ACLA WORKING PAPER SERIES

The 5D model: towards a more comprehensive approach for improving education

ILJA CORNELISZ
CHRIS VAN KLAVEREN

Working Paper 20212

https://acla.amsterdam/workingpapers-wp20212/

AMSTERDAM CENTER FOR LEARNING ANALYTICS
van der Boechorststraat 7
1081 BT, Amsterdam
April, 2021

The 5D model: towards a more comprehensive approach for improving education

Chris van Klaveren Ilja Cornelisz

April 2021

Abstract

Human capital theory, the dominant paradigm in the economics of education, suggests that education and training are investments that improve individual and societal outcomes. Yet, this is largely a black-box demand-side approach, providing little guidance for educators and policy makers on how to organize educational offerings. For this purpose, empirical production function studies estimate the marginal contributions of educational inputs in promoting cognitive and non-cognitive skills. This endeavor has given impetus to randomized trials in education and the rise of evidence-based education. This focus on experiments has met with criticism, in part due to relatively disappointing results and a lack of sustainable improvements in education policy and practice.

This paper introduces an iterative 5D model as a more comprehensive approach for improving education. This model acknowledges the importance of causal evaluations in determining the intervention effectiveness. Yet, causal evaluations are placed in a broader framework and are scaled equally in terms of importance with other important stages. Proper diagnosis, for example, ensures that interventions can be effectively targeted, such that conventional scattergun approaches are avoided. The practical intuition of the 5D model is outlined by linkage to a stylized teacher shortage example.

1. Introduction

Evidence-based policy acknowledges that students deserve the best educational programs rooted in the most rigorous evidence available (Slavin, 2002). Notwithstanding different perspectives regarding what the objectives of education ought to be, the dominant paradigm in the economics of education is human capital theory which suggests that education and training are investments that improve individual and societal outcomes. Most notably, human capital theory identifies education as increasing the marginal productivity of labour, which in turn drives earnings and economic growth. With these propositions, human capital theory provided an economic justification for public investment in education and in expanding educational opportunities (Marginson, 2019). The theoretical foundations of human capital theory, as laid out by Mincer (1958), Schultz (1959), and Becker (1964), have fueled a thriving research programme in the economics of education and – correspondingly- many thousands of empirical studies.

Human capital theory is inherently a demand-side approach to understanding investments in education and a 'black box' approach to transforming educational inputs into outputs. What this means is that it remains largely unclear what the internal mechanisms are underlying particular educational investments. We know that certain inputs go into the 'box' (e.g. learning materials, human resources) and we can observe particular outputs (e.g. students' skills and performance). Yet, the intricate workings of learning and interactions that take place throughout this process are mostly unobserved. In order to open up the black box of *how* resources are to be allocated in improving educational outcomes, supply-side economics emerged on the scene with the introduction of production-function analysis (Cohn & Geske, 1990). These input-output analyses are based on the assumption that education corresponds to some production technology involving all kinds of educational inputs (e.g. teachers, curriculum, peers) and outputs (e.g. standardized test scores, wages). The results of these econometric analyses should provide policy-makers with an empirical basis for deciding on the most productive 'mix' of inputs. Over the past decades, human capital theory and education production function analyses have become dominant in policy and public thinking, with policies aimed explicitly at accumulating human capital through investing in specific educational inputs.

Allocating scarce public resources to education policies raises the important question of whether these means are invested both effectively and efficiently. Educational effectiveness is determined by whether a specific mix of inputs has a positive effect on educational outputs. Educational efficiency, instead, refers to either maximizing the output that can be generated by a given set of inputs or to whether comparable levels of output can be obtained using fewer inputs (Lockheed and Hanushek, 1994). To inform such decision-making processes, empirical research in the economics of education thus seeks to identify causal relationships between educational offerings and a variety of cognitive and non-cognitive outcomes. This endeavor has supported the rise of evidence-based education as a model in which policy-makers and educators use empirical evidence to make informed decisions about the effectiveness and efficiency of educational interventions. The ability to make decisions based on empirical evidence, in turn, supports sustainable innovation in education towards improving overall educational quality. While this appears a straightforward approach, evidence-based policy and practice face several challenges, and the emphasis on outcomes and performing randomized

experiments in education has been subject to criticism (Cooke, 2002; Biesta, 2010; Nelson & Campbell, 2017).

In this chapter, we acknowledge the relevance of evaluating the impact of (policy) interventions on educational outcomes, but argue that conducting randomized experiments is only one piece of the puzzle in systematically and sustainably improving education. Thus, we suggest a more comprehensive approach; referred to as the 5D model. The 5D model sets out to provide a five-step approach for iteratively addressing educational issues. In this model, suggestions for improving educational practice and policy are the result of a sequential procedure of *Detect*, *Diagnose*, *Design, Determine* and *Decide*. In this model, interventions (*Design*) are carefully crafted, implemented and targeted based on the outcomes of the preceding phases of signaling the problem at hand (*Detect*) and targeting which interventions would benefit whom *(Diagnose*). Causal evaluations (*Determine*) are crucial for determining what actually works for whom in practice, after which the findings should be carefully interpreted and translated to sustainable changes in educational practice and policy (*Decide*).

We argue that targeting interventions based on valid diagnoses are crucial in that doing so acknowledges the potential of different causes and solutions for addressing educational challenges for different students and schools. Such heterogeneity has important implications for how one should think about assigning interventions and allocating scarce resources efficiently across students over time. These can be decision processes that operate either at the micro level (e.g. deciding which students should receive what type of one-on-one tutoring), or at a meso or macro level (e.g. attracting and retaining more teachers at specific hard-to-staff schools). Ultimately, the 5D model provides a general framework for systematically improving decision-making for educators and policy makers alike.

This paper first briefly introduces the increased role of evidence-based education and some of the challenges it is exposed to (Paragraph 2). It then introduces the 5D model as a more comprehensive approach for thinking about sustainable improvements in education (Paragraph 3). This is followed by a stylized example in which we portray how this approach might facilitate effective developments in addressing teacher shortage (Paragraph 4). The paper concludes by discussing the potential and pitfalls of this approach and how future avenues of research can address these (Paragraph 5).

2. The Quest for Evidence Based Education

Empirical evaluations in education are not new. Ever since data such as student test scores and educational attainment have been available, educators and policy-makers have tried to assess whether their educational offerings are yielding the desired results. Yet, it is intrinsically difficult to find out whether a new approach or policy works better than what's already in place. Simply comparing results before and after a new educational intervention is rolled out, while tempting, can be very misleading, as other unmeasured (and often unmeasurable) factors may have changed at the same time (Goldacre, 2013).

Observing that it is often unclear whether the many developments that take place in educational thinking and practice are better, or worse, than the regimes they replace has given impetus to the rise of evidence-based education (Davies, 1999). Yet, evidence-based education does not emerge instantaneously. For example, it requires methodological approaches that facilitate rigorous causal evaluations in educational practice. Such evaluations often involve random assignment to avoid unobserved differences can provide alternative explanations when comparing different educational offerings. Organizing such successful large-scale field experiments in education requires close collaboration between educational practitioners and researchers, for example in so-called education labs. Last but not least, the evidence-base generated through such research needs to be shared with the larger educational community to facilitate overall innovation throughout the system.

Partly in response to observing that -currently- educational practice and policy is often not evidence-based (Goldacre, 2013; Slavin, 2004), a so-called 'What Works' movement has emerged. This is observed most notably in Anglo-Saxon countries, such as the Education Endowment Foundation (EEF) in the UK, cc, and the What Works Clearinghouse (WWC) in the US. This increased emphasis on an evidence-based education paradigm has also been met with criticism from different stakeholders. For example, educational practitioners point out that educational practice should be evidence-informed, rather than evidence-based. Advocates of evidence-informed practice emphasize that scientific evidence is just one of several factors that influence decision-making by educators and emphasize the need to apply professional judgment, rather than being driven solely by research evidence or data (Nelson & Campbell, 2017). Others take an even more critical stance, in claiming that evidence-based education is both unfeasible and undesirable, emphasizing that it serves to only distract from the real values and purposes of education (Biesta, 2010). In summarizing, Cook (2002) identifies several philosophical, practical, conceptual and methodological arguments used by educational evaluators to oppose randomized trials in education. The fact that systematic reviews of experimental evidence in education often point to relatively small gains –if any- and that results often highlight considerable heterogeneity across educational contexts, (non-)cognitive domains and subpopulations of learners further serves to undermine broad-scale support for advocating randomised trials in education..

This chapter adds to the discourse on the potential and pitfalls of experimental evidence in education. It does so by making explicit that experimental evidence in education 1) is only one important piece of the puzzle in improving education, 2) can only yield promising results if educational questions are adequately detected, diagnosed, and intervened upon and 3) will only serve to promote sustainable development if results are adequately interpreted and fed back to this iterative process of improving educational quality. The next chapter integrates these observations into a 5-step model that can serve as a guideline in making explicit the prerequisites for moving towards sustainable improvement in education; thereby incorporating evidence-based research.

3. An Iterative Approach Towards Sustainable Improvement in Education

The point of departure for the proposed iterative approach –what we refer to as the 5D model (see Figure 1)- is that educational outcomes can only be structurally improved if each sequential phase is carefully addressed; thereby utilizing both quantitative and qualitative methodologies from a variety

of academic disciplines. Below, we do not present an exhaustive list of what these analytical approaches and instruments in each phase might be, yet we touch upon some for illustrative purposes.

Detect: Observing the issue at hand

The first step towards improvement is to **D**etect whether a particular issue in education (e.g., dropping reading skills) is adequately observed ("*Does the problem exist?*"). This phase can thus involve careful analysis of monitoring instruments, such as surveys of pupils' scholastic performance, administrative data, and well executed expert interviews. The main objective in this phase is to gain rich insight into the educational phenomenon under investigation as to ensure a clear exposition of what exactly needs to be addressed and resolved in the following phases of the 5D approach.

Diagnose: Defining target population

Next, careful and detailed **D**iagnose should take place ("*Who experiences this problem when and why?*") in order to pave the way for effective action. Diagnosing which (groups of) students, teachers or schools ought to be targeted when implementing changes in educational programs and policy is crucial in ensuring the effectiveness and efficiency of such interventions. This phase benefits from deploying diagnostic tools (e.g., inventories, checklists) that are validated for identifying the educational issue at hand. In addition, recent developments in computational modeling and so-called statistical machine learning techniques create data-driven opportunities for early risk detection; thereby enabling a timely and targeted intervention. Furthermore, rigorous diagnosis sheds light on the underlying cause and mechanisms of why a particular issue is observed for whom and thus what intervention(s) ought to be considered.

Design: Developing the intervention

A detailed understanding of a problem -and to whom it applies- then enables setting up a well-targeted intervention developed to directly address the issue at hand. This **D**esign phase ("*Who receives what treatment and when*?") contains several elements, such as considering off-the-shelf interventions, adapting approaches that seem to have been effective in other settings (e.g., countries) to the educational context at hand, or even configuring completely new interventions. Small-scale feasibility studies can be conducted to learn what obstacles and practical considerations arise such that treatments are implemented correctly and consistently (i.e., treatment fidelity). Ideally, then, this phase not only determines who will be targeted with what specific action, but also generates essential knowledge on what is required for successful implementation in practice.

Determine: Evaluating what works

Then, to correctly **D**etermine whether the proposed action is indeed successful ("*What works for whom*?"), a rigorous evaluation exploiting random assignment or some other high-quality quasi-experimental designs should take place. In particular, the educational outcomes of those participating in the evaluation ought to be carefully scrutinized in the next phase. Evaluation studies conducted in this phase should be adequately powered, such that the relative effectiveness of different treatment arms can be estimated empirically with high precision; thereby allowing for potential heterogeneity

in treatment effects across subgroups of students, teachers and schools. This asks for large-scale studies in which the unit of randomization (e.g., student-level, school-level) is appropriate for the intervention and delivers the required level of statistical power. If performed successfully, the determine phase will provide an unbiased and precise answer to what has worked for whom.

Decide: Changing policy and practice

Last but not least, the results of this evaluation ought to be shared, discussed, and interpreted such that educators and policy-makers can **D**ecide ("*How to proceed based on these results*?") and make an evidence-informed decision on how to incorporate lessons learned from the evaluated interventions in educational practice and policy. If the previous phases have made clear that a particular intervention is highly effective, the decision can be to start thinking about what the implications are -for both educational policy and practice - of across-the-board implementation. Yet, this last part of the process will likely also identify unresolved issues and ideas on how these are to be addressed in future work; thereby facilitating a feedback loop in the 5D model.

Figure 1- is that education can only be structurally improved if a problem (e.g. teacher shortage in disadvantaged schools) is first appropriately **D**etected ("*Does the problem exist?*"). Next, adequate **D**iagnose should take place ("*Who experiences this problem and why?*") to pave the way for a well-targeted theory of action. Understanding a problem and to whom it applies enables setting up an appropriate intervention (**D**esign) designed to directly address the issue at hand. Then, to correctly **D**etermine whether the proposed solution is indeed successful, rigorous evaluation should take place; thereby often resorting to random assignment (or other high-quality quasi-experimental designs). Last but not least, the results of this evaluation ought to be discussed, shared and interpreted such that educators understand how to embed successful policies and programs in their educational practice (**D**ecide). This last part of the process also highlights what unresolved questions there still are and how these can be detected in future work; thereby facilitating a feedback loop in the 5D model.

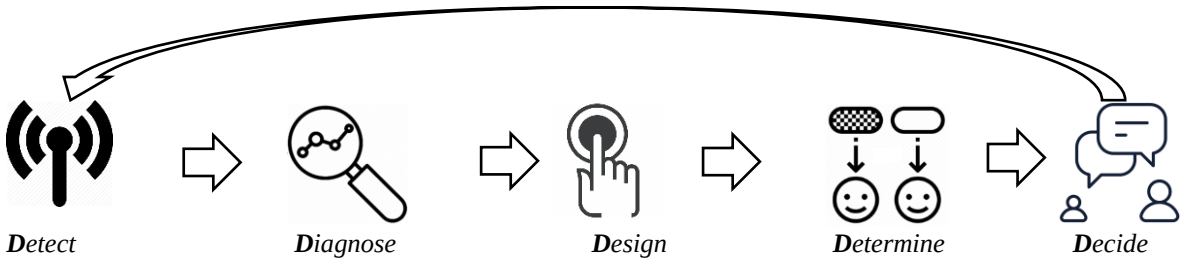

Figure 1: The 5D Model for Iterative Sustainable Improvement in Education

The aforementioned 5 steps all seem fairly trivial and neither of them are -in and of itself- likely to come across as particularly ground-breaking. Therefore, it is relevant to emphasize that it is the overall picture, sequence and interlinkage of activities that this model portrays which can be of significance. For example, there is an abundance of examples in educational policies and interventions that seem to be rolled out (Design) in direct response to an apparent problem (Detect). Often more than not, these interventions (e.g. financial bonuses for teachers) are implemented after observing symptoms

(e.g. high teacher turn-over in disadvantaged schools), without properly understanding who these teachers are and why it is difficult to retain these teachers in these schools (Diagnose). Failing to act upon the outcome of a sound diagnostic process increases the risk of implementing an unsuccessful intervention. Besides, such policy measures are often not implemented in a way that even allows for causal evaluation (Determine) and this absence of robust findings on what has worked for whom further compromises the ability to learn which changes should be embedded structurally and how (Decide). What this highlights is that all 5 stages in the 5D model for improving education should be addressed sequentially and that omitting 1 is likely to put at risk the others. Whereas success in each part of the 5D model is to some extent (in)directly contingent on success in all other parts, particular emphasis is put on a solid diagnostic process.

In this diagnostic phase, it is important that rich (multimodal) data are incorporated from the detect phase. In this phase, answers should be generated to questions such as who is affected by a particular challenge in education, why this is the case and –for example- who is likely to run the risk of being affected by this situation in the foreseeable future. In the absence of the proper diagnosis, interventions are implemented without targeting the right population with the most promising solution. This is akin to a "scattergun approach" common to how many educational policies have been rolled out in the past, which entails scarce educational resources to be allocated both inefficiently and ineffectively. In contrast, departing from a detailed diagnostic phase also opens up pathways to propose different interventions to different target groups of recipients, thereby addressing the inherent heterogeneity observed in many educational settings.

In addition to improving resource allocation in education through effective targeting, the diagnostic phase is also vital in enabling success in later stages of the 5D model. An intervention that results from a proper diagnosis is more effective in articulating the proposed theory of change and –as such- causal evaluations will be better equipped in demonstrating what works for whom. This then facilitates a high-quality discussion and interpretation of the results in which lessons learned can be determined. In this final stage, findings –if positive- should be solidified into educational policy and practice, together with feeding back to the 5D model things that needs to be observed and adjusted in a next iteration.

4. A Stylized Example: Teacher Shortage

It has become widely accepted that teachers are crucial in facilitating student learning. A vast body of empirical work has examined the importance of teachers in explaining observed differences in student performance (Hanushek, 2011). Whereas important differences in teacher effectiveness exist, it is more difficult to empirically pinpoint what specific teacher characteristics contribute to teacher quality (Coenen et al., 2017). Yet, it is clear that there is much at stake for schools to attract and retain enough high-quality teachers. In practice, schools often fail in realizing this objective. For example, in the US, as in other countries, it is found that teacher shortage is real, large and growing (Garcia & Weiss, 2019). It is thus concerning to observe that many countries face severe teacher shortages. Yet, simply 'pouring more teachers into the system' in response, without clearly understanding the underlying causes, is likely not effective in resolving this issue (Ingersoll and Smith, 2003). To show the potential of addressing this matter more comprehensively, we show how teacher shortage can be addressed by applying the 5D model below.

Detect: observe clear patterns of teacher retention

In evaluating patterns of teacher shortage, empirical studies, conducted mostly in Anglo-Saxon contexts point out that in these countries teacher shortage does not seem to result from a lack of teacher program graduates, but in the inability to retain teachers in the education system (Ingersoll, 2001, 2002; Boyd et al., 2006; Sass et al., 2012; Buchanan et al., 2013 and White et al., 2006). In addition to overall concerns of educational quality, teacher shortage also induces specific educational equity issues, since many studies show that particularly schools with a large proportion of students coming from disadvantaged backgrounds (Garcia & Weiss, 2019), belonging to ethnic minorities and/or facing learning difficulties face much higher teacher turnover rates (Dupriez et al., 2016). Similarly, Darling-Hammond et al. (2016) find that the proportion of uncredentialed teachers in high-minority schools in California to be more than twice that in low-minority schools.

It is analyses like these that are required to gain insight into the magnitude and dimensions of the problems associated with teacher shortage. Further zooming in on teacher characteristics associated with a higher propensity to exit the profession can reveal that both the youngest and oldest teachers are most likely to dropout (Borman & Dowling, 2008; Boyd et al., 2011). In addition, it is important to detect whether turnover rates are higher for teachers in specific fields such as mathematics, science and special education (Podolsky et al., 2016).

To summarize, this phase of the 5D model should indicate to what extent teacher shortage is prevalent and should distinguish between a lack of supply and/or a lack of retention. Also, the detection phase should shed light on what types of schools and regions are more likely to suffer from this phenomenon, and whether there are particular subgroups of teachers more prone to exit the teaching profession. These results can then be useful in guiding the diagnostic process.

Diagnose: understand why specific teachers leave particular schools

Understanding the underlying reasons for teacher shortage is crucial in determining which intervention or policy measure can benefit which specific schools. Detailed data analysis can facilitate this process of arriving at the correct diagnosis. For example, in understanding the role of several of these factors, Garcia & Weiss (2019) compare teachers who stayed in teaching with those who quit teaching, and observe that larger shares of staying teachers had received early support in the form of an assigned mentor, induction programs, engaged in useful subject-specific professional development activities, worked in highly cooperative environments, and experienced more autonomy. Similarly, survey results from the National Center of Education Statistics (NCES) indicate that a majority of leaving teachers indicate that a salary increase could have induced them to stay in the profession, while also pointing out important differences between teachers who either move to a different school versus those who leave the education system altogether (Ingersoll, 2003). Adding to the importance of acknowledging detailed data analysis, results in Dupriez et al. (2016) point to a very close relationship between job conditions over the first year in the profession and exit rates. Based on survey data from teacher education graduates, Perryman and Calvert (2020) find that for these novice teachers workload was the most frequently cited reason for already having left the profession.

The diagnostic process can benefit greatly from performing qualitative studies; for example for understanding which formal and informal school structures contribute to a supportive school environment that encourages novice teachers to remain in the school (Podolsky et al., 2016). Similarly,

the qualitative data in Perryman and Calvert (2020) point out that in addition to just the quantity of workload, the nature of the workload is also an important factor to consider in understanding why young teachers leave the profession.

To summarize, the diagnostic process will reveal the underlying mechanisms for why different groups of teachers decide to choose for, leave or stay in the teaching profession. Similarly, it can highlight important heterogeneities across regions, school types, or even subjects. Understanding why teacher shortages arise across these different dimensions is important in deciding which interventions are considered most promising.

Design: developing and targeting interventions to specific groups of teachers and schools

There exists a vast body of research on improving teacher recruitment and retention. Podolsky et al. (2016) identify five major factors that can influence teachers' decisions to enter, stay in, or leave the teaching profession, generally, and high-need schools, specifically:

1. Salaries and other compensation.
2. Preparation and costs to entry.
3. Hiring and personnel management.
4. Induction and support for new teachers.
5. Working conditions.

Targeting the right schools and teachers is an important prerequisite for the effectiveness of an intervention aimed at resolving teacher shortage. If important heterogeneities exist across schools and teachers in the underlying causes of high exit rates, this can imply that several interventions ought to be considered. For example, it could be that differential pay is to be preferred in providing financial incentives for retaining STEM teachers if these are the teachers that have relatively high-paying alternative job opportunities outside of teaching. In contrast, if results for teachers working in special-needs education point to excessive workload as a major determinant in leaving the profession, effective interventions ought to address this by for example offering classroom assistants. Similarly, results in Ingersoll (2003) suggest that for novice teachers improving the lack of administrative support, poor student discipline and motivation, and lack of participation in decision-making proof to be more successful improvement efforts.

The *Diagnose* and *Design* stages of the model ensure that "scattergun approach" is avoided in which a one-size-fits-all intervention is rolled out across schools and teachers. Instead, in exploiting the insights from an elaborate diagnostic process, this phase is designed to develop and target specific interventions to specific teachers and schools. After deciding on which interventions are considered most promising for whom, these interventions can be developed, implemented and the actual effectiveness should then be determined through rigorous empirical evaluations.

Determine: evaluate which approach is most successful for which teachers and schools

The stages of Detect and Diagnose have guided the way to decide on which interventions are targeted to whom (Design). Yet, a focus on rigorous experiments evaluating replicable programs and practices is essential to build confidence in educational research among policymakers and educators (Slavin, 2002). For example, Glazerman et al. (2013) evaluate the effect of offering bonuses for the highest-performing teachers to move into schools serving the most disadvantaged students using a multi-site randomized field experiment. They find that the transfer incentive successfully attracted high value-added teachers and had a positive impact on test scores in targeted elementary classrooms.

Furthermore, the results point out that this transfer incentive also had a positive impact on teacher-retention rates during the pay-out period and that retention rates dropped to comparable levels after the last pay-out of the incentive. Empirical results like these inform policy-makers on what the actual impact and sustainability of an intervention is. If field experiments involving explicit randomization is not feasible, high-quality quasi-experimental approaches can be considered such as regression-discontinuity or difference-in-differences approaches or –alternatively- high-quality matched designs.

In addition to the evaluation design, the Determine stage should also address treatment fidelity (i.e. whether the intervention in a research study is conducted consistently and reliably) as to better understand the results. For example, if the causal results point out that the intervention did not improve educational outcomes, it is important to know whether this can be attributed to the intervention itself or to the quality of how the intervention was implemented. For example, Springer et al. (2016) evaluate the retention bonus program for effective teachers in Tennessee's Priority Schools using a fuzzy regression discontinuity approach and find some null findings that might be the result of concerns, including the timing of the application process and observed noncompliance in bonus distribution.

In summary, the main objective of this stage is to perform a high-quality causal impact evaluation, using a (quasi-)experimental research design. Furthermore, treatment fidelity analysis of the interventions evaluated should be carried out to better attribute results to the intervention itself or to how it was implemented in the context of the research study. Together, this information fuels the final stage of improving education which concerns itself with how desirable changes can be systematically embedded in educational policy and practice.

Decide: translate findings to educational practice and policy

The objective of this final phase of the 5D model is to synthesize all that has been learned in the preceding four phases and how this new knowledge translates to potential changes in educational practice and policy. First, it should be clear what the external validity is of the empirical evaluations conducted. Most often it is the case that schools and teachers participating in randomized trials need to volunteer and provide some sort of consent. This can be relevant when deciding on translating these internally valid results to across-the-board changes in educational policy and practice.

Another important element in this respect is that the results observed in terms of effectiveness of the different interventions should be augmented with the costs associated. Cost-effectiveness analysis enables comparing alternatives for their relative costs and results and provides guidelines on which of the alternatives evaluated to provide the most impact relative to cost (Levin, 2001). Related to this, it should be considered whether interventions are scalable in terms of both the physical and financial resources required for implementing these changes. For example, Thompson and Goe (2009) review models for effective professional development and point to the importance of developing scalable systems of delivery for large-scale effective adoption of such professional development trajectories.

To conclude, this final stage translates lessons learned from the previous stages of the 5D model to actual recommendations for sustainable changes in educational practice and policy. For this, it is –

among other aspects- important to understand how results would generalize when adopting across-the-board policy changes, what the cost-effectiveness is of the different alternatives and to what extent effective interventions are scalable. Lastly, this stage will also identify gaps in the required knowledge for alleviating teacher shortage and this is information is important feedback for the next iteration in the 5D model.

5. Conclusion and Discussion

The evidence-based policy movement has arisen to ensure that policy is made based on empirical scientific literature that provides rigorous evidence with the objective that the intended policy works. This chapter recognizes the importance of rigorous empirical evaluations, but argues that evidence-based education policy could be designed more sustainably and effectively if other important research-process stages are also recognized. Evaluation studies, for example, often evaluate the effectiveness of a particular intervention to solve a certain problem, without providing a proper diagnosis that the offered solution by the intervention is indeed the underlying cause of the problem.

This may explain why many evaluation studies with solid methodological designs find insignificant results; it is like a `scattergun approach' is adopted which increases the probability that the evaluated intervention is ineffective, and when persons need different treatments related to the cause of the problem, then the average treatment effect may not provide an accurate indication of the effectiveness of the intervention. The latter is because the intervention may be effective when it is specifically targeted at persons who experience the problem that the intervention aims to solve.

This paper proposes an iterative *5D Model* to provide a more comprehensive and sustainable approach for improving education, The model describes 5 stages that must be followed in succession to optimize the evaluation as a whole. The Detection stage shows that the problem is existent and the Diagnostic stage then identifies the underlying cause of the problem ("*Who experiences this problem and why?*"). Given that existence and the root of the problem are revealed, one or multiple interventions can be developed and implemented (Design) with the objective to solve the underlying problems experienced. The Determination stage rigorously evaluates the effectiveness of the implemented interventions. This stage concurs with the evidence-based policy movement, in the sense that both have the objective that policy is backed up with rigorous empirical evidence. Finally, the Discussion stage interpret the overall results such that educators and policy makers understand how to embed successful policies and programs in the educational practice. This stage also pinpoints the unresolved problems, which should then (once again) be detected, diagnosed etc.

The aforementioned 5 stages seem perhaps fairly trivial, also because neither of them are likely to come across as particularly ground-breaking. The relevancy and strength come from the overall picture, the sequence, and the interlinkage of stages that this model portrays. For example, providing an argument that a particular outcome is too low and rigorously evaluating a certain policy that could potentially improve this outcome is less optimal than showing solid evidence that the outcome considered is too low, providing a thorough diagnosis of the problem enabling effective development of interventions and rigorous evaluation of these interventions. When the emphasis is only on the rigorous evaluation of certain policies, then this might seriously hamper the effective and sustainable improvement of education. The chapter at hand provides a stylized example of teacher shortage to discuss all stages in more detail.

Given that policy is usually determined for the short- or mid-term and are prone to political forces, the iterative character of the *5D* model is important to discuss. In high iterative settings, the *5D* model

functions well because the iterative character ensures that problem existing and diagnosis checks occur frequently and as a response interventions can be updated and (again) rigorously evaluated, and based on the results a discussion can take place. However, when policies are (partly) politically driven and have a short- or mid-term focus then it may prevent that policies are iteratively updated and improved, and as such prevent a long-term focus to achieve structural and sustainable educational improvements. This discrepancy between achieving sustainable educational improvements and restrictions faced due to political and organizational factors should never prevent that one of the mentioned 5 stages is neglected. Effective educational improvement thus requires that the interplay between these stages is coordinated effectively.

REFERENCES

Becker, G. 1964. Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education. Chicago, IL: University of Chicago Press

Biesta, G. J. (2010). Why 'what works' still won't work: From evidence-based education to value-based education. *Studies in philosophy and education*, *29*(5), 491-503.

Borman, G. D. & Dowling, N. M. (2008) Teacher attrition and retention: A meta-analytic and narrative review of the research, Review of Educational Research, 78(3), 367–409.

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., Michelli, N. M. & Wyckoff, J. (2006) Complex by design. Investigating pathways into teaching in New York City schools, Journal of Teacher Education, 57(2), 155–166.

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S. & Wyckoff, J. (2011) The influence of school administrators on teacher retention decisions, American Educational Research Journal, 48 (2), 303–333.

Buchanan, J., Prescott, A., Schuck, S., Aubusson, P., Burke, P. & Louviere, J. (2013) Teacher retention and attrition: Views of early career teachers, Australian Journal of Teacher Education, 38(3), 112–129.

Coenen, J., Cornelisz, I., Groot, W., Maassen van den Brink, H., & Van Klaveren, C. (2018). Teacher characteristics and their effects on student test scores: A systematic review. *Journal of economic surveys*, *32*(3), 848-877.

Cohn, E., & Geske, T. G. (1990). Production and cost functions in education. The Economics of Education, 3rd Edition, Pergamon Press, Oxford.

Cook, T. D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational evaluation and policy analysis*, *24*(3), 175-199.

Darling-Hammond, L., Furger, R., Shields, P. M., & Sutcher, L. (2016). Addressing California's Emerging Teacher Shortage: An Analysis of Sources and Solutions. *Learning Policy Institute*.

Davies, P. (1999). What is evidence-based education?. *British journal of educational studies*, *47*(2), 108-121.

Dupriez, V., Delvaux, B., & Lothaire, S. (2016). Teacher shortage and attrition: Why do they leave?. *British Educational Research Journal*, *42*(1), 21-39.

García, E., & Weiss, E. (2019). The Teacher Shortage Is Real, Large and Growing, and Worse than We Thought. The First Report in" The Perfect Storm in the Teacher Labor Market" Series. *Economic Policy Institute*.

Glazerman, S., Protik, A., Teh, B. R., Bruch, J., & Max, J. (2013). Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment. NCEE 2014-4003. *National Center for Education Evaluation and Regional Assistance*.

Goldacre, B. (2013). Building evidence into education.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education review*, *30*(3), 466-479.

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 37*(3), 499–534.

Ingersoll, R. M. (2002). The teacher shortage: A case of wrong diagnosis and wrong prescription. *NASSP Bulletin, 86*(631), 16–30.

Ingersoll, R. (2003). Is there really a teacher shortage?.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational leadership*, *60*(8), 30-33.

Levin, H. (2001). Waiting for Godot: cost-effectiveness analysis in education. *New directions for evaluation*, *2001*(90), 55-68.

Lockheed, M. E., & Hanushek, E. A. (1994). Concepts of educational efficiency and effectiveness. Washington, DC: World Bank.

Marginson, S. (2019). Limitations of human capital theory. *Studies in Higher Education*, 44(2), 287-301.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of political economy*, *66*(4), 281-302.

Nelson, J., & Campbell, C. (2017). Evidence-informed practice in education: meanings and applications. *Educational REsEaRch*, *59*(2), 127-135.

Perryman, J., & Calvert, G. (2020). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*, *68*(1), 3-23.

Podolsky, A., Kini, T., Bishop, J., & Darling-Hammond, L. (2016). Solving the Teacher Shortage: How to Attract and Retain Excellent Educators. *Learning Policy Institute*.

Sass, D. A., Flores, B. B., Claeys, L. & Perez, B. (2012) Identifying personal and contextual factors that contribute to attrition rates for Texas public school teachers, Education Policy Analysis Archives, 20(15), 1–26.

Schultz, T. W. (1959). Investment in man: An economist's view. *Social service review*, *33*(2), 109-117.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational researcher*, *31*(7), 15-21.

Slavin, R. E. (2004). Education research can and must address “what works” questions. *Educational researcher*, *33*(1), 27-28.

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, *38*(2), 199-221.

Thompson, M., & Goe, L. (2009). Models for effective and scalable teacher professional development. *ETS Research Report Series*, *2009*(1), i-35.

White, P., Gorard, S. & See, B. H. (2006) What are the problems with teacher supply?, Teaching and Teacher Education, 22(3), 315–326.