

Impact evaluation within the education sector – Current perspectives from Denmark and the USA

Within the education sector, as in other public policy sectors, recent years have been characterized by a rise in demand for “evidence-based” policies and practices. What this means is not always entirely clear, but in most cases it draws on the idea of applying impact evaluation to provide policymakers and educators with knowledge about “what works”. Education evaluation in Denmark – as in the rest of Europe – has to a large extent been carried out using qualitative methods within a discourse that has emphasised a quality assurance purpose. As a result, the demand for more rigorous methods and impact oriented analyses constitutes a real and exciting challenge to the field of education evaluation (and research).

EVA has looked to the USA for inspiration on how to complement qualitative and process oriented approaches with more quantitative and impact oriented methods. This paper presents some of the developments in the methodological approaches over recent years in both the USA and Denmark.

In the USA, concern about lack of evidence of the effectiveness of programmes, policies and educational practices has led to the strategic choice of promoting experimental methods (randomized controlled trials (RCTs)) to provide evidence-based knowledge to educators and policy-makers. One of the main agents has been the What Works Clearinghouse (WWC) founded in 2002 by the Institute of Education Sciences under the auspices of the US Department of Education. The WWC funds RCTs-based impact evaluations of educational interventions and undertakes meta-analyses in order to synthesize findings about impacts of different educational initiatives. This knowledge is circulated to educators through practice-guides.

The founding of the WWC has led to a change in the scope of impact evaluations. Lately, more emphasis has been placed on evaluating the effectiveness of educational methods, like Reading Recovery and Dialogic Reading, than on the “traditional” impact evaluation of large scale programmes and interventions (clearly, the boundaries between educational methods and programmes/interventions are flexible at certain points, e.g. a drop-out prevention programme can be perceived as both).

The WWC has been able to document the positive effects of some initiatives. More often, however, the result of the multimillion dollar investment in RCTs and meta-analyses has

Dato
24.09.2008

Ref
nms/tvj

Side
1/3

been negative, i.e. it has not been possible to prove the effects of many of the education interventions under review. In response, the WWC has been dubbed “The nothing works clearinghouse” by its enemies.

The WWC has also been criticised by other quantitatively orientated researchers for employing a very narrow conception of what constitutes appropriate methods for impact evaluation. This is due to the fact that the WWC almost exclusively gathers evidence from RCTs and is very reluctant even to accept knowledge generated from quasi-experimental studies. Perhaps this very hard methodological stance has been one of the reasons for the WWC’s limited success. Lately, there has, however, been a softening of attitudes towards other kinds of studies, especially quasi-experimental methods like Regression Discontinuity Design, etc.

The approach of the WWC and like-minded agencies is also criticisable from the perspective of realistic evaluation. In the eyes of realistic evaluators like Nick Tilley and Ray Pawson, the WWC exemplifies a player in the field of evidence-based policy, with a very naïve old-school positivist belief in acquiring knowledge about universal laws and paying too little attention to the fact that the context in which an intervention is implemented is decisive for whether it works or not. An initiative that works in inner city schools in Washington DC, doesn’t necessarily work in schools in rural Arkansas, and vice versa. Instead of asking “What works?” the WWC should be asking itself, “What works for whom and under which circumstances?”

Side
2/3

This critique is not necessarily a fair one. The researchers of the WWC are not so naïve to believe that if an intervention is proven effective in several RCTs, it is guaranteed always to work at every school. Their view is that the results of RCTs and meta-analyses provide strong evidence that an intervention is effective, but schools are still advised to monitor the implementation and results of interventions to check whether the expected outcomes are achieved or not.

The development of education evaluation in Denmark has been different to that of the USA. Although in recent years there has been focus on impact evaluation, and Danish education R&D has been criticized by the OECD for its lack of good experimental studies and quantitative research in general, there hasn’t been any real move towards these methods yet. A Danish clearinghouse for education R&D has been established, but it does not promote experimentally based impact studies or evaluations – partly because such studies are practically non-existent, even in a Scandinavian context. The social policy area stands in contrast to this. The Danish national centre for social research has created a Nordic centre for the Campbell collaboration and has started disseminating results from this international collaboration and conducting its own meta-analyses.

Although large public registers give some obvious potentials for conducting experimental and quasi-experimental studies, the education sector in Denmark does not seem ready for this kind of impact evaluation. There is a rather limited focus on what this kind of impact evaluation would mean for policy-formulation and implementation. Paradoxically, it is often decided after a program has been active for a couple of years to try and evaluate its impacts. In many circumstances this presents a lot of methodological problems in the process of constructing a rigorous impact evaluation design.

Another obstacle to impact assessment is the Danish tradition of bottom-up development of new education methods. National programmes and initiatives to combat problems very often only invite local schools and authorities to create their own methods. This way of doing politics encourages creativity and innovation at the local level, but results in a lack of standardisation, mutual learning and generalized knowledge that impact evaluations on a national level might provide.

As an alternative approach to impact evaluation such as the RCTs, realistic evaluation is to some extent employed in schools and by local authorities for small-scale evaluations of local initiatives. However, realistic evaluation has not yet been used for larger or more professional evaluation studies within the education sector (or even other public policy sectors).

EVA is currently gaining its first experiences with both quasi-experimental impact evaluation and realistic evaluation. First signs are promising, but there are some obstacles. Firstly, the lack of earlier studies makes it impossible to find inspiration from good practice examples of how to apply the methods in a Danish context. Secondly, one of the experiences so far is the vital importance of having good measurements of outcomes. In many cases, it is going to take a lot of political will, resources, and coordinated action to secure the gathering of this information, which is an essential part of impact evaluations.