

Exploring Evaluation with John Killeen

Michael White, Policy Studies Institute

In *The Economic Value of Careers Guidance* (Killeen *et al.*, 1992), John Killeen outlined the idea of an evaluation of guidance services to establish their economic impact on clients. Before long, John was carrying out an evaluation, on behalf of the Department for Education and Employment, of the Gateway to Learning pilot programme for unemployed people (Killeen, 1996e). The results of this were sufficiently encouraging to lead to a more ambitious evaluation project (also funded by DfEE) to assess the impacts of adult careers guidance on employed people. This project, on which John and I co-operated, spanned five years from feasibility study to report (Killeen & White, 2000) and was, I believe, a 'first' in two respects. It was the first labour market evaluation in Britain focussing upon a group of employed people, and it was the first study to apply a rigorous econometric evaluation method to careers guidance in any country. The latter comment is made without disrespect to previous attempts. They did not have the benefit of the large developments in evaluation methodology taking place (chiefly in the USA) in the mid-to-late 90s, just in time for our project.

The practical tasks involved in the adult guidance evaluation were formidable. In the absence of any national programme of careers guidance for employed people, there was a frantic search for local schemes, most of which were being funded by Training and Enterprise Councils on soft money. There was then the delicate task of persuading each under-pressure scheme manager to participate, and of negotiating the research procedures, which had to be adapted to how the local scheme operated. John handled most of this side, without which there would have been no evaluation. He also designed the personal interview, which was used to chart educational and employment events in the year following guidance (or, in the case of the non-guidance 'comparison group', in the period since they were recruited to the study). This was to be the key instrument for establishing outcomes and impacts, and all depended on gathering reliable and detailed information. Yet our budget, and considerations of response rate, required that this interview should be restricted to half-an-hour. With great skill and tenacity, John crafted a questionnaire which met this constraint but made no compromises (as I am certain I would have done) in defining the respondents' experiences and resolving the many potential ambiguities in educational and employment transitions.

I have already indicated that the study used a comparison group design. To estimate the impacts of guidance we decided

to combine this with a method, new at the time but now well-established, known as propensity score matching. In essence, this method uses retrospective data to estimate individuals' probability of taking part in the programme under scrutiny - here, the probability of being a client for guidance - and then matches each participating individual with another who, while having *a priori* the same probability of participating, in fact did not do so. This method overcomes the problem which formerly undermined the credibility of matching, namely that only a few characteristics or circumstances could be taken into account. Propensity score matching can take account of any number of background variables.

John saw that clients for adult careers guidance were likely to have distinctive *attitudes* - for instance, to be more dissatisfied with their jobs or more frustrated about further education and training. He argued that these aspects needed to be covered in the initial questionnaires which guidance clients and comparison group members were asked to complete when they were recruited to the study. My instinct was to match only on 'objective' variables such as age, type of job, and prior qualifications. Fortunately, John's view, backed by his unfailingly friendly but firm argumentation, prevailed over mine. In particular, adult guidance clients proved to have extraordinarily low levels of job satisfaction, and this was the most powerful among more than 20 variables which were eventually used in matching.

If job satisfaction had not been used in matching the samples, the evaluation would misleadingly have found that adult guidance leads to subsequent increases in satisfaction, relative to non-participants. But once we selected the non-guidance sample to have the same levels of initial dissatisfaction as guidance clients, they proved to make equally large gains in job satisfaction over time. This was a somewhat depressing result of the evaluation from the viewpoint of careers guidance practitioners, since increased job satisfaction has been claimed as a result of guidance by previous research. The point is, of course, not that guidance 'fails' in this respect, but that people cannot put up with job dissatisfaction for long and find various kinds of help and self-help - apart from guidance - to escape from it.

There were, in any case, much more positive conclusions from the evaluation. Especially, adult guidance clients took part in more further education activities (independently of their employers), and these activities were more successful

ARTICLES

in terms of completion rates and qualification rates. These positive results came through strongly, despite including previous education and training experiences, and educational aspirations, along with the other matching variables. In a subsequent statistical paper (White & Killeen, 2002), we showed that the educational impacts were also robustly maintained in the face of quite severe sensitivity tests which we imposed.

John and I were keen to develop the findings about educational impacts. In 2001, we worked together on another evaluation, concerning innovative provision of adult literacy and numeracy courses. Here John once more displayed his virtuosity in questionnaire design, but his illness prevented him from taking part in the later stages. John was also developing an imaginative evaluation design for self-funded further education, which would use secondary data sources. On a day when we were due to meet to discuss his initial ideas, I learned that he was being taken into hospital. He continued to think about and discuss the evaluation of further education throughout his illness.