Journal : Small 11125 | Dispatch : 29-10-2014 | Pages : 12
Article No. : 9328 | LE | TYPESET
MS Code : PROS_4404_Wagner | CP | DISK

Prospects
DOI 10.1007/s11125-014-9328-8

TRENDS/ CASES

Learning at the bottom of the pyramid: Constraints, comparability and policy in developing countries

Daniel A. Wagner · Nathan M. Castillo

Abstract United Nations development goals have consistently placed a high priority on the quality of education—and of learning. This has led to substantive increases in international development assistance to education, and also to broader attention, worldwide, to the importance of children's learning. Yet, such goals are mainly normative: they tend to be averages across nations, with relatively limited attention to variations within countries. This review provides an analysis of the scientific tensions in understanding learning among poor and marginalized populations: those at the bottom of the pyramid. While international agencies, such as UNESCO and OECD, often invoke these populations as the "target" of their investments and assessments, serious debates continue around the empirical science involved in both research and policy. The present analysis concludes that the UN post-2015 development goals must take into account the critical need to focus on learning among the poor in order to adequately address social and economic inequalities.

Keywords Learning · Low- and middle-income countries · Poor and marginalized populations · Learning outcomes · Constraints · Comparability · Education policy

The World Conference on Education for All in Jomtien, Thailand, was a watershed moment in international education and development. Held in 1990, the conference embraced two key challenges: significantly increase access to education for children in poor countries, and promote the quality of learning in education. A decade later, at the 2000 Education for All (EFA) conference in Dakar, these same two challenges were expanded into a detailed list of six education targets in the Dakar EFA Framework for

D. A. Wagner (✉) · N. M. Castillo
Graduate School of Education, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104-6216, USA
e-mail: wagner@literacy.upenn.edu

N. M. Castillo
e-mail: ncast@gse.upenn.edu

Journal : Small 11125 | Dispatch : 29-10-2014 | Pages : 12
Article No. : 9328 | ☐ LE | ☐ TYPESET
MS Code : PROS_4404_Wagner | ☑ CP | ☑ DISK

Action. The aims were to promote early childhood care, make primary school compulsory, address learning needs for all, promote adult literacy, reduce gender disparities, and develop quality measures of learning outcomes (UNESCO 2003, p. 28). They were reinforced again in the UN Millennium Development Goals (MDGs) for 2015, where universal primary education was made the second of eight major goals (United Nations 2000). These global efforts led international development agencies to substantively increase their assistance to education; it also led the broader public to pay more attention to children's learning on a global scale.

There is a large and diverse empirical research base in the area of human learning. However, much of the available research is substantially limited by constraints of various kinds. Most prominent among these constraints is the limited ability to generalize from findings in one population context to other distinct population contexts. Similarly, research methods may vary greatly between one set of studies and another, making it difficult to discern whether the findings vary because of the methods or other factors. These are, of course, classic problems in the social sciences.

In this article, we analyse the scientific tensions in understanding learning among poor populations, those that Prahalad (2006) called the bottom of the pyramid (BOP). While international agencies often claim to target investments to populations most in need, serious debates continue about the empirical science needed to implement appropriate policies, with inevitable consequences for effective implementation in developing or low- and middle-income countries (LMICs).

Constraints

Learning in two South African classrooms

> Shayandime Primary School, with buildings fashioned of adobe walls and zinc roofing, is located in a small rural village in the northern province of Limpopo, South Africa. Just a few dozen miles from the border of Zimbabwe, the area is dotted with traditional houses called *rondavels*, an adapted version of the southern Africa style hut. Baboons roam the school grounds freely and are known to slip through the space between the red-mud block and corrugated roofing and vandalize the classrooms at night. Despite the occasional broken window, the school is not without resources. It is one of many establishments in the region that received a donation of early-model desktop computers where upper primary learners spend time working on basic typing activities. However, disadvantaged learners with weak English proficiency have no access to the computers since no programmes have been written in their local language, Venda. In the classroom, learners spend most of their time copying sentences from the chalkboard, and are rarely stimulated to participate in activities that support creativity and critical thinking skills.

> By contrast, 4 hours away in the provincial capital of Polokwane sits Central Elementary School. With brick paths around the perimeter and a state-of-the-art computer lab, it has flat-screen monitors and a smart board with projector. The computer lab, which rivals that of the local university, was acquired in part through revenue earned by renting out the school's event hall to the community. There are no broken windows, the teachers present structured lesson plans, and the parents are an integral

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 LE TYPESET
MS Code : PROS_4404_Wagner CP DISK

part of the school culture. Given its appealing learning environment, the provincial officials proudly exhibit this urban school to visiting national and international education planners. Many students have mobile phones, and give the appearance of being motivated to learn and to be connected to South Africa's future. (Author's note: the school names have been changed, and these profiles combine details from several schools.)

Comparisons of rural and urban contexts in LMICs often consist of these types of observations of infrastructural and social characteristics. The South Africa Annual National Assessment (ANA), administered at the end of each school year, measures progress in learner achievement in grades 1–6 and 9 (DBE 2013). It tends to confirm the subjective account offered above. South African schools are categorized according to a poverty index based on the relative wealth or poverty of the community and are grouped into quintiles. Rural Shayandime Primary belongs to the lowest quintile. When the ANA was conducted there, only five learners in grade 3 scored above the national norm while the large majority scored in the bottom 10%, creating a bimodal distribution. By contrast, Central Elementary ranks in the middle (third) quintile with normally distributed scores; these are somewhat below the national urban norms for the Mathematics and Home Language reading competencies for grade 3 (DBE 2013).

The contrast in ANA test performance between these schools raises two related questions: Why does Shayandime have a handful of high achievers, with the rest of its students clustered around the lower achievement continuum? And, within urban Central Elementary, why are the scores more normally distributed, though below the national average? Studies often point to the role that social and family influences have on predicting learning outcomes, mainly in terms of what we call power, parents, and privilege. For example, various authors have studied the impact of social stratification on school results in terms of such factors (Benedict and Hoag 2004; Buchmann and Hannum 2001; Korinek and Punpuing 2012; Lu and Treiman 2011). In many cases, the learners with the right combination of these influences are the ones who tend to sit closer to the front in large classrooms and have greater focus and motivation for learning.

These findings have serious implications. Our understanding of the psychological science of learning derives primarily from data from wealthy OECD countries, where educational outcomes, including standardized test scores, are, by design, normally distributed. However, variation around the world, and especially in BOP contexts such as we noted in rural South Africa, may be bimodal: a few top scorers, and many low-scoring students. Where the contrast is so dramatic, and where the focus is on the bottom end of the normal curve, we are forced to reconsider notions of statistical normality, and with it, the idea that BOP contexts are simply an extension of the typical normal curve.

The idea that learning—in and out of school—may vary significantly across cultures is hardly new (Cole, Gay, Glick, and Sharp 1971; Wagner 1993, 2014). Nonetheless, the continued increase in pressure to globalize data collection on education has pushed both researchers and policy makers to ignore, or minimize, such differences (Benavot and Tanner 2007). We believe that the failure to give serious consideration to learning at the bottom of the pyramid distorts reality and may also lead to ill-considered interventions on behalf of poor students.

Beyond South Africa, examples of distinctive learning styles abound in the research literature. Some of this work originated decades ago in the United States with pioneering studies of individual differences in learning (Kagan, Moss, and Sigel 1963; Witkin, Moore, Goodenough, and Cox 1977). Today, learners in poor LMIC schools often struggle to understand the language of instruction and the language of reading. In such situations, teachers often emphasize rote

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 LE TYPESET
MS Code : PROS_4404_Wagner CP DISK

learning and memorization; though this remains a very common learning strategy around the globe, it is widely derided by modern (Western-trained) pedagogues (Wagner 1983). Another example of such contrasts is the way learning is constructed in various societies, such as those strongly influenced by Confucianism (Li 2003).

Learning is ubiquitous and takes many forms in everyday life. In education, learning is measured with instruments that can reliably estimate both processes and outcomes—or learning assessments. If it is necessary for an assessment to be representative of an entire population of a country, or valid across multiple countries in a comparative framework, then it will likely cost more in terms of both time and resources. Thus far, researchers have controlled time and resources by delimiting the range of *skills* to be assessed (the *skills sample*), and by constraining the *population* to be included (the *population sample*). It is important to understand these two forms of boundary constraints in terms of technical and statistical requirements, as well as policy requirements and outputs. Each of these issues poses empirical and statistical challenges.

Skill sampling and assessments

It is widely accepted that humans learn by sampling their environment, beginning by using their built-in senses from the moment of birth. Clearly, no infant, child or adult could possibly survive by taking in the totality of information available in the environment. In other words, human systems are designed to discriminate so they can sample for the information that will best help them handle learning challenges (Kahneman 2011). Indeed, parents typically prepare young children to adapt, learn and survive precisely by exposing them to the range of situations they will likely encounter in their lives. Of course, not all these learning environments may be similarly well adapted for a child's future in educational settings.

When it comes to scientific research on learning, we humans do best when we take samples of our informational environment, whether in educational institutions or via word of mouth or, increasingly, via Internet search engines such as Google. This relatively simple observation is very relevant here: one of the most vexing problems researchers encounter in studying and evaluating learning is how to generalize from one sampling of skills to another. Thus, sampling a finite set of skills, and knowing about the contextual situations in which they are used, are key elements of all learning assessments.

In designing learning research and evaluation strategies, researchers make highly complex decisions: they choose contextual and demographic variables (e.g., child's age, year in school, gender, socio-economic status), and select the skills to be assessed and the type of research methodology to apply. Each option is tied to a set of assumptions and compromises, and the selections included in the final research design will influence the validity, reliability and practical feasibility of the chosen approach (Braun and Kanjee 2006; Wagner 2010, 2011a). Furthermore, research must be designed so it can respond to dynamic changes over time. And, as expectations of literacy, numeracy and higher-order skills adapt to changes in social and economic environments, the measurement methods must also be adapted so they align with educational goals (Wagner, Murphy, and de Korne 2012).

Learning and population samples

Population sampling also matters. For example, about 5% of the world's population resides in the United States, but nearly 95% (Arnett 2008) of scientific publications on psychological development are based on populations that are WEIRD (western, educated,

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 LE TYPESET
MS Code : PROS_4404_Wagner CP DISK

industrialized, rich, and democratic; Heinrich, Heine, and Norenzayan 2010), and living primarily in OECD countries. Moreover, of the research on psychological development conducted in the United States, about 80% is on "majority" ethnic groups (those of European origin), though these groups account for only about 50% of the current U.S. population (Arnett 2008). It seems obvious that researchers should explicitly address questions of representativeness and external validity, but often they do not. These critiques also apply to international research, as much of the available research on learning is constrained in important ways by scientific data sets and research studies drawn from population samples living mainly within middle- to high-income countries. Fortunately, this trend is now beginning to change (Wagner 2014).

In international large-scale educational assessments (LSEAs), key parts of BOP populations may be excluded from, or under-represented, in samples that are said to be national (Engel and Feuer 2014; Wagner 2011b). Gender has been a leading reason why children in LMICs do not attend school, although recent decades have seen significant progress. Still, in the poorest countries, fewer girls than boys are present in schools at the two points when achievement is often measured: the entry to primary and to post-primary school. The systematic exclusion of girls in low-income countries usually results in fewer adolescent girls attending school; those who do attend often earn lower scores on national assessments compared to boys. For example, in the SACMEQ regional assessment in 6th grade, undertaken in 2007, Saito (2011) found that, averaged over 15 African countries, boys generally outperformed girls in mathematics, while girls outperformed boys in reading. However, national differences in gender disparities varied widely in both reading and math. Similar trends arise in national assessments that oversample the easier-to-reach urban areas in low-income countries. Further, in some LMICs, the difficulty of literally tracking down nomadic children can make it onerous and expensive for education authorities to include them in schools (UNESCO 2010).

Another issue is the language variation across ethnic groups that exists in nearly every country. Many of these groups, sometimes termed ethno-linguistic minorities, are well integrated into a national mix, as in Switzerland; but in other situations this variation may contribute to civil strife. Latin America, with over 500 indigenous languages, is one region where intercultural bilingual education is expanding to promote social change; to date, 12 governments have institutionalized multilingual pedagogy (Cortina 2014). Often, social and political forces try to help resolve differences, usually including policy decisions that result in a hierarchy of acceptable languages to be used in schools and governance structures. In such situations, whether in OECD countries or LMICs, it is not unusual for children who speak minority languages to be excluded from research and assessments of learning.

This process of exclusion also occurs in regions where civil conflict or economic distress leads to substantial cross-border migration, where immigrant groups (and their children) are treated as transients, and where children are provided with little or no schooling (Pigozzi, Carrol, Hayden, and Ndaruhutse 2014). The 2010 Global Monitoring Report describes how marginalization can threaten educational attainment as these children face many challenges. The world's most marginalized learners are generally faced with "inequalities, stigmatization, and discrimination linked to wealth, gender, ethnicity, language, location and disability" (UNESCO 2010, p. 5). The degree to which groups are, or are not, included in population samples has serious implications as researchers develop norms for learning outcomes. The majority of those in the population of interest may treat "others" as an inferior group that "cannot learn". Ironically, in South Africa, where the poor are in the numerical majority, it is the poor rural students who feel the most

Journal : Small 11125 | Dispatch : 29-10-2014 | Pages : 12
Article No. : 9328 | □ LE | □ TYPESET
MS Code : PROS_4404_Wagner | ✔ CP | ✔ DISK

marginalized and powerless (Babson 2010). In sum, both skills and population samples vary, as do the learning processes that individuals deploy and the contexts in which they take place.

Finally, we must consider the stakeholders who *do* the sampling. Whether they are policymakers, psychometricians, or local teachers, they all come to the task of sampling skills and populations with their own experiences and points of view. Choices about which skills to sample, among which populations, and in which languages and contexts, also add potential bias to an already complex set of sampling issues. In order to address such biases, researchers can use a range of methods including tailored sampling and subsample designs, matching samples, oversampling marginalized populations, and mixed methods designs. The consequences of these various constraints can have an important impact on educational policy and practice, and on global educational governance (Meyer and Benavot 2013).

Methodological credibility

Research that can be converted into policy depends on its credibility—which means that well-trained specialists must achieve a consensus on the merits of a particular objective set of findings, even if they might disagree with the interpretation of such findings. The two most often-cited dimensions of credibility in learning research are validity and reliability.

The validity of any learning measurement tool can be determined in several ways. First, internal validity is determined by the degree to which findings can be credibly linked to the conceptual rationale for the intervention by minimizing systematic error, or bias. For example, do questions on a multiple-choice test really relate to a child's ability to read, or to the ability to remember what he or she has read earlier? Validity can vary significantly by context and by population, since a test that might be valid in London may have little validity in Lahore. Similarly, a reading test used effectively for one language group of mother-tongue speakers may be quite inappropriate for children who are second-language speakers of the same language. This second type of validity is appropriately referred to as external: the concern is whether findings are replicable across contexts. If data continues to be aggregated without regard to local context, assessments may misrepresent learners in BOP contexts.

A third type of validity concern has been raised with respect to international LSEAs: how valid are the choices of test items and how appropriate is their content, when they are applied to local cultures and local school systems? While much learning research takes the form of quantitative testing, qualitative and ethnographic methods can also contribute, particularly with respect to cultural variation in learning processes in diverse contexts.

Reliability is often measured in two quantitative ways. In general, reliability means the degree to which an individual's results on a test are consistently related to additional times that the individual takes the same (or equivalent) test. High reliability usually means that the rank ordering of individuals taking a given test would be very similar on a second occasion. A second, and easier, way to measure reliability is to look at the internal function of the test items: Do the items in each part of an assessment have a strong association with one another? This is inter-item reliability, measured by Cronbach's *alpha* statistic. Of course, reliability implies little about the validity of the instrument: the researchers' consensus that the instrument is relevant to educational outcomes.

Seen in a qualitative perspective, reliability would be achieved when context-sensitive ethnographers, for example, agree on a set of observations of learning processes that they have gathered independently in a particular context. This is an example of "team ethnography", which is increasingly being used in education research in the United States and Europe (Bartlett and Garcia 2011; Blackledge and Creese 2010). Further, the use of randomized control trials (RCT) is seen as an important way to increase the credibility of research findings, by comparing interventions with control groups. Recent reviews by Kremer and Holla (2009), Banerjee and Duflo (2011) and Bruns, Filmer, and Patrinos (2011) support the use of RCTs for improving research credibility in international development work, while others (e.g., Castillo and Wagner 2014) suggest some serious limitations of the use of RCTs for the design of educational policy.

The diversity of learning outcomes is most often summarized in terms of an average or normal range that can be mapped along the predictable dimensions of a bell-shaped curve (Gurn 2010). As many have observed, the notion that human behavior falls along some normal curve, with the majority of observations concentrated around a discernable average, oversimplifies the range and diversity of human experiences (Dudley-Marling and Gurn 2010). Society and culture influence almost every aspect of the human condition, from intelligence to height and weight, in many non-random ways. In the domain of learning and international development, the overreliance on interpreting findings through a prism of normally distributed data contains inherent biases.

Why does that matter? It is potentially misleading to base claims about human learning, and make predictions about it, in an assumption of normal distributions. We have argued that, for learners at the BOP, learning science may be substantially different than for those in more favored populations. One useful approach would be to focus on what BOP learners bring to learning rather than what they are missing. For example, Harper (2012) frames anti-deficit research in U.S. education contexts and Moll, Amanti, Neff, and Gonzalez (1992) describe an orientation that looks at "funds of knowledge" or assets. Each of these approaches supports the notion that there needs to be greater focus on what and how learning takes place at the bottom of the pyramid.

Comparability of learning outcomes across contexts

Comparability is central to global education databases, such as the large-scale data collection carried out by the UNESCO Institute for Statistics (UIS) and OECD. Nonetheless, if the primary goal is comparability, less attention may be paid to the local and cultural validity of the definitions and classifications of learning. Further, the data may become less meaningful and potentially less applicable at the local level. This is a natural and essential tension between universalistic *etic* and context-sensitive *emic* approaches to measurement, and it is particularly relevant to the study of BOP populations. In one well-known example, emic approaches are those that are consciously focused on local cultural relevance, such as local words or descriptors for an "intelligent" person. Etic approaches are those that might define "intelligence" as a universal concept, and try to measure individuals across cultures on that single concept or definition.

Can both comparability and sensitivity to context be appropriately balanced in learning research? Should countries with below average scores be tested on the same scales with countries that have much higher average scores? If some countries, or groups of students, are located at the "floor" of a scale, some would say that the solution is to drop the scale to a lower level of difficulty. Others might say that the scale itself is flawed, and that there are different types of skills that could be better assessed, especially if the variations are evidently caused by cultural, ethnic, linguistic and related variables that lead one to question the test as much as or more than the group that is tested. Yet some say that having different scales for different groups or nations is an unacceptable compromise of the

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 ☐ LE ☐ TYPESET
MS Code : PROS_4404_Wagner ☑ CP ☑ DISK

benchmarks that are sought by international policy makers, such as the Learning Metrics Task Force (Brookings Institution 2013) or the UN Global Education First Initiative (GEFI 2014). If the most important goal is to improve learning at the BOP, how credible are the findings at the tail of the distribution from international (or even national) assessments?

To the extent that comparability can be achieved (and no learning assessment claims perfect comparability), the results allow policymakers to consider their own national or regional situation relative to others. This seems to have most merit when the choices to be made apply to proximal situations, rather than distal ones. For example, consider an African country that has adopted a particular bilingual education program that appears to work well in primary school. If the education minister in a neighboring country believes that the case is similar enough to his or her own national situation, then it makes good sense to compare the scores on, say, primary school reading tests. A more distal comparison might be to observe that a certain kind of bilingual education program in Canada seems to be effective, but to doubt the prospects for applying it in a quite different context in Africa. But proximity is not always the most pertinent feature; for example, in the United States and Japan rivalries between educational outcomes and economic systems have been a matter of serious discussion and debate over many years (Stevenson and Stigler 1982). In a more recent example, senior officials in Botswana were interested in knowing how Singapore came to score first in mathematics on several LSEAs (Gilmore 2005; see also Sjoberg 2007).

The key issue here is the degree to which it is necessary to have full comparability in learning outcomes, with all individuals and all groups on the same measurement scale. Or if a choice is made not to "force" the compromises needed for a single unified scale, what are the gains and losses in terms of comparability? Can international goals and commensurate statistics be maintained as stable and reliable if localized approaches are chosen over international comparability? The responses to these questions have led to situations where some LMICs may be tempted to participate in international learning assessments, but hesitate because their results may appear to be very low. Or, they may feel that the cost to participate does not add sufficient value to decision-making at the national level (Greaney and Kellaghan 1996). Others may participate because they do not want to be viewed as having benchmarks that are inferior to those used in OECD countries; for a recent discussion of some of these issues, see OECD (2014) and Bloem (2013).

In the end, international research on learning requires some form of comparability, but perhaps in more varied ways than usually considered today. For example, international and regional assessments are aimed specifically at cross-national comparability, while hybrid assessments (Wagner 2011b) are more focused on local contexts and increased validity. The latter try to combine aspects of large-scale and small-scale assessments, and may be thought of as smaller, quicker and cheaper. An early hybrid assessment was UNESCO's Literacy Assessment Project (ILI 2002); later versions may be seen in the early grade reading assessments that have grown in popularity (Gove and Wetterberg 2011). Hybrid assessments offer localized comparability that large-scale assessments do not, and can offer more focused results for improving learning and interventions among poor and disadvantaged populations. Which types of comparability are most important depends on the policy goals desired, as well as timing and resource considerations.

What roles do stakeholders play?

Many stakeholders—including policymakers, ministers of education, community leaders in rural villages, teachers, parents and education specialists—should be held to account for

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 LE TYPESET
MS Code : PROS_4404_Wagner CP DISK

what and how children learn. Journal editors and universities can play a role by requiring that researchers offer more intentional explanations of the representation and inherent implications of the samples they include in published studies. Yet, even today, educational specialists and statisticians in most countries have been the primary "guardians" of learning processes and their importance for school and economic success. One major reason for this restricted access to knowledge about learning is the complexities of the empirical science of learning, as described above.

A second reason is insufficient knowledge—and at times erroneous beliefs—among both parents and children about how important learning and schooling are for their life chances. Much evidence, from many societies, suggests that people in poor communities underestimate the value of learning and schooling; for example, Stevenson and Stigler (1982) compared parental beliefs in the United States, China and Japan. Today, it is more important than ever before to involve multiple stakeholders in educational decision-making. In many countries, the public has become more interested in children's learning and school achievement in comparative perspective, probably due to increasing globalization, the influence of international agencies, the efforts of NGOs, greater community activism and parental interest. Some field studies have involved strong community engagement that has led to governments incorporating findings for policy change; see Bhattacharjea, Wadhwa, and Banerji (2011) in India, and Piper and Korda (2010) in Liberia.

This type of multilevel information exchange is another way of linking science to accountability and expectation. Whose problem is it if a child, teacher, school, district or nation is not performing at a given level of learning? Indeed, how are such expectations even built? Whose expectations should be taken into account? Knowledge about the importance of learning—and how it can be achieved in formal and non-formal settings, and in structured and informal ways—has the potential to break new ground in research, policy development, community and family participation and local ownership. This is nowhere more apparent than at the bottom of the pyramid, where parents and communities are only now becoming more aware of the role learning can play in their children's lives.

Policy conclusions

Research on how to improve learning in low-income countries and in poor and marginalized communities—BOP populations—is, in principle, no more difficult to conduct than similar research in wealthier communities. However, given where most of the scientific (human and fiscal) resources are located (i.e., largely in OECD countries), it can be much less convenient for those with the advanced training needed to do the work. That fact, among others, is why so much remains to be known about learning in BOP contexts.

The way that learning is studied in LMICs, and specifically in BOP populations, could have great scientific significance, for both researchers and education planners. As we move forward from Jomtien, Dakar, and the United Nations MDGs towards the post-2015 development goals, it is clear that social and economic inequalities will persist unless we maintain a serious focus on learning among the poor. In his seminal book on new approaches for reaching BOP consumer markets, Prahalad (2006) challenged corporations to adopt a new philosophy of service delivery for this historically overlooked population. By transforming the way learning is understood in contexts at the bottom of the pyramid, we can begin to understand how to better promote policies that will enhance educational quality and increase the learning consequences among those hardest to reach.

References

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602–614.

Babson, A. (2010). *The place of English in expanding repertoires of linguistic code, identification and aspiration among recent high school graduates in Limpopo Province, South Africa*. Ph.D. dissertation, University of Michigan, Ann Arbor.

Banerjee, A. V., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight poverty*. New York: Public Affairs.

Bartlett, L., & Garcia, O. (2011). *Additive schooling in subtractive times: Bilingual education and Dominican immigrant youth in the Heights*. Nashville, TN: Vanderbilt University Press.

Benavot, A., & Tanner, E. (2007). *The growth of national learning assessments in the world, 1995–2006.* Background paper prepared for EFA Global Monitoring Report 2008. Paris: UNESCO.

Benedict, M., & Hoag, J. (2004). Seating location in large lectures: Are seating preferences or location related to course performance? *The Journal of Economic Education, 35*(3), 215–239.

Bhattacharjea, S., Wadhwa, W., & R. Banerji, R. (2011). *Inside primary schools: A study of teaching and learning in rural India*. New Delhi: ASER. http://img.asercentre.org/docs/Publications/Inside_Primary_School/Report/tl_study_print_ready_version_oct_7_2011.pdf.

Blackledge, A., & Creese, A. (2010). *Multilingualism: A critical perspective*. London: Continuum.

Bloem, S. (2013). *PISA in low and middle income countries*. OECD education working papers, no. 93. Paris: OECD.

Braun, H., & Kanjee, A. (2006). Using assessment to improve education in developing nations. In J. E. Cohen, D. E. Bloom, & M. B. Malin (Eds.), *Educating all children: A global agenda* (pp. 303–353). Cambridge, MA: MIT Press.

Brookings Institution (2013). *Toward universal learning: Recommendations from the Learning Metrics Task Force*. Washington, DC: Brookings Institution Center for Universal Education. http://www.brookings.edu/~/media/Research/Files/Reports/2013/09/learning%20metrics%20task%20force%20universal%20learning/LTMF%20RecommendationsReportfinalweb.pdf.

Bruns, B., Filmer, D., & Patrinos, H. A. (2011). *Making schools work: New evidence on accountability reforms*. Washington, DC: World Bank.

Buchmann, C., & Hannum, E. (2001). Education and stratification in developing countries: A review of theories and research. *Annual Review of Sociology, 27*, 77–102.

Castillo, N. M., & Wagner, D. A. (2014). Gold standard? The use of randomized controlled trials for international educational policy. *Comparative Education Review, 58*(1), 166–173.

Cole, M., Gay, J., Glick, J., & Sharp, D. (1971). *The cultural context of learning and thinking*. New York: Basic Books.

Cortina, R. (Ed.) (2014). *The education of indigenous citizens in Latin America*. Clevedon, UK: Multilingual Matters.

DBE [Department of Basic Education], South Africa (2013). *Report on the Annual National Assessment of 2013: Grades 1 to 6 & 9*. Pretoria: DBE. http://www.education.gov.za/LinkClick.aspx?fileticket=Aiw7HW8ccic%3D&tabid=36.

Dudley-Marling, C., & Gurn, A. (Eds.) (2010). *The myth of the normal curve*. New York: Peter Lang.

Engel, L. C., & Feuer, M. J. (2014). Five myths about international large-scale assessments. *Quality Assurance in Education, 22*, 18–21.

GEFI [Global Education First Initiative] (2014). *Global education first initiative: An initiative of the United Nations Secretary-General*. http://www.globaleducationfirst.org/files/GEFI_Brochure_ENG.pdf.

Gilmore, A. (2005). *The impact of PIRLS (2001) and TIMSS (2003) in low- and middle-income countries: An evaluation of the value of World Bank support for international surveys of reading literacy (PIRLS) and mathematics and science (TIMSS)*. New Zealand: IEA.

Gove, A., & Wetterberg, A. (2011). *The early grade reading assessment: Applications and interventions to improve basic literacy*. Research Triangle Park, NC: RTI.

Greaney, V., & Kellaghan, T. (1996). *Monitoring the learning outcomes of education systems*. Washington, DC: World Bank.

Gurn, A. (2010). Conclusion: Re/visioning the ideological imagination in (special) education. In C. Dudley-Marling & A. Gurn (Eds.), *The myth of the normal curve* (pp. 241–256). New York: Peter Lang.

Harper, S. R. (2012). *Black male student success in higher education: A report from the National Black Male College Achievement Study*. Philadelphia: University of Pennsylvania, Center for the Study of Race and Equity in Education.

Heinrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral Brain Science, 33*(2–3), 85–135.

ILI [International Literacy Institute] (2002). *Towards guidelines for the improvement of literacy assessment in developing countries: Conceptual dimensions based on the LAP project*. Philadelphia: ILI/UNESCO.

Kagan, J., Moss, H. A., & Sigel, I. E. (1963). Psychological significance of styles of conceptualization. *Monographs of the Society for Research in Child Development, 28*(4), 73–112.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.

Korinek, K., & Punpuing, S. (2012). The effect of household and community on school attrition: An analysis of Thai youth. *Comparative Education Review, 56*(3), 474–510.

Kremer, M., & Holla, A. (2009). Improving education in the developing world: What have we learned from randomized evaluations? *Annual Review of Economics, 1*, 513–542.

Li, J. (2003). U.S. and Chinese cultural beliefs about learning. *Journal of Educational Psychology, 95*, 258–267.

Lu, Y., & Treiman, D. J. (2011). Migration, remittances and educational stratification among Blacks in Apartheid and Post-Apartheid South Africa. *Social Forces, 89*(4), 1119–1143.

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

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132–141.

OECD (2014). *PISA for development*. Paris: OECD. http://www.oecd.org/pisa/aboutpisa/pisafordevelopment.htm.

Pigozzi, M. J., Carrol, B., Hayden, J., & Ndaruhutse, S. (2014). Fragile and conflict-affected situations. In D. A. Wagner (Ed.), *Learning and education in developing countries: Research and policy for the post-2015 UN development goals* (pp. 58–73). New York: Palgrave Macmillan.

Piper, B., & Korda, M. (2010). *Early grade reading assessment (EGRA) plus: Liberia*. Program evaluation report: Prepared for USAID/Liberia. Research Triangle Park, NC: RTI International.

Prahalad, C. K. (2006). *The fortune at the bottom of the pyramid: Eradicating poverty through profits*. Upper Saddle River, NJ: Wharton School Publishing.

Saito, M. (2011). *Gender equality in education: Looking beyond parity*. Paris: UNESCO IIEP. http://www.sacmeq.org/sites/default/files/sacmeq/publications/7_1_1_5_mioko_saito_eng_version_2012_02_15.pdf.

Sjoberg, S. (2007). PISA and 'real life challenges': Mission impossible? In S. T. Hopmann, G. Brinek, & M. Retzl (Eds.), *PISA according to PISA: Does PISA keep what it promises?* (pp. 203–224). Vienna: LIT Verlag. http://folk.uio.no/sveinsj/Sjoberg-PISA-book-2007.pdf.

Stevenson, H. W., & Stigler, J. W. (1982). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit.

UNESCO (2003). *Education for all global monitoring report: Gender and education for all—The leap to equality*. Paris: UNESCO.

UNESCO (2010). *Education for all global monitoring report: Reaching the marginalized*. Paris: UNESCO.

United Nations (2000). *United Nations millennium declaration*. Resolution adopted by the General Assembly. United Nations A/RES/55/2. http://www.un.org/millennium/declaration/ares552e.htm.

Wagner, D. A. (1983). Rediscovering 'rote': Some cognitive and pedagogical preliminaries. In S. Irvine & J. W. Berry (Eds.), *Human assessment and cultural factors* (pp. 179–190). New York: Plenum.

Wagner, D. A. (1993). *Literacy, culture and development: Becoming literate in Morocco*. New York: Cambridge University Press.

Wagner, D. A. (2010). Quality of education, comparability, and assessment choice in developing countries. *Compare: A Journal of Comparative and International Education, 40*(6), 741–760.

Wagner, D. A. (2011a). What happened to literacy? Historical and conceptual perspectives on literacy in UNESCO. *International Journal of Educational Development, 31*, 319–323.

Wagner, D. A. (2011b). *Smaller, quicker, cheaper: Improving learning assessments in developing countries*. Paris, Washington: UNESCO IIEP and EFA Fast Track Initiative of Global Partnership for Education.

Wagner, D. A. (Ed.) (2014). *Learning and education in developing countries: Research and policy for the post-2015 UN Development Goals*. New York: Palgrave Macmillan.

Wagner, D. A., Murphy, K. M., & de Korne, H. (2012). *Learning first: A research agenda for improving learning in low-income countries*. Washington, DC: Brookings Institution.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research, 47*, 1–64.

Journal : Small 11125 Dispatch : 29-10-2014 Pages : 12
Article No. : 9328 LE TYPESET
MS Code : PROS_4404_Wagner CP DISK

Daniel A. Wagner (United States) is UNESCO Chair, professor of education, and director of the International Literacy Institute and International Educational Development Program at the University of Pennsylvania. He has extensive experience in national and international educational issues, and has served as an advisor to UNESCO, UNICEF, the World Bank, USAID, DFID, and other agencies on international development issues. He is a fellow of the American Psychological Association, the American Anthropological Association, and the American Educational Research Association.

Nathan M. Castillo (United States) is in the Ph.D. program at the Graduate School of Education and a research associate for the International Literacy Institute at the University of Pennsylvania. Before coming to Penn, he was based at Education Development Center's International Development Division in Washington, DC. From 2006 to 2008, he was a health education coordinator, while serving with the United States Peace Corps in Guatemala. He holds a master's degree in International Education Policy from Harvard University.