments that collected and report quantifiable data on education worldwide. In other words, the educational data available for analysis and interpretation are enormous and complex.

Chapters in this volume emphasize that quantitative research evidence is often the most legitimized among national educational policymakers and international organizations that influence national educational policymaking because it is perceived to be more accurate and trustworthy. In fact, the phrase "scientific" for international or national educational agencies often means "based in empirical research"—typically quantitative. By contrast, useful qualitative data are not often considered "scientific" enough for national or international educational funding. The measurability of achievement scores from large-scale datasets, for example, is often perceived by policymakers to be clearer and more direct than that of learning potential or the transferability of ideas that may be measured with less quantitative instruments or methods. However, chapter authors discuss how these rich data are not always used to their full potential by policymakers or educators because of the predominant focus on student achievement and ranking systems. While student achievement data can offer great insight on educational systems, the unique country-level background data available through large international datasets, for example, provide opportunities for scholars and policymakers to develop greater insight into the social and cultural factors that influence education systems around the world.

Chapters in this volume accomplish three sub-goals. First, they identify and discuss ways that a testing and assessment infrastructure contributes to improved educational planning, policy, and implementation. Second, they discuss specific ways to build capacity in relevant knowledge, skills, attitudes, and values about large-scale education data for policymaking. And, finally, they create a rough template for sustainable testing and assessment models and procedures to push national education systems forward, both regionally and globally. The volume also has two practical outcomes: short-term and long-term. The short-term outcome is that it builds a network of experts and colleagues, creates a set of policy recommendations related to large-scale education data and policymaking, and contributes to scholarship on both evidence-based educational policy and educational reform in diverse systems around the world.

Long-term outcomes of this volume are to develop a framework for sustainable collaboration and a platform for the systematic influence of education policy using large-scale assessments and quantitative data.

Alexander W. Wiseman
IPES Series Editor
Volume Co-Editor

This page intentionally left blank

CHAPTER 1

THE RHYTHMIC APPLICATION OF EVIDENCE-BASED POLICY IN NATIONAL EDUCATIONAL SYSTEMS WORLDWIDE

Alexander W. Wiseman and Petrina M. Davidson

ABSTRACT

The shift from data-informed to data-driven educational policymaking is conceptually framed by institutional and transhumanist perspectives. Examples of the shift to large-scale quantitative data driving educational decision-making suggest that data-driven educational policy will not adjust for context to the degree as done by the data-informed or data-based policymaking. Instead, the algorithmization of educational decision-making is both increasingly realizable and necessary in light of the overwhelmingly big data on education produced annually around the world. Evidence suggests that the isomorphic shift from localized data and individual decision-making about education to large-scale assessment data has changed the nature of educational decision-making and national educational policy. Big data are increasingly legitimized in educational policy communities at national and international levels, which means that algorithms are assumed

Cross-nationally Comparative, Evidence-based Educational Policymaking and Reform
International Perspectives on Education and Society, Volume 35, 1–17

ISSN: 1479-3679/doi:10.1108/S1479-367920180000035001

to be the best way to analyze and make decisions about large volumes of complex data. There is a conceptual concern, however, that decontextualized or de-humanized educational policies may have the effect of increasing student achievement, but not necessarily the translation of knowledge into economically, socially, or politically productive behavior.

Keywords: Educational policy; evidence-based policymaking; national education systems; large-scale assessment; policy borrowing; algorithmization

There is growing evidence of an established and increasingly taken-for-granted rhythmic application of large-scale educational assessment data to national educational policy agendas, reforms, and practice. Both the growth of large-scale assessment (LSA) systems and the collection and availability of big data in education have contributed to the growing influence of quantitative data on national educational policy worldwide (Best et al., 2013; Wiseman, 2010). Alongside this growing phenomenon are critiques of educational data collection and use, most of which focus on the morality of globalization and policy borrowing rather than the function of LSAs in educational reform and policymaking (Klees & Edwards, 2014; Lingard, Martino, & Rezai-Rashti, 2013). As a result, this analysis is among the first—beyond the groundbreaking policy-borrowing model developed by Phillips and Ochs (2003)—to identify, explain, and provide both explanatory and critical evidence regarding the rhythmic application of evidence-based policy in national educational systems worldwide.

Of critical importance is the subtle shift in agency and control from policymakers and educators to increasingly automated data and assessment systems. This shift from policymaker- and educator-driven decision-making to data- and assessment-driven policymaking and practice increasingly legitimizes the functions and institutionalizes the impacts of data systems independent from educational context or educator experience in systems worldwide. Although there are many potential concerns about this arrangement, one is that this legitimization and institutionalization process may replicate itself even further in national educational systems around the world until the policymakers and educators in every system are following directives independently established by the data rather than by policymakers and educators, who interpret and evaluate the data as an evidence base for wider decision-making and reform of educational policy and practice. The legitimized and increasingly global model of evidence-driven rather than evidence-based decision-making

conceptually lends itself to an institutional and increasingly transhumanist framework. Evidence-driven policymaking includes policies that begin with and are therefore driven by quantitative evidence, whereas evidence-based decision-making is spurred by a human interaction with the context and evidence to adopt the most appropriate response.

CONCEPTUAL FRAMEWORK

The conceptual framework for understanding and analyzing the rhythmic application of LSA data to educational decision-making, therefore, incorporates neo-institutional with transhumanist theoretical frameworks. The conceptual framework for identifying, explaining, and estimating the impact of the rhythmic application of evidence-based policy in national educational systems worldwide is based upon neo-institutional theory as applied to comparative and international education (Baker & Wiseman, 2006). Questions about decision-making control rest in more agency-oriented theories, which suggest that power imbalances in data-based decision-making legitimize evidence-driven rather than evidence-based educational policymaking and reform. The decision-making power shifts beyond human control to create transhuman systems based on the decision-making power of LSA data and algorithms.

Institutional Framework

The main elements of neo-institutional theory (i.e., legitimacy-seeking, scripting, isomorphism, and coupling) facilitate understanding of the phenomenon of quantitative data application and its incorporation into national and international level educational policies (DiMaggio & Powell, 1991). This neo-institutional base, in turn, serves as a foundation for the "scientization" of education (Wiseman, Damaschke-Deitrick, Bruce, Davidson, & Taylor, 2016a). Scientization of education comprises four principle components: quantification of education, commodification of education, cyborg dialectic, and education as a panacea. Of the four components of the scientization of education, the most relevant to the rhythmic application of evidence-based policy in education are "quantification" and the "cyborg dialectic." Therefore, the synthesis of neo-institutional theory and scientization suggests that the shared influence of legitimacy-seeking, isomorphism, coupling,

quantification, and the cyborg dialectic contributes to what Danaher (2016) calls "algocracy," which provides the script for the rhythmic application of data to policymaking in education worldwide. A brief explanation of each component and its contribution to the conceptual framework for this analysis is given below.

Although neo-institutional perspectives in comparative and international education research have been critiqued for their lack of attention to direct power or agency (Wiseman, Astiz, & Baker, 2014), the power to control without a direct agent is at the heart of normative isomorphism. The institutionalization of practices, norms, values, and culture creates a taken-for-granted expectation and acceptance of practices that would be actively resisted if imposed overtly or directly. Instead, a slow introduction of values, norms, and practices over time creates an opportunity for cultural norms and values to shift without drawing the attention of critics or spurring resistance. In this case, the century-long introduction of large-scale, quantitative data collection, and use within educational systems for an increasingly diverse array of purposes and outcomes has led to the establishment and thriving existence of large data systems, which are ubiquitous in the 21st century educational systems worldwide (Rutkowski, 2008; Smith & Baker, 2001).

These large-scale, national education data systems engage in the collection, storage, analysis, and application of quantitative data to educational issues across a broad spectrum of content, issues, and agendas. This slow change has resulted in the widespread legitimization of quantitative data for educational policymaking at local, national, and international levels. Yet, even though the data itself may not be tightly coupled with every context, issue, or agenda in educational systems, the data provide a loosely coupled, legitimized evidence base, which is rationalized to fit and support policy- and decision-making in systems worldwide. Therefore, normative isomorphism creates the legitimacy needed to perpetuate LSAs and quantitative data as taken-for-granted components of national educational policymaking through the coupling of quantitative data with decision-making.

The scientization of education phenomenon further explains how quantitative data use is rationalized for use by policymakers and how the quantification of education is automated in a way that also takes advantage of the process of normative isomorphism to gradually introduce: (1) quantitative data as necessary to policymaking; and (2) the computerized approach to data use. In fact, a cornerstone of the scientization of education is its "quantification" (Wiseman et al., 2016a). Through the quantification process the personal or more affective links between education and policy become increasingly formulaic and less dependent on human intuition or context.

Yet, the "cyborg dialectic" suggests that there is an increasing symbiosis between policymakers, educators, and the technology used to monitor educational performance, collect data, and compute "best practices" based on quantitative measures of educational quality (Wiseman et al., 2016a).

Transhumanist Framework

As the cyborg dialectic becomes and remains normalized and educational organizations and their cultures accept large-scale quantitative data collection and use as a normal and valued component of educational decision-making, the conversion of decision-making to more automated processes also becomes normalized. The volume and complexity of educational data available, as well as the need to identify multilevel and multi-staged patterns within this data, lead inexorably to the use of algorithms.

Algorithms are mathematical procedures for solving problems or recognizing patterns usually calculated by computer software due to the large scale and complexity of data (Iyengar et al., 2016). Algorithmization is an embodiment of the institutionalization of cultural norms and values through educational structures, policies, and practices, which incorporate algorithms to analyze data, understand educational phenomena, and solve educational problems (Wiseman, 2017). This algorithmization signals the tight coupling of educational systems to data, but it also leads to the establishment of an educational algocracy. Danaher (2016, p. 247) defines "algocracy" as

> a system in which algorithms are used to collect, collate and organize the data upon which decisions are typically made and to assist in how that data is processed and communicated through the relevant governance system.

Danaher (2016, p. 247) further asserts that when systems, such as national educational systems, use algorithms to collect and process the data upon which policies and other decisions are made, then

> the algorithms structure and constrain the ways in which humans within those systems interact with one another, the relevant data and the broader community affected by those systems.

This caveat to algocracy points toward a more relevant critique of evidence-based educational policymaking than the moral objections to globalization typically raised by LSA critics (Meyer & Benavot, 2013). In other words, the rhythmic application of evidence-driven policy in national educational systems worldwide follows a mathematical pattern based on reams

of data collected by LSAs or amassed over time. And, this evidence-driven policy may be driven more by the data as a function of its legitimacy and automated, algorithmic analyses than by the real needs, contexts, or decisions of policymakers representing districts, political agendas, and local educators planning for or responding to the needs of students and their community.

THE LEGITIMACY OF DATA

Education policymakers around the world consistently report that numbers matter when establishing an evidence base for both existing and new policy and reform agendas (Wiseman & Baker, 2005). In fact, critics and proponents of quantitative data alike have argued that quantitative data are useful in making educational policymaking and institutionalizing reform (Smith, 2014). Quantitative data are easily collected compared to qualitative data and seemingly easily interpreted as well, although the distinction between "easily" and "appropriately" interpreted is important (Creswell, 2013). Quantitative data are more easily transferred and compared within and between systems because they give the illusion of being independent of context and are both value- and bias-free (Ragin, 2013). Quantitative data can be mined and analyzed rapidly and that also in an automated fashion, meaning that educational practices and expectations can be operationalized using quantitative data and then statistically analyzed using sophisticated software to improve the efficiency and supposed effectiveness of data use for policy- and decision-making in educational systems (Scheuer & McLaren, 2011). In fact, educational data mining (EDM)

> is concerned with developing, researching, and applying computerized methods to detect patterns in large collections of educational data—patterns that would otherwise be hard or impossible to analyze due to the enormous volume of data they exist within. (Scheuer & McLaren, 2011, p. 1)

Quantitative data have the possibility to be generalized from a sample to a population; a distinct advantage over anecdotal and qualitative data where interpretation is restricted to the sample only (Silverman, 2006). This possibility of generalization is one of the reasons that EDM is becoming increasingly ubiquitous in discussions about LSAs and their application to system-wide educational policymaking and practice. In other words, the legitimization of quantitative data has led to a proliferation of such data to the point where the sheer volume of it is overwhelming to many educators and policymakers. However, a trust in and reliance on the increasingly ubiquitous

"knowledge society" mean that there is less resistance to use of quantitative data (Hazelkorn, 2008).

As a result, quantitative data have triumphed over qualitative and anecdotal evidence in terms of the speed and degree of impact on educational policy and practice. Here again the scientization of education, and specifically quantification, serves to further illustrate this point. Consisting of explicit references to measuring progress and a reliance on research-based practices, the quantification of education was the most commonly occurring category across national education policy documents from Australia, Germany, and the USA (Wiseman et al., 2016a). Policy documents from Wiseman et al.'s (2016a) study highlighted a focus on data use for evidence-based decision-making. For example, the representative policy documents from Australia, where student assessment was positioned as an individual and holistic endeavor, emphasized quantitative data as a primary means to promote public support of education.

In a subsequent, and yet unpublished study (Wiseman, Damaschke-Deitrick, Davidson, & Bruce, 2016b), documents from a wide variety of international organizations also strongly emphasized a focus on quantification of education. For example, a document published by the Center for Universal Education at Brookings stated as follows: "Disaggregating data by sex, age, and ethnicity is critical to identifying excluded social groups and regions and monitoring their progress" (Brookings Institute, 2011, p. 46). While this quote cannot represent the discourse from all international organizations, it does highlight the reliance on quantitative data to determine the success of education delivery and quality for marginalized populations. This refrain was echoed across the education policy documents from both national and international levels.

Yet, in spite of the popularity and persuasive impact of quantitative evidence in national education policy, there are dangers inherent in the over-reliance on data out of context for decision-making, and numerical data in particular. From a researcher's perspective, one danger is the loss of perspective that triangulation provides with both quantitative and qualitative data. However, a larger and more looming danger results from the loss of policymakers' and educators' ability to understand and make decisions based on their own expert interpretation of the data, and the context in which policies are or will be implemented. Since quantitative data are so large and complex, there is a reliance on technology to interpret the data based on static values and commands set up by remote data experts rather than local policymakers and the educators implementing the policies in schools and classrooms. This is an embodiment of the cyborg dialectic (Wiseman et al., 2016a).

ASSESSMENT INFRASTRUCTURE AND CAPACITY BUILDING

The systematic collection, interpretation, and application of quantitative data within national educational systems is a constant drain on assessment infrastructure and the capacity of education officials', administrators', and educators' assessment-relevant knowledge and skills. LSA requires the coordination of several branches of a nation's educational system, including central national leadership and oversight, the marshalling of resources ranging from material supplies for the manufacture and distribution of test booklets and answer sheets, to personnel ranging from test content developers to those responsible for the fair and procedurally appropriate administration of the tests on site, to those educators and administrators in selected testing school sites who must prepare students for the tests as well as proctor and safely collect and store completed answer sheets and test booklets.

These many infrastructure and capacity (i.e., knowledge and skills) needs are not always present within a nation's educational system or available when LSA is first introduced and planned. Instead, large-scale, standardized assessments administered within national educational systems are key moments for the development and institutionalization of both testing infrastructure and capacity, but they often lack sustainable local support, leadership, and forward movement. The sustainability of LSAs then relies not only on the infrastructure and capacity-building, which occur either in preparation for or as a result of participation in and administration of LSAs, but also on the degree to which invested stakeholders are empowered and incorporated into the decision-making apparatus of LSAs at each level of interaction in the national educational system.

The rhythmic application of quantitative LSA data, therefore, is contingent upon many factors. These factors—as mentioned above—determine the extent to which LSAs and quantitative data collection, analysis, and application are determined by the infrastructure, capacity, and sustainability of the national educational system itself. Part of what makes the algorithmization of evidence-based policy so important is that with adequate testing, infrastructure and capacity necessary for the sustainability of LSAs, and other quantitative data systems, is minimal. At a minimum, the necessary infrastructure and capacity include resources, knowledge, and skills needed to develop the assessment instruments and basic system, plus the programming and resources for the establishment of a computer-based quantitative data collection and analysis system. In other words, buy-in and empowerment of policymakers and educators are not needed when the process is largely automated.

THE INSTITUTIONALIZATION OF EXTERNAL COMPARISON

The trend in national educational policymaking since the 1960s international development era began has been to rely increasingly on the availability of international education data (Wiseman & Baker, 2005). The highly problematic trend is to then use this quantitative data beyond its practical scope, and apply the data beyond its policy relevance (Heyneman & Lee, 2012). These trends—both positive and negative—are possible because of the ubiquitous availability of internationally comparative, quantitative educational data in the 20th and 21st centuries. The ubiquitous availability of international education data is partly the reason behind the global spread of transnationally standardized educational models (Schriewer, 2000), yet it has not historically been the data itself that "made" policy or reform. That is a more recent trend. Instead, as Holzinger and Knill (2005, p. 784) asserted:

> Transnational problem-solving typically occurs within transnational élite networks or epistemic communities, defined as networks of policy experts who share common principled beliefs over ends, causal beliefs over means and common standards of accruing and testing new knowledge. Common educational and normative backgrounds typically facilitate joint development of common policy models in such constellations.

What Holzinger and Knill (2015) found was that beliefs and standards among policymakers and policy models during much of the 20th century followed shared norms for making meaning. This historical approach to data use for policymaking asserts the primacy of meaning-making over the data itself.

Both the phenomenon of policy borrowing and externalization were explained frequently as hegemonic processes of control, whereby dominant systems' cultural norms and values were either forced on less-dominant and periphery nations' systems through direct coercion or indirectly through borrowing of models that "work" (Phillips & Ochs, 2003). Yet, the primacy of dominant and often malicious agency is only a part of the reason for the institutionalization of external comparison. Other, less direct, agency explanations offer more accurate explanations of the externalization process. For example, Vavrus (2004, pp. 141–142) argues as follows:

> Up to now, the process of externalization has been considered primarily from an international perspective that looks at how educational models and policy language in one state are appropriated by another … an equally compelling approach to the study of borrowing and lending explores how the language of policy—in education and in other sectors—is externalized without reference to a specific lending country or agency. Thus, I argue that the intensification of global networks in recent years has contributed to the universalization of keywords.

Although Vavrus (2004) does not say it explicitly, she presents an implicit argument for the process of normative isomorphism occurring through the externalization of policy language to legitimize and institutionalize universal keywords. These universalized keywords contribute to the development of rationalized scripts among educational policymakers and educators, which institutionalize external comparisons of education using internationally collected and shared LSA data.

Universalized and rationalized scripts lay the foundation for algocracy of education through the rhythmic application of evidence-driven policy in national educational systems by: (1) establishing shared expectations and values about what education is and how it should be structured or studied; (2) constructing education in a rationalized way that emphasizes capacity, diversity, and efficiency within the boundaries of shared expectations; and (3) creating rationalized constructions, which are quantifiable, and therefore data reliant. Furthermore, an indulgence in data, which represents quantifiable rationalized constructions of globally legitimized educational outcomes, is a key part of the institutionalization of the external comparison of educational data worldwide and evidence-driven educational policy. For example, increasingly electronic, and sometimes automated, analysis and interpretation of quantitative assessment data occurs within national educational systems (Koedinger, D'Mello, McLaughlin, Pardos, & Rosé, 2015).

Data analysis software systems are regularly introduced by international corporations such as Pearson and Educational Testing Service (Mislevy, Behrens, Dicerbo, & Levy, 2012). National educational systems and policymakers worldwide are increasingly implementing aligned LSA systems for both internal use and international data sharing (Hinze-Pifer & Ramsey, 2011). In turn, big education firms such as Pearson Education, the ETS, Houghton Mifflin Harcourt, and McGraw-Hill spend millions lobbying for pro-testing policies (Strauss, 2015). In fact, reports suggest these firms spent more than $20 million between 2009 and 2014 lobbying at the state and national level for standardized assessment policies. There is a symbiotic relationship between government and business related to LSA and quantitative data for evidence-driven policy. For example, in the USA, the Department of Education pushes for the use of LSA data, and privatized education companies make major financial contributions to education policy lobbyists, who then directly benefit from the adoption of testing policies.

There are other international examples of data indulgence and the institutionalization of external comparison. For example, the Systems Approach for Better Education Results (SABER) from the World Bank includes measurement data for national education management information systems. There are many examples related to the use of SABER in various national ministries of

education (Robertson, Mundy, & Verger, 2012). These examples suggest that data management in educational systems and the evidence produced from these efforts drives much of the national level educational policy and reforms in these countries. And, this evidence-driven policy is both assumed and descriptively shown to improve student achievement. As a result, increasingly more educational data are collected and shared worldwide.

Expansion of data or availability of data beyond its immediate or intended scope and purpose has become the norm in both educational policymaking and comparative and international education research that supports it. For example, educational policymakers and comparative education researchers can use data from Indonesia even though they are not Indonesian, nor do they have a stake in Indonesian education or a specific need for the data related to Indonesia. The Indonesian data may be compared externally to local or regional data merely because of its convenience or to support a policy agenda without reference to or knowledge of the Indonesian educational context. Policymakers regularly compare their educational systems with high scoring nations on international LSAs, such as the Program for International Student Assessment (PISA) or the Trends in International Mathematics and Science Study (TIMSS), even though the contexts, history, and cultures of the compared systems are extremely dissimilar or even contradictory to each other (Crossley & Broadfoot, 1992; Phillips & Ochs, 2003).

The easy availability and legitimacy of LSA data also generates an addiction of sorts to this data among policymakers and the public alike. This addiction is enacted when quantitative data from these assessments are used as evidence for particular policies or reforms regardless of the measurable impact of the policy or reform when implemented. There is also often no sense among data-addicted policymakers about whether data-driven policies and reforms are "right" or "wrong" for a particular national educational system. Instead, an addiction to the data demonstrates itself when policymakers use quantitative evidence regardless of the value of a particular policy or practice, but instead only because it is technically, statistically, or procedurally possible to use the data as evidence.

THE IRONY OF EXTERNALLY DRIVEN INTERNAL ACCOUNTABILITY

The global institutionalization of externalization by national education systems worldwide is increasingly demonstrated by "shadow" versions of

LSA results, reports, and reforms. For example, the nationally representative results from the National Assessment of Educational Progress (NAEP) administered in the USA have been reported in the *Nation's Report Card* for several decades (Haertel, 2016). The system of annual testing and reporting is an example of nation-level accountability, but has also been replicated in every state of USA through state-level annual assessments pushed by the No Child Left Behind policy in the early 2000s (Jacob, 2017). These system-wide assessments of students' academic performance are mimicked in national education systems worldwide as well.

Since the turn of the century, not only has the quantity of and focus on international assessments increased but so has attention to national assessments too. For example, in the early 2000s in South Africa, findings from international assessments, including the Progress in International Reading Literacy Study (PIRLS), TIMSS, and the Southern African Consortium for Monitoring Education Quality (SACMEQ) revealed that South African students scored among the lowest in the world on standardized assessments. As a result, the South African Department of Education adopted several national initiatives to improve their ranking on international assessments. These initiatives included curriculum revisions and the incorporation of national standardized tests for students in grades 1–6 and 9 with the stated purpose of increasing student performance on international assessments such as PIRLS, TIMSS, and SACMEQ (Carrim, 2013).

Previous studies have found that even nations which choose not to participate in these LSAs develop their own national testing regimes. For example, with the help of the Educational Evaluation Research Consortium and funding from the United States Agency for International Development, the Dominican Republic has implemented a long-term assessment project (Kamens & McNeely, 2007). The Conference des Ministres de l'E´ducation des Pays Ayant le Franc¸ais en Partage in Francophone Africa assists in the development of national assessments. The national assessments in both of these cases, and in many others, resemble the international model established by the NAEP in the USA, and PISA, TIMSS, and PIRLS driven by national educational systems participating in assessments sponsored by either the Organization for Economic Cooperation and Development or the International Association for Evaluation of Educational Achievement (Kamens & McNeely, 2007).

There are also regional intra-national comparisons with national educational systems. For example, several within-country states or regions participate in international assessments, such as TIMSS, as benchmarking communities or district participants. For example, there have been almost 30 benchmarking

communities participating in TIMSS since the 1995 cycle, including many repeat participants (National Center for Educational Statistics (NCES), 2017). Between these types of participation in established LSAs and shadow versions of LSAs developed independently by national educational systems along legitimized and commonly recognized lines, there are many forms of externally driven internal accountability. Even South Africa, which completely withdrew from international LSAs in the early 2000s is participating again in them, but was also focused on the LSA model format and content while withdrawing in order to participate more fully when re-engaged (Reddy, 2005).

THE PATH TO DATA-DRIVEN POLICY

As much of the evidence presented here suggests, data-driven policy moves from data-informed to data-based, then finally to data-driven policy. *Data-informed* policy is when quantitative data are legitimizing. This type of data then informs decision-making, but is not a required element of the policy-making process. *Data-based* policy is when quantitative data are necessary, and a required foundation for decision-making, but policymakers still make informed policy decisions based on the data, implementation needs, and context. Finally, *data-driven* policy is when quantitative data are automatically collected and generate either policy recommendations or actions automatically without the filter of human understanding or reasoning (Hess & Little, 2015). There are, therefore, three stages to the path of data-driven educational policy.

Data-informed policy is among the older forms of comparative, cross-national educational policymaking established among the first education specialists as "travelers' tales" (Noah & Eckstein, 1969). As educators have traveled beyond their local or national communities, they have often documented the educational habits, traditions, and systems of other communities where they traveled. In the USA, one of the most famous examples of data-informed policy comes from Horace Mann's trip to Prussia in 1843. From this trip, and Mann's observations of Prussia's national educational system, he brought many organizational structures and educational values back to his educational system in Massachusetts, which then spread across the USA. In this stage, the evidence may select or suggest policies, but can only be implemented with the command or agency of policymakers and educators.

The next stage on this path is data-based policy. When policy is based on quantitative data it may trigger certain automatic responses, but there is oversight and the possibility of intervention by policymakers and educators.

In a data-based system, policymakers often set the threshold, standards, or minimum criteria. Quantitative data or evidence may be set to spur certain responses or actions automatically, but if needed, these automatic responses can be modified by policymakers and educators. An example of this is the "basic competency" standards the policymakers set for schools, which may trigger automatic sanctions if not met. Again, the USA provides an example of data-based policy in its "adequate yearly progress" (AYP) requirement instituted as part of the No Child Left Behind policy of the early 2000s (US Department of Education, 2002). Schools and districts that did not meet the threshold for AYP were automatically put on probation or defined as "at risk" schools. These designations were publicly reported and often led to the takeover of school districts by cities or private firms among other significant changes (Wong, 2006). Yet, even with these dramatic consequences, policymakers were still able to intervene and adjust the response based on context or other factors.

The final stage along this path is data- or evidence-driven policy. This stage signals the full algorithmization of educational data and attendant policy. In this type of system, the quantitative data alone may trigger certain policies and their implementation and policymakers and educators are then either expected or bound to carry out those evidence-driven policies without oversight or the possibility to intervene or override. While there are no active examples of this extreme form of data-driven control, the ubiquity of LSA data and comparative analysis of this type of data indicate that local educators and policymakers have already lost control of the data and its effect. Instead, evidence suggests that policymakers have become more passive consumers of theirs and others' LSA data than decision-makers.

CONCLUSION

Although the evidence-driven policy is not fully implemented worldwide, there is evidence that educational policymaking is increasingly algorithmized. How might national educational policymaking get to the point of full algorithmization, and what does the rhythmic application of evidence-based policy look like worldwide? Evidence presented here suggests that this rhythmic application of quantitative data to educational policy and practice occurs through the measurable implementation of several symbiotic, system-level phenomena, including: (1) legitimacy seeking and the legitimization of quantitative data for educational decision-making; (2) constant development and

re-development of national educational assessment infrastructure and capacity; (3) reliance on external comparison of quantitative measures of educational quality; (4) development of internal accountability systems reliant on LSA data; and (5) institutionalization of quantitative-based educational decision-making within national educational systems across a broad range of national social, economic, and political contexts (Wiseman & Waluyo, 2017).

REFERENCES

Baker, D. P., & Wiseman, A. W. (Eds.). (2006). The impact of comparative education research on institutional theory. In *The international perspectives on education and society series* (Vol. 7). Oxford: Elsevier Science.

Best, M., Knight, P., Lietz, P., Lockwood, C., Nugroho, D., & Tobin, M. (2013). *The impact of national and international assessment programmes on education policy, particularly policies regarding resource allocation and teaching and learning practices in developing countries.* Final Report. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.

Brookings Institute. (2011). *A global compact on learning: Taking action on education in developing countries.* Washington, DC: Author.

Carrim, N. (2013). Approaches to education quality in South Africa In Y. Sayed, A. Kanjee, & M. Nkomo, (Eds.), *The search for quality education in post-apartheid South Africa: Interventions to improve learning and teaching* (pp. 39–60). Cape Town, South Africa: HSRC Press.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage Publications.

Crossley, M., & Broadfoot, P. (1992). Comparative and international research in education: Scope, problems and potential. *British Educational Research Journal*, *18*(2), 99–112.

Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy & Technology*, *29*(3), 245–268.

DiMaggio, P. J., & Powell, W. W. (Eds.). (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.

Haertel, E. (2016). *Future of NAEP long-term trend assessments: A white paper prepared for the national assessment governing board*. Stanford, CA: Stanford University, CA. Retrieved from https://www.nagb.org/content/nagb/assets/documents/newsroom/naep-releases/naep-long-term-trend-symposium/long-term-trends.pdf. Accessed on August 31, 2017.

Hazelkorn, E. (2008). Learning to live with league tables and rankings: The experience of institutional leaders. *Higher Education Policy*, *21*(1), 193–215.

Hess, F. M., & Little, B. (2015). '*Moneyball' for education: Using data, evidence, and evaluation to improve federal education policy.* Washington, DC: American Enterprise Institute.

Heyneman, S. J., & Lee, B. (2012). Impact of international studies of academic achievement on policy and research. In L. A. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large scale assessment: Background, technical issues, and methods of data analysis* (pp. 37–74). London: Chapman and Hall Publishers.

Hinze-Pifer, R., & Ramsey, D. S. (2011). Evaluating education information systems: Implementation of longitudinal student data systems in six school districts. *Policy Perspectives*, *18*(1), 71–91.

Holzinger, K., & Knill, C. (2005). Causes and conditions of cross-national policy convergence. *Journal of European Public Policy*, *12*(5), 775–796.

Iyengar, R., Mahal, A. R., Aklilu, L., Sweetland, A., Karim, A., Shin, H., … Pokharel, P. (2016). The use of technology for large-scale education planning and decision-making. *Information Technology for Development*, *22*(3), 525–538.

Jacob, B. (2017). The changing federal role in school accountability. *Journal of Policy Analysis and Management*, *36*(2), 469–477.

Kamens, D. H., & McNeely, C. L. (2007). Globalization and the growth of international educational testing and national assessment. *Comparative Education Review*, *54*(1), 5–25.

Klees, S., & Edwards, D. B. (2014). Knowledge production and technologies of governance. In T. Fenwick, E. Mangez, & J. Ozga, (Eds.), *Governing knowledge* (pp. 31–43). London: Routledge.

Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(4), 333–353.

Lingard, B., Martino, W., & Rezai-Rashti, G. (2013). Testing regimes, accountabilities and education policy: Commensurate global and national developments. *Journal of Education Policy*, *28*(5), 539–556.

Meyer, H. D., & Benavot, A. (Eds.). (2013, May). *PISA, power, and policy: The emergence of global educational governance*. London: Symposium Books.

Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining. *JEDM-Journal of Educational Data Mining*, *4*(1), 11–48.

National Center for Educational Statistics (NCES). (2017). *TIMSS participating countries*. Retrieved from https://nces.ed.gov/timss/countries.asp. Accessed on December 14, 2017.

Noah, H. J., & Eckstein, M. A. (1969). *Toward a science of comparative education.* London: Macmillan.

Phillips, D., & Ochs, K. (2003). Processes of policy borrowing in education: Some explanatory and analytical devices. *Comparative Education*, *39*(4), 451–461.

Ragin, C. C. (2013). *The comparative method: Moving beyond qualitative and quantitative strategies.* Berkeley, CA: University of California Press.

Reddy, V. (2005). Cross-national achievement studies: Learning from South Africa's participation in the trends in international mathematics and science study (TIMSS). *Compare: A Journal of Comparative and International Education*, *35*(1), 63–77.

Robertson, S., Mundy, K., & Verger, A. (Eds.). (2012). *Public private partnerships in education: New actors and modes of governance in a globalizing world*. London: Edward Elgar Publishing.

Rutkowski, D. (2008). Towards an understanding of educational indicators. *Policy Futures in Education*, *6*(4), 470–481.

Scheuer, O., & McLaren, B. M. (2011). Educational data mining. In *The encyclopedia of the sciences of learning*. New York, NY: Springer.

Schriewer, J. (2000). World system and interrelationship networks. In T. S. Popkewitz (Ed.), *Educational knowledge: Changing relationships between the state, civil society, and the educational community* (pp. 305–343). Albany, NY: SUNY Press.

Silverman, D. (2006). *Interpreting qualitative data: Methods for analyzing talk, text and interaction*. Thousand Oaks, CA: Sage.

Smith, T. M., & Baker, D. P. (2001). Worldwide growth and institutionalization of statistical indicators for education policy-making. *Peabody Journal of Education*, *76*(3–4), 141–152.

Smith, W. C. (2014). The global transformation toward testing for accountability. *Education Policy Analysis Archives*, *22*(116), n116.

Strauss, V. (2015), March 30). Report: Big education firms spend millions lobbying for pro-testing policies. Washington Post. Accessed online 22 March 2018 at https://www.washingtonpost.com/news/answer-sheet/wp/2015/03/30/report-big-education-firms-spend-millions-lobbying-for-pro-testing-policies/?utm_term=.e70fe459ba66.

US Department of Education. (2002). *No Child Left Behind Act*. Washington, DC.

Vavrus, F. (2004). The referential web: Externalization beyond education in Tanzania. In G. Steiner-Khamsi (Ed.), *The global politics of educational borrowing and lending* (pp. 141–153). New York, NY: Teachers College Press.

Wiseman, A. W. (Ed.). (2010). The impact of international achievement studies on national education policymaking. In *International perspectives on education and society* (Vol. 13). Bingley: Emerald Group,.

Wiseman, A. W. (2017). The algorithmization of national educational assessments and policymaking. Paper presented at Transcending humanity? Transhumanism, algorithms, & scientization as phenomena of late modernity, University of Tübingen, Germany.

Wiseman, A. W., Astiz, M. F., & Baker, D. P. (2014). Comparative education research framed by neo-institutional theory: A review of diverse approaches and conflicting assumptions. *Compare: A Journal of International and Comparative Education*, *44*(5), 688–709.

Wiseman, A. W., & Baker, D. P. (2005). The worldwide explosion of internationalized education policy. In D. P. Baker & A. W. Wiseman (Eds.), *Global trends in educational policy* (Vol. 6, pp. 1–21). London: Elsevier Science.

Wiseman. A. W., Damaschke-Deitrick, L., Bruce, E., Davidson, P., & Taylor, C. S. (2016a). Transnational scientized education discourse: A cross-national comparison. In J. Schmid, K. S. Amos, J. Schrader, & A. Thiel (Eds.), *Internationalisierte welten der bildung—Bildung und bildungspolitik im globalen vergleich* (pp. 121–146). [*Internationalized worlds of education—Education and training policies in global comparison.*] Baden-Baden, Austria: Nomos.

Wiseman, A. W., Damaschke-Deitrick, L., Davidson, P., & Bruce, E. (2016b, June). Scientization discourse in education: Globalization and contextualization. Paper presented at the research symposium on "The scientization of the world: A global cultural cognitive pattern in late-modern societies", Eberhard Karls Universität Tübingen, Germany.

Wiseman, A. W., & Waluyo, B. (2017). The dialectical impact of PISA on international educational discourse and national education reform. In L. Volante (Ed.), *The PISA effect on global educational governance* (pp. 33–48). Oxford: Routledge Press.

Wong, K. (2006). The political dynamics of mayoral engagement in public education. *Harvard Educational Review*, *76*(2), 164–177.