

**The policeman and the statistician**  
**On the quality of the raw data in official statistics**

**Anton Färnström**

**Crime prevention council, Stockholm, Sweden, [anton.farnstrom@bra.se](mailto:anton.farnstrom@bra.se)**

As state institutions are implementing increasingly sophisticated computerized systems, tremendous amounts of data are being constantly generated. This has opened up a wide range of possibilities for the production of statistics that are highly relevant to virtually all sectors of society. However, despite a large consensus regarding the importance of such statistics, little research examining the specific quality issues associated to administrative data is available. Here, we report the results of two studies conducted by the Swedish Crime Prevention Council evaluating the quality of the raw data used for the official crime statistics in Sweden. Our analyses revealed that the Police assigned in average 12% of the crime codes incorrectly, with severe consequences for the final statistics. Furthermore, different interpretations and over-usage of a particular code have rendered national statistics on police decisions nearly unusable. Based on our observations regarding the main causes of these errors, we propose a few recommendations likely to be applicable to both users and producers of statistics based on administrative sources. With official statistics being at the backbone of evaluations and strategic decisions, we strongly believe that investigating and continuously improving the quality of the raw data is a must for all the national statistical agencies.

Key words: Administrative register, coding error, measurement error, crime statistics