Article

HOW IMPORTANT IS "REGRESSION TO THE MEAN" IN AREA-BASED CRIME PREVENTION RESEARCH?

David P. Farrington[1] and Brandon C. Welsh[2]
[1]Institute of Criminology, University of Cambridge, UK
[2]Department of Criminal Justice, University of Massachusetts Lowell, Massachusetts, USA

Correspondence: David P Farrington, Institute of Criminology, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DT, UK
E-mail: dpf1@cam.ac.uk

Abstract

If an area with a relatively high crime rate is chosen to receive a crime prevention programme, its crime rate is likely to decrease after the programme because of fluctuations in crimes over time caused by fluctuations in influencing factors. This is called "regression to the mean". If this high-crime experimental area is compared with a control area with a lower crime rate, crimes are likely to decrease more in the experimental area than in the control area because of these fluctuations, possibly creating the illusion that the programme had an effect when in fact it did not. This article estimates the importance of "regression to the mean" using recorded crime rates in police Basic Command Units (BCUs) in England and Wales in 2002–03 and 2003–04. It is concluded that, in reasonable comparisons between areas with high and moderately high crime rates, this effect may cause a 4% decrease in crimes. Hence, it could not account for much of the 19% decrease in crimes following improved street lighting.

Keywords
crime prevention; evaluation; regression to the mean; street lighting; meta-analysis

Crime Prevention and Community Safety (2006) **8**, 50–60.
doi:10.1057/palgrave.cpcs.8150017

Introduction

In an earlier issue of this journal, Marchant (2005) published a thinly veiled critique of our research on street lighting and crime (Farrington and Welsh, 2002). His three arguments were as follows:

1. Area-based crime prevention programmes should be evaluated using a randomized experimental design.
2. Fluctuations in the numbers of crimes from one year to the next are so large that they mask any effect of an intervention.
3. Where an experimental area (receiving an intervention) is compared with a non-equivalent control area, "regression to the mean" will cause the number of crimes in the experimental area to decrease more than the number of crimes in the control area, and this decrease may mistakenly be interpreted as a consequence of the intervention.

It is unfortunate that we were not invited to respond to Dr. Marchant's critique in the same issue of the journal as it was published. We were able to respond to his earlier critique published in the *British Journal of Criminology* (see Farrington and Welsh, 2004; Marchant, 2004). The main aim of the present article is to respond to his 2005 critique and especially to provide new analyses of the importance of "regression to the mean". While Marchant (2005) focuses mainly on street lighting and crime, his arguments apply very generally to any evaluation of an area-based intervention, including physical, situational and policing interventions (e.g. CCTV, target hardening, neighbourhood watch, hot spots policing, community policing, problem-oriented policing, etc.).

Randomized experiments

There is nothing new in the suggestion that crime reduction programmes should be evaluated using randomized experiments. We have been prominent both in recommending this and in reviewing existing criminological experiments (Farrington, 1983, 2003a, b; Farrington and Painter, 2003; Farrington and Welsh, 2005, 2006).

There are, however, severe practical problems in evaluating an intervention such as improved street lighting using an experimental design, and such an evaluation has never been conducted. Marchant (2005) suggests that the unit of randomization should be the household or dwelling. In order to provide improved street lighting, pavements have to be dug up, cables have to be laid and lamp-posts have to be erected. How would members of the public react if their house was chosen to have street lights outside or to be in darkness on a random basis? It would be a little more practical to allocate streets at random to have better or worse lighting, but again this would lead to practical problems and complaints from the public. In the Dudley evaluation of improved street lighting (Painter and Farrington, 1997), tenants on the control estate were so angry that the lighting had been improved on the experimental estate that they formed a tenants' association and campaigned for the local authority to improve their lighting as well.

Variance of the number of crimes

In our meta-analysis of the effects of improved lighting on crime (Farrington and Welsh, 2002), we had to base our measure of effect size on the numbers of crimes in experimental and control areas before and after the intervention, because this was the only information that was regularly provided in these evaluations. The basic data for each individual study was as follows:

	Before	After
Experimental	a	b
Control	c	d

where *a*, *b*, *c* and d are numbers of crimes.

The odds ratio (OR) was used as a measure of effect size: OR= $(a*d)/(b*c)$. The OR compares the odds of a crime after compared with before in the control area (d/c) with the odds of a crime after compared with before in the experimental area (b/a). The OR is intuitively meaningful because it measures the relative change in crimes in the control area compared with the experimental area. For example, OR = 2 could be obtained if crimes doubled in the control area and stayed constant in the experimental area, or if crimes decreased by half in the experimental area and stayed constant in the control area (etc.).[1]

The statistical significance of an OR is calculated by reference to the variance V of LOR, its natural logarithm. The usual formula is: $V(\mathrm{LOR}) = 1/a + 1/b + 1/c + 1/d$. This formula implies that the variance of the number of crimes in a year is the same as its mean, so that for example $V(a) = a$. Marchant (2004, 2005) argues that the variance is much greater than the mean. In response, Farrington and Welsh (2004) showed that, even if $V(a)$ were five times a, their conclusions about the effect of improved lighting on crime would be unchanged. This is true even if $V(a)$ were 10 times a.

Since Farrington and Welsh (2004) wrote their response to Marchant (2004), there have been two further developments. First, monthly crime data were collected for 70 areas in a national evaluation of CCTV (Farrington *et al*, 2005) and were used to estimate the yearly variance (12 times the monthly variance). This estimate of the yearly variance is too high because, unlike yearly figures, monthly numbers of crimes are influenced by seasonal factors.

For each area in each year, the total number of crimes N was compared with V/N, where V is the estimated variance of the yearly number of crimes.

It was clear that V/N increased with the total number of crimes. The correlation between V/N and N was 0.77 ($P < 0.0001$). A linear regression analysis showed that: $V/N = .0008 * N + 1.2$. Hence, for values of N up to 1,000 (which was true of the vast majority of evaluations of improved lighting), V/N was between 1 and 2. On the basis of this research, the "overdispersion" factor seemed likely to be no greater than 2.

Second, Jones (2005) compared six statistical methods of estimating a weighted mean effect size in a meta-analysis and found that both of the most common methods (the so-called "fixed effects" and "random effects" methods; see Lipsey and Wilson, 2001) gave implausible results. A multiplicative variance adjustment method gave plausible results (in agreement with other methods) and exactly fitted the data. For the lighting meta-analysis, a variance adjustment factor of 4.4 was required to fit the data exactly. This reflects both overdispersion and the heterogeneity of effect sizes. The weighted mean OR was 1.23 (95% confidence interval 1.10–1.39, $z = 3.48$, $P = 0.0005$). Even after allowing for overdispersion, we can conclude that, on average, improved lighting led to a significant decrease in crime of about 19% (from 1/1.23) – similar to our original conclusion.

Regression to the mean

The idea of "regression to the mean" usually depends on the assumption that a measured score (e.g. on an IQ test) reflects partly the true score and partly an error component, which varies randomly over time (Campbell and Kenny, 1999). Hence, someone with a high measured score on one test would tend to have an above-average error component on that test, and on the next test, this component would tend to decrease towards the mean. Just as high scores would tend to decrease from one test to the next, low scores would tend to increase.

There is nothing new in the suggestion that "regression to the mean" might create the illusion of a reduction in crime (see, e.g. Maltz *et al*., 1980). This concept is usually applied to individuals. Applying it to crime data over time, we might hypothesize that there are both persisting influences and varying influences on the measured number of crimes committed in an area. To the extent that influences persist (e.g. demographic features of the population, design of housing estates), high crime areas in one year will still tend to be high crime areas in the next year. To the extent that influences vary in a way that might be regarded as random (e.g. the locking up or release of prolific offenders, changes in the local unemployment rate), high crime areas in one year may "regress to the mean" and decrease in the next year. This is only a problem for area-based crime prevention research if the experimental area has a different (e.g. higher) crime rate compared with the control area before the intervention. If the two areas are truly comparable in their crime rates, both might be expected to regress to the same extent.

Marchant (2005) investigated regression to the mean using data on the number of household burglaries in 124 areas collected by Tilley *et al.* (1999). Professor Tilley kindly provided us with the data to reanalyse. Table 1 summarizes Marchant's key argument. For areas whose burglary rate was above the mean in year 1, the number of burglaries decreased from 432.4 to 361.3 on average in year 2. For areas whose burglary rate was below the mean in year 1, the number of burglaries increased from 218.5 to 224.4 in year 2. Over all areas, the number of burglaries decreased from 299.6 to 276.3 on average. The correlation between the burglary rates in years 1 and 2 was 0.80.

While Marchant (2005) demonstrated that regression to the mean *could* occur (which again is not new), he did not quantify its extent. The OR for Table 1 is large (1.23), but not statistically significant.[2] However, this example is equivalent to comparing an area whose burglary rate is at the 81st percentile with one whose burglary rate is at the 31st percentile.[3] This seems equivalent to setting up a straw man to be knocked down. It is not surprising that regression to the mean would be important in an evaluation carried out with such different areas. In our systematic review (Farrington and Welsh, 2002), we only included evaluations where there was a *comparable* control area. We explicitly excluded comparisons between a high-crime experimental area and the rest of the police division, which perhaps could involve such disparate crime rates.

Future crime prevention evaluators should seek to document the relative crime rates of experimental and control areas. This is estimated in studies based on victim surveys. For example, in the Dudley evaluation of improved street lighting (Painter and Farrington, 1997), the prevalence of victimization was very similar in experimental and control areas (42% compared with 39%) before the intervention, but the incidence of victimization was 40% higher in the experimental area.

Most area-based crime prevention evaluations report only police-recorded numbers of crimes, not crime rates. It is quite difficult to estimate crime rates in many cases, because of the problem of measuring the number of persons (or other units) at risk. While it is possible to measure the number of households at risk of household burglary, it is harder to measure the number of persons in an area who are at risk of becoming a victim of violence (for example), except

Table 1 Changing numbers of burglaries

	Year 1	*Year 2*	*Change*
Above	432.4	361.3	−71.1
Below	218.5	224.4	+5.9
Total	299.6	276.3	−23.3

Note: Above=Above average burglary rate at time 1; Below=Below average burglary rate at time 1; The data show average number of burglaries in an area; Based on 116 areas known in years 1 and 2; Data from Tilley *et al.* (1999).

by labour-intensive observational methods. In some cases, the denominator should be a combination of persons and times at risk.

The data collected by Tilley *et al* (1999) are limited because the areas were self-selected and unrepresentative and the information concerned only household burglary. We (Farrington and Welsh, 2004) investigated regression to the mean by comparing recorded crime rates of all 38 English police force areas in successive years. The crime rate in one year was very highly correlated (0.98–0.99) with the crime rate in the following year. Furthermore, the average change in crime rates for police forces in the highest quartile was very similar to (within 1% of) the average change for police forces in the second quartile.

While this analysis showed little evidence of regression to the mean, it might be objected that police force areas are much larger than the size of areas in crime prevention evaluations. We (Farrington and Welsh, 2004, p 465) stated that "Whether different results would be obtained in an analysis based on smaller areas is an empirical question for the future". We now address that empirical question.

Basic command unit (BCU) data

In recent years, the Home Office has published national recorded crime data for police BCU areas – basically police divisions – on its web site at the same time as it releases its annual report on *Crime in England and Wales* (e.g. Dodd *et al*., 2004). However, because there was a major change in police recording practices in April 2002, it is difficult to compare crimes before that date with crimes after. We are able to analyse recorded crime data for six major types of offences (violence, sex, robbery, burglary, motor vehicle theft and theft from vehicles) in BCUs in England and Wales in 2002–03 and 2003–04. Excluding the smallest BCUs (four with a population below 50,000), there were 250 BCUs with data available on population and crimes in both years. The crime rate (number of crimes per 1,000 population) in 2002–03 was highly correlated with the crime rate in 2003–04: 0.949 for violence, 0.955 for sex, 0.986 for robbery, 0.972 for burglary, 0.964 for motor vehicle theft, 0.970 for theft from vehicles and 0.985 for total crimes (total of these six types).

Table 2 investigates regression to the mean, dividing the 250 BCUs into five quintiles of 50 each, according to their crime rates in 2002–03. To the extent that there is regression to the mean, the average number of crimes in the highest quintile (from the 80th to the 100th percentile) should decrease, while the average number of crimes in the lowest quintile (from the 1st to the 20th percentile) should increase.[4] There was no change in the total number of crimes, which averaged 8,946 per BCU in 2002–03 and 8,952 per BCU in 2003–04. The average population per BCU was about 200,000. Some recorded crimes

Table 2 Percentage changes in crimes in rate quintiles

Quintile	*VIOL*	*SEX*	*ROB*	*BURG*	*MVT*	*THFV*	*TOT*
(1) 0–20	26.5	22.9	2.3	−0.4	−3.9	2.2	7.2
(2) 20–40	26.9	12.1	−5	−4.4	−5	−7	2.1
(3) 40–60	12.3	6.6	−6.6	−5.2	−4.7	−8.1	1
(4) 60–80	11.9	6.3	−1.2	−4.6	−12.7	−6.5	0.3
(5) 80–100	8.3	2	−7.8	−13.6	−11.1	−12.3	−3.8
Total	14.8	7.5	−6	−7.7	−8.8	−8.6	0.1
OR 5 vs 4	1.033	1.042	1.072	1.104	0.982	1.066	1.042
OR 5 vs 3	1.037	1.045	1.013	1.098	1.073	1.048	1.049
OR 5 vs rest	1.09	1.083	1.051	1.109	1.039	1.069	1.059

Note: VIOL=violence, ROB=robbery, BURG=burglary; MVT=motor vehicle theft, THFV=theft from vehicle; TOT=total (of these 6); OR=Odds Ratio; 50 BCUs in each quintile, total=250.

increased (violence and sex) while others decreased; some recorded crimes had only a few hundred cases per BCU on average (sex and robbery), while others had between 1,000 and 3,500 cases per BCU. Hence, these data allow tests of regression to the mean under varying conditions.

The figures in Table 2 show percentage changes between 2002–03 and 2003–04. For example, BCUs in the lowest quintile on the total crime rate in 2002–03 showed a 7.2% increase in total crimes by 2003–04, while BCUs in the highest quintile showed a 3.8% decrease. All the figures indicate that some regression to the mean is occurring.

The importance of regression to the mean was assessed using the OR. For example, crimes decreased by 3.8% for the highest quintile (5) and increased by 0.3% for the next highest quintile (4). Comparing quintile 5 (from the 80th to the 100th percentile in crime rate in 2002–03) with quintile 4 (from the 60th to the 80th percentile), OR = 1.042. The implication is that, if a researcher compared an experimental area at the (median) 90th percentile with a control area at the (median) 70th percentile, this could create the illusion of a 4% decrease in crime. The ORs ranged from 0.982 for vehicle theft to 1.104 for burglary.

If a researcher were so unwise as to compare a BCU in quintile 5 (at the 90th percentile) with a BCU at quintile 3 (at the 50th percentile), the OR for total crimes was only slightly higher, at 1.049. Even more unwisely, if a researcher compared a BCU in quintile 5 (90th percentile) with the rest of the division (median 40th percentile), the OR for total crimes was only slightly higher, at 1.059. The OR was highest for burglary in all cases. Assuming BCUs with average numbers of crimes, none of these ORs would have been statistically significant.

Table 3 Percentage changes in rate deciles

Decile	*VIOL*	*SEX*	*ROB*	*BURG*	*MVT*	*THFV*	*TOT*
(6) 50–60	11	8.4	−14.2	−2.3	−3	−5.1	1
(7) 60–70	9.3	11.9	−3.1	−4.1	−11	−7.3	1.3
(8) 70–80	14.2	0.5	−0.1	−5	−14.1	−5.9	−0.6
(9) 80–90	7	2.7	−5.5	−10.5	−5.1	−10.6	−2.6
(10) 90–100	9.3	1.5	−9	−15.6	−16	−13.6	−4.5
OR 10 vs 9	0.979	1.011	1.039	1.06	1.13	1.035	1.02
OR 10 vs 8	1.045	0.99	1.098	1.125	1.023	1.089	1.041
OR 10 vs 7	1	1.102	1.064	1.136	1.059	1.074	1.061
OR 10 vs 6	1.015	1.068	0.943	1.158*	1.155	1.099	1.057
OR 10 vs rest	1.025	1.051	1.047	1.118	1.094	1.071	1.048

Note: VIOL=violence, ROB=robbery, BURG=burglary; MVT=motor vehicle theft, THFV=theft from vehicle; TOT=total (of these 6); OR=Odds Ratio; 25 BCUs in each decile; $^*P<0.05$, assuming average-sized BCUs.

It might be argued that researchers are likely to compare areas that are more comparable than in the above examples. Hence, Table 3 divides the BCUs into deciles according to their crime rates in 2002–03 and again shows percentage changes in numbers of crimes between 2002–03 and 2003–04. For example, crimes decreased by 4.5% for the highest decile (10) and by 2.6% for the next highest decile (9). Comparing decile 10 (from the 90th to the 100th decile on crime rate in 2002–03) with decile 9 (from the 80th to the 90th percentile), OR = 1.020. Therefore, if a researcher compared an experimental area at the median 95th percentile with a control area at the median 85th percentile, this could create the illusion of a 2% decrease in crime.

Because there are only 25 BCUs in each decile, the ORs fluctuate more in Table 3 than in Table 2. However, for BCUs with average numbers of crimes, only one OR would have been (just) statistically significant. If a researcher had compared a BCU in decile 10 (at the 95th percentile) with a BCU in decile 6 (at the 55th percentile) on burglary, OR = 1.158, confidence interval = 1.00–1.34, $z = 1.98$, $P = 0.048$). Most researchers would attempt to choose a more comparable control area.

One problem with these analyses is that we do not know which areas have received interventions designed to decrease their crime rates. However, it is reasonable to assume that areas with high crime rates (i.e. those in the highest decile or quintile) would be most likely to receive such interventions. Hence, part of what appears to be regression to the mean might be attributable to the effects of interventions. It is conceivable that the greater "regression" effects observed for burglary might have been caused by the large number of anti-burglary programmes implemented as part of the Crime Reduction Programme (see Kodz and Pease, 2003). Therefore, we may be overestimating the effects of regression to the mean.

Conclusions

Regression to the mean is only a problem if researchers compare two areas with very different crime rates. Providing researchers compare two areas that are fairly similar (within 20 percentile points) in their crime rates, these analyses suggest that regression to the mean is unlikely to cause more than a few (e.g. 4) % reduction in all crimes. However, it may be more serious for studies of burglary. Regression to the mean does not threaten our previous conclusion that improved street lighting causes a decrease in crimes, given our weighted mean OR of 1.23 and our confidence interval of 1.10–1.39 ($P = 0.0005$). As mentioned, this corresponds to a decrease in crime of 19% in experimental areas compared with control areas. Further analyses of crime rates in a large representative sample of smaller areas would be desirable.

In his *British Journal of Criminology* article and in his Home Office critiques, Marchant (2004) ignored studies showing no effect of improved lighting on crime and criticized studies showing a crime-reducing effect of improved lighting. In contrast, in our systematic review (Farrington and Welsh, 2002), we treated all evaluations (that met our inclusion criteria of methodological quality) equally. Some of the studies showing desirable effects of lighting are among the highest-quality 10% of situational crime prevention evaluations (Farrington and Welsh, 2004, p 457). The clear implications of Dr Marchant's critiques, if taken seriously, are that "nothing works" and that virtually all previous evaluations of area-based crime prevention programmes should be disregarded as useless. If taken seriously, this could return us to the dark ages of the Martinson (1974) era, with one important difference.

Dr Martinson was trying to discover "what works" by counting how many results were or were not statistically significant. The problem was that his "vote-counting" method was less scientifically rigorous than meta-analysis, which was developed later to obtain more accurate estimates of effect size. However, Dr Martinson was not explicitly trying to prove that "nothing works", and he later (Martinson, 1979, p 252) rejected his original conclusions in light of new evidence. In contrast, Dr Marchant is straining every sinew to try to think of every conceivable statistical objection to studies showing a desirable effect of improved lighting. This of course maximizes the likelihood of a Type II statistical error: concluding that no effect is present when there is in fact an effect. How many evaluations of criminological programmes could withstand such a determined and destructive statistical assault? This is surely not helpful in advancing knowledge about the effectiveness of crime prevention programmes.

Acknowledgements

We are very grateful to Shane Johnson for insightful comments on an earlier version of this paper.

Notes

1 For an alternative method of estimating effect size using time series data, see Johnson *et al.* (2004).
2 *V*(LOR) was estimated using the following equation:

$$V(LOR) = \left[\frac{V(a)}{a^2} + \frac{V(b)}{b^2} + \frac{V(c)}{c^2} + \frac{V(d)}{d^2}\right]$$

where *a*, *b*, *c* and *d* are the numbers of crimes and *V*(*a*) is the variance of *a* (etc.). $V(a)/a$ was estimated from the equation $V(a)/a = 0.0008 * a + 1.2$. In this example, OR = 1.23, confidence interval = 0.93–1.62, $z = 1.45$, ns.
3 Because crime rates are skewed, 62.1% of areas were below the mean. The highest crime rate area is at the 100th percentile. Hence, above-average areas ranged from percentile 62.1 to the 100th percentile, with the median at the 81st percentile. Similarly, the median below-average area was at the 31st percentile because these areas ranged from percentile 0.9 to percentile 62.1.
4 Since there were 250 BCUs, the lowest quintile actually ranged from 0.4% to 20%, while the highest ranged from 80.4% to 100%.

References

Campbell, D.T. and Kenny, D.A. (1999). *A Primer on Regression Artefacts*. New York: Guilford Press.

Dodd, T., Nicholas, S., Povey, D. and Walker, A. (2004). *Crime in England and Wales, 2003/2004*. London: Home Office (Statistical Bulletin 10/04).

Farrington, D.P. (1983). Randomized Experiments on Crime and Justice. In Tonry, M. and Morris, N. (eds.) *Crime and Justice*, Vol (4). Chicago: University of Chicago Press, pp. 257–308.

Farrington, D.P. (2003a). British Randomized Experiments on Crime and Justice. *Annals of the American Academy of Political and Social Science* (589), pp. 150–167.

Farrington, D.P. (2003b). Methodological Quality Standards for Evaluation Research. *Annals of the American Academy of Political and Social Science* (587), pp. 49–68.

Farrington, D.P. and Painter, K.A. (2003). How to Evaluate the Impact of CCTV on Crime. *Crime Prevention and Community Safety* 5 (3), pp. 7–16.

Farrington, D.P. and Welsh, B.C. (2002). *Effects of Improved Street Lighting on Crime: A Systematic Review*. London: Home Office (Research Study No. 251).

Farrington, D.P. and Welsh, B.C. (2004). Measuring the Effects of Improved Street Lighting on Crime: A Reply to Dr. Marchant. *British Journal of Criminology* (44), pp. 448–467.

Farrington, D.P. and Welsh, B.C. (2005). Randomized Experiments in Criminology: What Have we Learned in the Last Two Decades? *Journal of Experimental Criminology* (1), pp. 9–38.

Farrington, D.P. and Welsh, B.C. (2006). A Half-Century of Randomized Experiments on Crime and Justice. In Tonry, M. (ed.) *Crime and Justice*, Vol 34. Chicago: University of Chicago Press (in press).

Farrington, D.P., Gill, M., Waples, S.J. and Argomaniz, J. (2005). The Effects of CCTV on Crime: Meta-Analysis of an English National Quasi-Experimental Multi-Site Evaluation. Unpublished paper.

Johnson, S.D., Bowers, K.J., Jordan, P., Mallender, J., Davidson, N. and Hirschfield, A.F.G. (2004). Evaluating Crime Prevention Scheme Success: Estimating "Outcomes", or How Many Crimes were Prevented. *Evaluation* (10), pp. 327–348.

Jones, H.E. (2005). *Measuring Effect Size in Area-Based Crime Prevention Research*.Unpublished M.Phil. Thesis, Statistical Laboratory, Cambridge University.

Kodz, J. and Pease, K. (2003). *Reducing Burglary Initiative: Early Findings on Burglary Reduction*. London: Home Office (Findings No. 204).

Lipsey, M.W. and Wilson, D.B. (2001). *Pratical Meta-Analysis*. Thousand Oaks, CA: Sage.

Maltz, M.D., Gordon, A.C., McDowall, D. and McCleary, R. (1980). An Artifact in Pretest–Posttest Designs: How it can Mistakenly Make Delinquency Programs Look Effective. *Evaluation Review* (4), pp. 225–240.

Marchant, P. (2004). A Demonstration that the Claim that Brighter Lighting Reduces Crime is Unfounded. *British Journal of Criminology* (44), pp. 441–447.

Marchant, P. (2005). What Works? A Critical Note on the Evaluation of Crime Reduction Initiatives. *Crime Prevention and Community Safety* 7 (2), pp. 7–13.

Martinson, R. (1974). What Works? Questions and Answers About Prison Reform. *The Public Interest* (55), pp. 22–54.

Martinson, R. (1979). New Findings, New Views: A Note of Caution About Sentencing Reform. *Hofstra Law Review* (7), pp. 243–258.

Painter, K.A. and Farrington, D.P. (1997). The Crime Reducing Effect of Improved Street Lighting: The Dudley Project. In Clarke, R.V. (ed.) *Situational Crime Prevention: Successful Case Studies* 2nd ed. Guilderland NY: Harrow and Heston, pp. 209–226.

Tilley, N., Pease, K., Hough, M. and Brown, R. (1999). *Burglary Prevention: Early Lessons from the Crime Reduction Programme*. London: Home Office (Crime Reduction Research Series Paper 1).