

POLYGRAPH PILOT STUDY

Final Report

Don Grubin
Newcastle University
Newcastle upon Tyne

July, 2006

EXECUTIVE SUMMARY

This report describes a two year pilot project run on behalf of the National Probation Service in which polygraph testing was introduced into the sex offender treatment programmes of 10 probation areas in England. Participation by offenders in the project was voluntary.

The pilot was an exploratory trial intended to ascertain whether polygraph testing, when conducted as part of a probation sex offender treatment programme:

- can be an effective way of gathering information about risk behaviours not otherwise disclosed;
- results in changes to risk assessment, treatment and supervision;
- is of general assistance to probation officers in their treatment and supervision of offenders, that is, whether it contributes to the management of an offender in ways apart from the immediate changes associated with a test.

As an exploratory trial, the project set out to clarify factors that would be of relevance to a future, more definitive, randomised controlled trial of polygraph testing in sex offenders, consistent with the Medical Research Council's Framework for the Development and Evaluation of Randomised Control Trials for Complex Interventions. The specific objectives of the project were therefore:

1. to elucidate issues associated with the incorporation of polygraph testing in the supervision and treatment of sex offenders by the probation service;
2. to ascertain the take up rate for polygraph testing when participation is voluntary;
3. to explore whether polygraphy is associated with an increase in clinically useful disclosures made by tested offenders;
4. to explore whether polygraphy is associated with changes to the risk assessment, treatment and supervision of tested offenders;
5. to establish whether polygraphy has other effects of a more general nature on the treatment and supervision of offenders, separate from those associated with specific changes in assessment, treatment and supervision;
6. to make recommendations regarding future implementation.

Numbers tested

As of 15 September 2005, 347 offenders had attended for testing, 33% of whom (116) were tested on two or more occasions. In total, 483 polygraph examinations were carried out.

The estimated uptake rate for testing by volunteer participants was 43%, with a range of 34 to 53% between areas. The retest rate was 47%. Because of a shortage of

administrative resources needed to collect the information, two areas were unable to provide figures for the numbers of men who declined to take part in the pilot.

Probation officer feedback

Case managers and treatment facilitators were asked to complete feedback forms following every test. These forms asked about new disclosures, changes in risk assessment, supervision and treatment, and ‘other effects’ resulting from the polygraph examinations. As shown in Table i.1 below, forms were received from either case managers or treatment facilitators in respect to approximately 70% of first tests, 60% of retests, and just over two thirds of all tests carried out.

Table i.1: Number and percent of Actions Taken forms completed by case managers and treatment facilitators in relation to first tests and retests.

	FIRST TEST	RETEST	TOTAL
Case Manager	183 (53%)	54 (40%)	237 (49%)
Treatment Facilitator	113 (33%)	33 (24%)	146 (30%)
Case Manager <i>or</i> Treatment Facilitator	247 (71%)	80 (59%)	327 (68%)
Not available	100 (29%)	56 (41%)	156 (32%)
TOTAL	347	136	483

For comparison purposes, between June and September 2004 feedback forms were sent to case managers of sex offenders who were waiting for, or who were on, treatment programmes in 4 probation areas where polygraph testing was not introduced – these offenders were therefore being supervised ‘as normal’. The forms were similar to those sent to case managers of polygraphed offenders, and asked about new disclosures that had occurred in the previous months, as well as about changes in risk assessment, supervision and treatment that occurred over this time (Time 1).

Forms were sent in relation to 308 comparison cases, representing nearly all of the relevant sex offenders in the 4 areas (as 28 of these offenders were not on the original lists provided to us, there may have been a small number of other sex offenders in these areas of whom we were not aware). Feedback forms were received in relation to 58% (180 of 308) of these offenders, with most case managers being responsible for more than one relevant offender. Returns from two of the comparison areas were nearly complete, but they were less good from the other two areas. Because of the uneven nature of the feedback, it is unclear the extent to which the findings in relation to the comparison cases are representative of all the relevant sex offenders in those areas, and caution is therefore needed when generalising from them.

Case managers who returned forms from the comparison areas were surveyed again 6 to 12 months later regarding the same offender in order to form a subgroup with which to compare the retested polygraphed offenders (Time 2). Second forms were

received in relation to 36% (64 of the 180) of these offenders. This limited response rate must be taken into account when interpreting this aspect of the findings.

When comparing outcome between polygraphed and non-polygraphed offenders, a number of caveats must be bourn in mind. In particular, the two groups were not randomised, which means that there may have been differences between them that could have influenced the results – variable amounts of missing data makes it difficult to determine how similar or otherwise the two groups were. This is particularly relevant given the voluntary nature of the study, as it is possible that offenders who volunteer for polygraph testing may differ from sex offenders on probation generally.

Test types and test outcome

Tests were of four types:

- *sex history disclosure* (41% of tests): this test explores an offender’s sexual background, in particular his history of sexually deviant behaviour;
- *maintenance* (43% of tests): in this test the offender is asked about behaviours associated with his licence conditions or relapse prevention plan;
- *specific issue-offence* (13% of tests): in this test the offender is asked about aspects of his index offence where his account differs from other information;
- *specific issue-other, or monitoring* (3% of tests): this test focuses on a single issue about which there is concern.

Test outcome is shown in Table i.2. It can be seen that examiners reported about a fifth of all tests as being ‘no deception indicated’ (i.e. the offender ‘passed’ the test), and over 40% as ‘deception indicated’ (that is, a ‘failed’ test). In about a quarter of tests the examiners were unable to reach an opinion (referred to as ‘inconclusive’). In a small number of cases the examination itself was not carried out because sufficient disclosures were made in the pre-test interview, usually in the context of specific issue-offence examinations in which men had been denying significant aspects of their index offence. Approximately 1 in 10 tests did not reach an outcome, either because of non-cooperation by the offender or on occasion equipment malfunction.

Table i.2: Test outcome in first test, retests, and all tests.

	First tests	Retests	All tests
No deception indicated	68 (20%)	35 (26%)	103 (21%)
Deception indicated	161 (46%)	53 (40%)	214 (44%)
Inconclusive	73 (21%)	35 (26%)	108 (23%)
Test not administered due to pre-test disclosure	12 (4%)	2 (1%)	14 (3%)
Examination not completed	33 (10%)	11 (8%)	44 (9%)
Total	347	136	483

New disclosures

Based on their knowledge of a case from probation records and discussions with probation officers, examiners reported that new disclosures relevant to treatment and supervision were made in 79% of first examinations and 78% of retests, regardless of whether the test was 'passed', 'failed', or 'inconclusive'. Nearly 30% of these disclosures took place in the post-test interview that followed the offender having been questioned while attached to the polygraph – in other words, after being challenged with the result of the test.

In the cases for which we had feedback from probation officers, case managers of polygraphed offenders reported new disclosures relevant to supervision being made in 70% of the first tests, compared with 14% of non-polygraphed offenders as reported by their case managers at Time 1. According to the polygraph case managers, disclosures were also made in 67% of the retests, compared with 13% reported by case managers of the non-polygraphed offenders at Time 2. As shown in Table i.3, these differences were highly significant.

It can also be seen in Table i.3 that in respect to both first tests and retests, the odds of a polygraphed offender making a disclosure relevant to his treatment or supervision were 14 times greater than a non-polygraphed offender making such a disclosure. The confidence intervals relating to these odds ratios are wide (in the case of first tests, for example, ranging from 8.5 to over 24), but they are well above 1, suggesting a meaningful clinical impact. However, caution in reaching this conclusion is necessary in view of the voluntary nature of the study, the absence of randomisation, missing data about offenders and the incomplete return of feedback forms.

Treatment facilitators of polygraphed offenders also reported new disclosures in over 70% of the tests for which they completed forms. Information was not available from a similar group of treatment managers of non-polygraph cases with which to compare.

Table i.3: Clinically relevant new disclosures reported by case managers of polygraphed offenders after first tests and retests, and by case managers of non-polygraphed offenders at comparable times. Because of the incomplete return of feedback forms and possible differences between the groups caused by the voluntary nature of the study and the absence of randomisation, these findings are indicative.

	Case Mangers Polygraph	Case Managers Non-polygraph	Significance
New Disclosure	First test	Time 1	$\chi^2 = 114.65$ (df=1) $p < .0001$ odds ratio = 14.4 (C.I. 8.5 – 24.4) ¹
Yes	128 (70%)	25 (14%)	
No	55 (30%)	155 (86%)	
Total	183	180	
New Disclosure	Retest	Time 2	$\chi^2 = 34.47$ (df=1) $p < .0001$ odds ratio = 14.0, (C.I. 5.6 – 35.0) ¹
Yes	36 (67%)	8 (13%)	
No	18 (33%)	56 (87%)	
Total	54	64	

¹confidence interval 95%

Seriousness of disclosures

Probation officers were asked to rate the seriousness of the disclosures made by offenders, using the following definitions:

- low: passive types of behaviour, such as masturbating to deviant fantasies;
- medium: possible preludes to reoffending, such as going to places where there are potential victims;
- high: specific breaches or actual offending.

In cases where probation officers reported that disclosures were made:

- case managers indicated that 25% of disclosures were in the *medium* category (46 of 185) and 19% (36 of 185) were *high*
- treatment facilitators indicated that 27% of disclosures were in the *medium* category (27 of 99), and 9% (9 of 99) were *high*

(It should be noted that in most instances the feedback received from case managers and treatment facilitators related to different offenders).

Disclosures that took place in sexual history or specific issue–offence examinations tended to relate to past actions rather than to current behaviours of concern. Maintenance and other specific issue (monitoring) examinations, however, focused on *current* behaviour, and were thus of particular relevance for the supervising case

manager. Of the 95 maintenance and monitoring examinations for which we feedback from case managers, disclosures rated as being of *medium* severity were made in 26% (25 tests), and of *high* severity in 7% (7 tests), amounting to a third of these examinations.

Risk assessment and intervention

The feedback forms completed by polygraph case managers requested information about revisions to risk assessment and changes in intervention that they attributed to the polygraph examination. For comparison purposes, non-polygraph case managers were asked about changes in risk or intervention that had taken place over the preceding 3 to 6 months. As an exploratory study, the findings reported here should be seen as indicative rather than definitive in view of the voluntary nature of the study, the absence of randomisation, missing data about offenders and the incomplete return of feedback forms.

Both polygraph and non-polygraph case managers reported making revisions in risk assessment with similar frequency – in relation to the former following the polygraph test and in the latter over the preceding months. There was, however, a significant difference in the direction that the risk assessments changed. As shown in Table i.4, non-polygraph case managers more typically *reduced* risk, with 55% of their changes in risk being downwards, while 81% of the risk modifications made by polygraph case managers represented an increase in risk. It should be noted, however, that the number of cases in which risk assessment actually changed was low in both groups.

Table i.4: Changes in risk assessment reported by case managers of polygraphed and non-polygraphed offenders. Information relates to changes made after first tests for polygraphed offenders, and at Time 1 for the non-polygraphed offenders.

	Polygraph case managers	Non-polygraph case managers	Significance
Increased risk	29 (81%)	14 (45%)	$\chi^2 = 7.60$ (df=1) p < .01 odds ratio = 5.0, (C.I. 1.7 – 14.6) ¹
Decreased risk	7 (19%)	17 (55%)	
Total	36	31	

¹confidence interval 95%

Findings regarding changes in risk were similar for the treatment facilitators of polygraphed offenders, with over 90% of their reassessments of risk being upwards. Information was not available from a similar group of treatment managers of non-polygraph cases with which to compare.

It was found that changes in risk assessment, supervision, treatment, or the initiation of a new intervention were reported more frequently by the cases managers of polygraphed offenders compared with the non-polygraph case managers: a change in

one of these variables was noted in respect to 41% of the polygraphed offenders for whom we had the information (73 of 177), compared with 27% (49 of 180) of the non-polygraphed offenders ($p < .01$, odds ratio 1.9, confidence interval 1.2 – 2.9).

Reported impact on risk assessment, supervision and treatment was most prominent following maintenance tests. In addition, case managers indicated that over a third of these tests resulted in information being passed on to the police.

Other effects

Apart from actual changes in risk assessment, treatment and supervision, case managers and treatment facilitators were asked whether the polygraph examination had any other effects on their interaction with the offender. We received feedback in relation to this from case managers following 235 tests, and from treatment facilitators following 142.

‘Other effects’ attributed to the polygraph examination were reported by case managers after 48% of tests (113 of 235), and by treatment facilitators after 40% (57 of 142). These ‘other effects’ related to qualitative aspects of a probation officer’s work with sex offenders, and included such things as confirmation of an existing risk assessment, improved co-operation with supervision, progress in overcoming denial, increased engagement in treatment, or obtaining information that is passed on to child protection agencies or public protection meetings. Illustrations in the form of case examples are given in the body of the report.

Overall impact on assessment, supervision, interventions and treatment

When the ‘other effects’ referred to above are taken into account, case managers reported that the polygraph examination had an impact on their treatment or supervision in 65% of the tests for which we had feedback (149 of 229), while treatment facilitators reported this after 62% of tests (89 of 144).

Utility

After every test, probation officers were asked to rate the impact of the polygraph test on their dealings with the offender. Their responses are shown in Table i.5. On the tests for which we had feedback, it can be seen that case managers and treatment facilitators both rated the polygraph examination as being *Somewhat* or *Very Helpful* in over 90% of cases. Very few tests were considered by probation officers to have had either no or negative impact.

Table i.5: Global rating of probation officers of the usefulness of polygraph tests in their interactions with offenders.

	Case managers	Treatment facilitators
Very detrimental	0	0
Somewhat detrimental	4 (2%)	0
No effect	10 (4%)	13 (9%)
Somewhat helpful	81 (35%)	61 (42%)
Very helpful	139 (59%)	71 (49%)
Total	234	145

Implementation issues

Based on feedback from examiners and probation officers, and on differences that were observed between areas, a number of issues affecting implementation were identified. The most important of these related to the commitment of programme and treatment managers to testing, and the need for case managers to have a good appreciation of their role in the process. The following are considered to be particularly relevant to effective implementation:

- the active involvement of the programme and treatment managers of the sex offender teams, and the support of their higher managers
- keeping case managers up to date with the nature of polygraph provision and their role in it
- establishing systems to support case managers in their understanding and use of polygraph findings
- developing protocols to integrate polygraph testing with treatment
- good administrative arrangements
- examiner competency in both interviewing skills and polygraph technique
- full time examiners who interact assertively with case managers and members of the treatment team, who are closely monitored and supervised, and who remain up to date in their knowledge and technique
- protocols to deal with offenders who default from examinations, or otherwise do not cooperate with testing

Conclusion

This pilot has demonstrated that the use of polygraph testing in probation sex offender treatment programmes is feasible, practical, and can be implemented effectively. Evidence arising from the pilot indicates that polygraphy has the potential to make an important contribution to the treatment and supervision of sex offenders on probation,

contributing to public safety by enabling probation officers to better monitor risk and to bring about more effective and timely interventions.

As an exploratory study, the findings reported here should be seen as indicative rather than definitive. Further research is needed in the context of a randomised trial to better assess the relationship between polygraphy and offender disclosures, and the impact or otherwise polygraphy may have on offender treatment and supervision. Apart from any policy considerations, this would best be carried as part of a mandatory testing programme in order to remove the self selection biases inherent in voluntary testing.

TABLE OF CONTENTS

	<u>Page</u>
Executive Summary	<i>i - ix</i>
1. Introduction	1
Objectives of the study	1
2. Background	2
The purposes of an exploratory study	4
3. Methods	5
Participants	5
Non-polygraph offenders	6
Research design	6
<i>Background information</i>	7
<i>Outcome measures</i>	7
Take-up rates	10
<i>Demographic data relating to take-up</i>	10
<i>Demographic summary</i>	14
Retests	14
4. Polygraph Procedures	16
The polygraph examination	16
Types of test	17
Retests	18
Interaction between examiners and probation officers and the use of information	18
5. Implementation	20
Examiners and examiner training	20
Inconclusive tests	21
Inter-rater reliability	22
Test administration	23
Integration with programmes	24
Defaults	28
Capacity analysis	28
Implementation conclusions	29

	<u>Page</u>
6. Findings	31
Examination outcome	31
<i>Retests</i>	32
Follow-up data	33
<i>Actions taken forms</i>	33
Disclosures	34
<i>Based on examiners' reports</i>	34
<i>Based on probation officer feedback</i>	37
<i>Agreement between examiners and probation officers</i>	39
<i>Seriousness of disclosures as reported by probation officers</i>	39
Risk assessment and intervention	41
<i>Agreement between case managers and treatment facilitators regarding risk</i>	44
<i>Police and MAPPAs</i>	44
<i>Retests</i>	45
Other effects	47
Utility	50
Overall impact of polygraphy on supervision and treatment	51
7. Discussion	53
Impact of polygraph testing	53
Seriousness of disclosures and changes in risk assessment	54
Is the polygraph the cause of the disclosures?	55
The use of polygraph results	56
Utility	57
Conclusion	57
8. Recommendations	58
9. References	60

1. INTRODUCTION

This report describes a two year pilot project run on behalf of the National Probation Service in which polygraph testing was introduced into the sex offender treatment programmes of 10 probation areas in England. Participation by offenders in the project was voluntary. The objective of the pilot was to assess the practicality of using polygraphy within probation sex offender programmes, and to determine the extent to which the addition of polygraph testing assists in the treatment and supervision of sex offenders managed by the probation service.

The pilot commenced in May 2003 with the training of examiners and the establishment of protocols in each of the 10 areas. The first polygraph examinations were carried out in September 2003, with testing established in all 10 sites by July 2004. The aim was to conduct polygraph examinations on 300 individuals in the two year period from September 2003 to September 2005, extending into a third year if necessary. In the event, this target was reached in June 2005, but it was nonetheless decided to continue the pilot for a further year, until the end of June 2006. As described in this report, the extension of the project means that potential solutions to a number of issues encountered during the pilot can be explored.

This report includes findings in respect of the 347 offenders who attended for polygraph testing up to 15 September 2005, 116 of whom were tested on two or more occasions. In addition, outcomes for these polygraphed offenders have been compared with sex offenders from four probation areas where polygraph testing was not introduced.

Objectives of the study

The aims of this research are:

1. to elucidate issues associated with the incorporation of polygraph testing in the supervision and treatment of sex offenders by the probation service;
2. to ascertain the take up rate for polygraph testing when participation is voluntary;
3. to explore whether polygraphy results in an increase in clinically useful disclosures made by tested offenders;
4. to explore whether polygraphy results in changes to the risk assessment, treatment and supervision of tested offenders;
5. to establish whether polygraphy is of general assistance to probation officers involved in the treatment and supervision of offenders, whether it contributes to their management of the case.;
6. to make recommendations regarding future implementation.

2. BACKGROUND

The polygraph, which is often referred to somewhat misleadingly as a lie detector, was developed in the early part of the twentieth century, primarily in the United States. It measures changes in breathing, cardiovascular activity and sweating (via skin conductance) which, in the context of a polygraph examination, are associated with the arousal engendered by being deceptive to emotionally salient questions. In the late 1930s and 1940s it became increasingly used by American law enforcement and intelligence agencies as a means of obtaining and verifying information, as well as in employment settings both to vet prospective employees and to screen existing ones for security breaches. Private examiners also flourished, finding particularly fertile ground in selling their services to carry out so-called fidelity tests of suspected spouses. This growth in polygraphy, however, has not been without controversy and claims of abuse: there have been a number of enquires into the reliability and validity of polygraphy, and federal legislation has been passed in the United States limiting its application by private employers. Nevertheless, the use of polygraphy remains extensive in North America and has spread to a number of other countries (the historical context is reviewed in Grubin & Madsen, 2005, Iacono & Patrick, 2005, and National Research Council, 2002).

Polygraphy was first applied to the supervision of offenders on probation and parole in the 1960s and 1970s by a small number of innovative judges in the United States. Follow-up suggested marked reductions in reoffending rates, albeit in the absence of well designed trials. It was not until the 1990s, however, that such post-conviction applications became prevalent, particularly in relation to sex offender treatment and supervision, to the extent that by the late 1990s it was estimated that probation and parole agencies in up to 35 states were using polygraphy in this capacity (English et al, 2000a; Holden, 2000). In many states polygraphy is now included routinely as a condition of a sex offender's probation or parole, or as a required component of a sex offender treatment programme; in one state, Colorado, polygraph testing of sex offenders is mandatory, comprising a fundamental element in what is referred to as a "Containment Approach" involving case managers, treatment providers and polygraph examiners (English et al., 2000b).

Critics of polygraph testing and reviews of it, although addressing its application in law enforcement and security rather than in post-conviction settings, have raised questions regarding its accuracy, regulation, and the way in which examination results are interpreted (British Psychological Society 1986; Cross & Saxe, 2001; Fiedler, Schmid & Stahl, 2002). Regarding the first of these concerns, a definitive review carried out by the National Academies of Science in the United States concluded that polygraph accuracy is in the region of 81-91% (National Research Council, 2002). The expert panel argued that this level of accuracy was not sufficiently robust for it to be used in security applications where the base rate of deception is likely to be low (as pointed out in the review, very few spies would be expected to be found in government agencies), but it also stated that polygraphy becomes more viable in settings where the expected base rate of deception is higher, putting the threshold at about 10% - a level that is almost certainly met in post-conviction sex offender testing. This distinction is sometimes not made in considerations of the implications of the National Academies report.

The emphasis in post-conviction testing is on the disclosure of information rather than on ‘passing’ or ‘failing’ the polygraph *per se* (in general terms, a ‘passed test’ is one in which the examiner concludes that there was ‘no deception indicated’ to the questions asked, while in a ‘failed’ test the conclusion is one of ‘deception indicated’), although a passed polygraph examination can provide useful confirmation and reassurance regarding an offender’s behaviour, while a failed one in the absence of disclosure may be an indicator of problematic behaviour. It is claimed that when used in the treatment and supervision of sex offenders, polygraph testing results in more reliable sexual histories (especially in respect to the onset and extent of deviant sexual behaviour), fuller accounts of offence behaviour with decreased denial, increased reporting of high risk behaviours, and reductions in reoffending, either because of early intervention or through offenders modifying their behaviour (see Grubin et al, 2004), although again the number of well designed studies in this area is limited. The ability of post-conviction polygraph testing to elicit disclosures from offenders has given rise to interesting legal arguments in the United States regarding the balance between self incrimination and public protection (for example, United States v Antelope, 2005).

Enthusiasm expressed by American proponents of polygraph testing of sex offenders, together with published reports regarding its utility, led to a small Home Office sponsored trial in England whereby sex offenders on treatment programmes in 3 probation areas were asked to undertake polygraph examinations (Grubin et al, 2004). Thirty-two offenders agreed to take part in the study (out of 115 who were approached), all of whom were thought to be well engaged in treatment and about whom probation officers had few concerns. In the course of their polygraph examinations, however, 31 of the men disclosed information of direct relevance to their therapy or supervision. Over 80% admitted to ongoing offence relevant fantasies in spite of denying this in treatment, but more worrying were disclosures made by a third of the offenders regarding high risk behaviours such as having unsupervised contact with children or going to places where they might encounter potential victims; one man disclosed prowling public toilets with a knife in search of young boys, another to over 50 incidents of frottage involving young girls on public transport, and a third to regular unsupervised “dates” with his previous 13 year old victim. Twenty-one of these offenders agreed to be retested 3 months later. On this second occasion, it was found that disclosures of high risk behaviour decreased in both frequency and severity, offenders were more likely to have already reported such behaviour to their probation officers, and more offenders passed the polygraph examination itself (Madsen et al, 2004).

Given these findings regarding the potential utility of polygraph testing in sex offenders, the National Probation Service commissioned the current pilot project. It is an exploratory study intended to determine the practicality of using polygraphy in probation settings, and to provide a basis for assessing whether, when properly administered and supervised, polygraphy can contribute significantly to the treatment and supervision of sex offenders to a degree that would make routine implementation worthwhile.

The purposes of an exploratory study

Randomised controlled trials (RCTs) are considered by many to be the ‘gold standard’ approach to the evaluation of new interventions. There are many issues that need to be resolved, however, before an RCT can be effectively designed and implemented. In recognition of this, the Medical Research Council published a framework for the development and evaluation of randomised control trials for complex interventions (Medical Research Council, 2000); a complex intervention is defined as an intervention that involves a number of separate components, acting both independently and “inter-dependently” with each other (as is the case in polygraphy). This framework identified five phases of investigation in the evaluation of complex interventions:

- theoretical
- modelling
- exploratory trial
- definitive RCT
- long term implementation

As indicated above, the use of polygraphy in sex offender treatment and supervision has in effect passed through theoretical and modelling phases, but an exploratory study to lay the foundation for a more definitive evaluation has yet to be carried out. According to the MRC framework, the aim of an exploratory study is to test out the intervention in practice, amongst other things gaining evidence to support the theoretically expected effect, identify an appropriate control group, establish outcome measures, reach estimates of the numbers of participants needed to demonstrate an effect, and ascertain problems in delivery (including factors such as differences in practitioner skill, training needs, and the impact of ‘learning curves’ on delivery). Whereas modifications of procedure and variations in practice are unacceptable in the context of an RCT, they are expected within an exploratory study in order to address the issues referred to above.

The evaluation of the polygraph pilot project described in this report, therefore, has been carried out in the context of an exploratory study as conceptualised by the MRC. Its findings are intended to be indicative of its impact on sex offender treatment and supervision, and to provide the information required to move on to the next phase of evaluation: a definitive randomised controlled trial.

3. Methods

A separate evaluation of the pilot was originally planned by the Home Office, with outcome for polygraphed offenders compared with offenders who refused polygraph examination, and with offenders in comparison probation areas where polygraph testing was not introduced. This study, however, was never commissioned, which has meant that the evaluation described here is limited by a number of constraints, referred to below.

Participants

All sex offenders taking part or waiting to take part in treatment programmes in the 10 English probation areas listed in Table 1 were eligible for inclusion in the pilot. A small number of offenders who were engaged in one-to-one treatment were also accepted onto the trial if requested by the probation officer involved, in order to assist in overcoming denial, gain a fuller sexual history, or as an adjunct to supervision. In addition, tests were occasionally carried out in situations where, although the offender was not involved in treatment, it was believed that polygraphy could make a significant contribution to the management of the case, for example, when it was requested by a Multi-agency Public Protection Panel (MAPPP).

The 10 probation areas involved in the pilot and the number of tests carried out are shown in Table 1. From the start of testing in September 2003 until 15 September 2005 (the cut-off date for this report), 347 offenders attended for polygraph examinations, of whom 116 men were tested twice, 17 three times, and 3 on four occasions. In total, 483 polygraph examinations were carried out over this time.

Table 1: The 10 pilot areas, and the number of tests carried out in each.

	<u>First tests</u>	<u>Re-tests</u>	<u>Total (% of total)</u>
Thames Valley	39	14	53 (11%)
Northumbria	82	36	118 (24%)
West Midlands	60	25	85 (18%)
Eastern Region ¹	36	18	54 (11%)
Northamptonshire	20	8	28 (6%)
Leicestershire	31	20	51 (11%)
Lancashire	21	1	22 (5%)
London (N & W)	23	9	32 (7%)
Manchester	22	2	24 (5%)
Devon & Cornwall	13	3	16 (3%)
TOTAL	347	136	483

¹ Comprising Hertfordshire, Cambridgeshire and Bedfordshire

All offenders who were tested were men, apart from one woman who was on a Community Rehabilitation Order for indecent assault, and taking part in one-to-one treatment.

Non-polygraph offenders

For comparison purposes, between June and September 2004 case managers in 4 probation areas where polygraphy had not been introduced were approached for information regarding sex offenders whom they were supervising. All were waiting for, or were on, sex offender treatment programmes (SOTP), and were being supervised 'as normal'.

The comparison areas were chosen so that:

- each of the three sex offender treatment programmes being used by the pilot sets was represented;
- there was a mix of offenders from rural, semi-urban and urban areas;
- information about a reasonable number of sex offenders would be obtained;

While they are representative of the areas in which testing was carried out, they are not a direct match for them.

Because only two of the four divisions that make up the London probation area were involved in the pilot (North and East), the rest of the London area (South and West) was chosen as one of the comparison sites. The other areas used for comparison were Hampshire, Humberside and West Yorkshire. Outcome data was sought for 308 cases, representing nearly all of the relevant sex offenders in the 4 areas (28 of these offenders, mostly from London, were not on the original lists provided to us, and there may therefore have been a small number of other sex offenders about whom we were not aware). Forms were returned in relation to 58% (180 of 308) of these cases.

Although used for the purposes of comparison, the non-polygraph offenders were not matched with those who were polygraphed, and one must be cautious in reaching conclusions about differences in outcome for the two groups. In particular, while the polygraph group is comprised of individuals who agreed to be polygraph tested, the non-polygraph sample contains offenders who may have refused testing had they been asked (by definition there are no such individuals amongst the polygraph volunteers). In addition, information regarding stage of treatment, including the number of men waiting to start treatment, was not readily available for either group, and there may be differences between the groups in this respect that could have had an effect outcome.

Research design

Participation by offenders in the pilot was voluntary. Polygraphy was explained to offenders, usually by case managers but in some instances by treatment facilitators, following which they were provided with an information sheet about the procedure.

If the offender agreed to be tested his signed consent was obtained. Offenders could withdraw from testing at any time, even during the course of the polygraph examination itself. Probation officers were advised not to draw any conclusions from an offender's refusal to agree to testing, or his subsequent withdrawal from it.

Background information

Background data was collected on all tested offenders from probation files by the polygraph examiners, augmented by information provided by the offenders during their tests. Background information for offenders who declined to be tested, and for offenders in the comparison areas, was intended to be obtained as part of the Home Office evaluation study. Because this did not take place, a research assistant was employed for a three month period in the summer of 2004 to gather data from probation files. This means that background information for 'refusers' and for the non-polygraph group is available only for those who were on waiting lists or in treatment at that time. As files were often incomplete and sometimes missing, there is in addition a relatively large amount of missing data for these offenders. There was nothing to indicate, however, that missing data or files were related to offender characteristics as opposed to simple oversights in data entry, or records being misfiled or in use elsewhere. This missing data has therefore been treated as 'missing at random'.

Outcome measures

Polygraphed offenders

'Actions Taken' forms were sent to case managers (probation officers responsible for the supervision of the offender) and treatment facilitators (probation officers who deliver treatment on the SOTP) at the same time as the polygraph examination reports, or shortly afterwards. These forms requested information regarding new disclosures made by the offender, the impact of the polygraph examination on risk assessment, treatment and supervision, and its affect on the probation officer's dealings with the offender generally.

The aim was to obtain forms from both case manager and treatment facilitator for each offender. This was not possible where offenders had yet to start treatment or were receiving one-to-one treatment from case managers, in which case no treatment facilitator was involved. It was also not possible in areas where treatment facilitators rather than case managers were more concerned with offender supervision, or when the case was being managed by duty officers, which meant that there was only limited case manager involvement. Thus, Actions Taken forms might not be available because none was returned, because there was no treatment facilitator involved with the offender, or because there was no active case manager.

Attempts were made to follow-up probation officers who did not return Actions Taken forms. This meant that forms were sometimes completed many months after the test had been carried out.

Actions Taken forms were obtained for over 70% of first tests, and for over two thirds of tests carried out in total (Table 2), a reasonable rate of return with which to evaluate outcome. As described in more detail in Chapter VI, there appeared to be minimal differences between cases for which Actions Taken forms were completed compared with cases for which they were not, although due to differences between the two groups which are discussed in Chapter 7 these findings are indicative.

Table 2: Completed Actions Taken forms (percentages relate to column totals).

	FIRST TEST	RETEST	TOTAL
Case Manager	183 (53%)	54 (40%)	237 (49%)
Treatment Facilitator	113 (33%)	33 (24%)	146 (30%)
Case Manager <i>or</i> Treatment Facilitator	247 (71%)	80 (59%)	327 (68%)
Not available	100 (29%)	56 (41%)	156 (32%)
TOTAL	347	136	483

Non-polygraph offenders

Actions Taken forms similar to those for polygraphed offenders were sent to case managers in the comparison areas. They asked about new disclosures made by offenders during the preceding months in the course of their normal supervision meetings or treatment, and about any changes in risk assessment, management or treatment that had taken place in this period. Returns were received in relation to 100% (43 of 43) of the offenders from Humberside, 80% (39 of 49) from Hampshire, 51% (69 of 135) from London, and 36% (29 of 81) from West Yorkshire. Overall, forms were received in relation to 58% of possible cases (180 of 308).

Case managers who had returned Actions Taken forms for non-polygraph offenders were surveyed again 6 to 12 months later regarding the same offender, with information again sought about new disclosures or changes in risk and treatment that had taken place in the preceding months. The purpose of this second survey was to form a subgroup with which to compare retested polygraphed offenders. Second forms were completed in relation to 36% of the relevant offenders (64 of 180), mostly from Hampshire (54%, 21 of 39 offenders) and Hull (77%, 33 of 43); fewer second forms were received from West Yorkshire (31%, 9 of 29) and London West and South (just 1 of the 69).

A small number of Actions Taken forms were received in relation to offenders who had not been on the initial lists of offenders provided to us, but who were nonetheless in treatment or waiting for treatment at the time. Background information regarding these individuals is therefore particularly limited.

Outcome data was also sought in relation to offenders in the pilot areas who had declined to be tested. However, there were particular difficulties in putting arrangements in place to achieve this: examiners did not know, and had no contact with, offenders who had refused testing, which meant that administrative support from within the areas was necessary to identify refusers and their case managers. Because most areas had little in the way of administrative slack, this support was not available. In the event, Actions Taken forms were completed by case managers for only 45 refusers, and by treatment facilitators for an additional 24 (as well as for 18 offenders who had also been reported on by case managers), representing less than 20% of the 'refusers'. Almost all of these forms came from just three areas. Because of the small number, this data has not been included in the analysis.

Other follow-up

Interviews with case managers, treatment facilitators and indeed with offenders themselves to determine outcome subsequent to an offender's polygraph examination would almost certainly have occurred had the Home Office research evaluation been commissioned as planned. In the context of the pilot, however, it was not possible to carry out such interviews. Some follow-up about individual offenders beyond that contained on the Actions Taken forms was obtained from treatment facilitators, programme managers and treatment managers during project review meetings, in the course of discussions that took place between examiners and probation officers about offenders, and from what offenders told examiners during their retests. This information is used in a number of the case studies described in Chapter 6; although useful, this does not provide a systematic follow-up of medium and long term outcome.

Examiners and Programme Managers were asked to provide information about what they perceived to be interesting cases, that is, cases in which the polygraph had a marked impact on supervision or treatment, either negative or positive in nature. No further guidelines were given, the aim being to collect a wide variation of cases. In practice many of the cases were first identified by the examiners, who then reviewed details with probation officers. Some of these cases are used to illustrate specific aspects of polygraph outcome in Chapter 6.

Feedback from examiners regarding their interactions with probation officers and with offenders, and their impressions of the way in which implementation of polygraph testing was working in the areas where they tested, was obtained routinely at the quarterly supervision meetings held throughout the pilot, and at other times when specific issues arose. Information was sought from them regarding their views on a range of matters, but their unprompted observations were also noted.

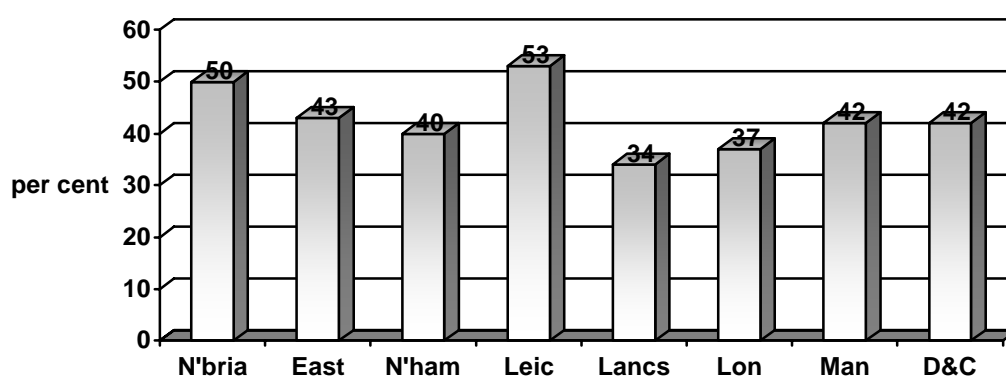
During the second year of the pilot, review meetings were held with the Programme and/or Treatment Managers from each area. During these meetings additional qualitative information regarding the pilot was obtained.

Take-up rates

Areas were requested to log the names of offenders who had been approached to take part in the pilot, but not all had the administrative capacity to liaise with case managers in this way. Good information on refusal rates was available from Devon & Cornwall, Eastern, Lancashire, London and Northamptonshire, while Leicestershire, Manchester, and Northumbria provided estimates based on information sought retrospectively from case managers regarding the offenders approached by them. Thames Valley and West Midlands could not supply any data on refusal rates at all. Thus, while the number of men who were tested is known, the figures for those who refused testing are less exact and should be treated as a good estimate.

Based on the information available from the 8 contributing areas (i.e., excluding offenders from Thames Valley and West Midlands), 248 offenders agreed to attend for polygraph testing and 311 declined. The overall take up rate, therefore, was 44%. It ranged from 34% to 53% across the 8 areas providing data (Figure 2).

Figure 1: Percentage of offenders attending for polygraph examinations in the 8 areas for which data is available.



Six men who kept their appointments withdrew their consent in the course of the pre-examination interview, before being attached to the polygraph, although one later changed his mind and was eventually tested. Because these individuals had commenced the examination, they are included in the polygraph group even though they were not tested.

Demographic data relating to take-up

There were no significant differences between polygraphed offenders, refusers, and non-polygraph offenders from the comparison sites in terms of age, legal status or offence type, but differences did emerge in relation to ethnicity, previous convictions for sex offences, and risk, as discussed below and shown in Tables 3 and 4.

Ethnicity: The difference between participants and refusers was not significant in terms of their ethnic backgrounds, but there was a difference between the non-polygraph group and both participants and refusers. There were significantly more

offenders from minority ethnic groups in the non-polygraph sample, mostly accounted for by offenders classed as African, Caribbean, or 'other black' in background, who made up 11% of non-polygraph offenders. This difference was accounted for entirely by offenders in the London comparison sites (East and South), where 34% were from minority ethnic groups.

Definitive conclusions can not be reached regarding this ethnicity data without further research, but the most likely explanation is that it is an artefactual finding caused by the lack of direct matching between non-polygraph and polygraphed offenders, and the low number of participants from London being outweighed by the larger number of London offenders in the non-polygraph group. This is supported by the finding that there did not appear to be any differences in the ethnic backgrounds of participants and refusers in London, albeit the numbers are small. However, the possibility that polygraph testing may be less acceptable to offenders from ethnic minority groups, and if so why, is something that warrants further exploration.

History of sex offence convictions: Significantly more participants had convictions for previous sex offences. The reason for this apparent willingness of men with a sex offending history to volunteer for polygraph testing is unclear. It is possible that such individuals may be more accepting of the need for treatment and see polygraphy as a possible benefit to them. Alternatively, it may be that they are more keen to prove they are 'low risk' and hope to demonstrate this through the polygraph, either legitimately or perhaps by 'beating' it. If the latter was the case, however, one might expect there to have been a higher proportion of failed or inconclusive polygraph tests amongst those with a sex offending history, which as described later in this report was not in fact found.

Risk: There was a significant difference between participants and refusers in terms of their Risk Matrix 2000 scores, with more offenders rated as 'very high risk' amongst participants and fewer as 'low risk'. There was also a significant difference between participants and non-polygraph offenders, but in this case there was a higher proportion of offenders classified as 'medium' in the latter, with 'very high risk' cases again more common in polygraphed offenders.

One must be cautious in reaching conclusions regarding these differences in risk, however. Because Risk Matrix provides an estimate of the likelihood of reconviction regardless of the severity of the reconviction, the greater proportion of very high risk men does not equate with this group being more 'dangerousness'. For example, 36% (10 of the 28) 'very high risk' cases amongst the participants were offenders whose index offence was indecent exposure; excluding indecent expositors, the proportion of 'very high risk' cases falls to 5% in the participant group, and the difference between participants and the other two groups is no longer significant (it is unclear why there were not more 'very high risk' indecent expositors in the other groups).

Deviance scores, which are based on psychometric testing and are relevant only for men who offended against children, did not differ significantly between the groups. This suggests that the groups were similar to each other in terms of overall risk. However, deviance scores were not readily available in the files and were difficult for the examiners to obtain; the limited number of offenders for whom this data is available makes conclusions regarding deviance scores tentative.

Table 3: Comparison between offenders who agreed to take part in polygraph testing and those who declined to be tested (refusers) from the same probation areas. Chi square used to test for significant differences except for age, where a t test was used.

	PARTICIPANTS (n=347)	REFUSERS (n=213)	Significance
Age (mean)	43 (sd 13.7) range: 19-90	43 (sd 12.1) range: 19-71	NS
Legal status Comm. Rehab. Order Parole (missing)	210 (61%) 137 (39%) (0)	102 (55%) 83 (45%) (28)	NS
Ethnicity white black minority ethnic (missing)	331 (95%) 16 (5%) (0)	144 (91%) 14 (9%) (55)	NS
Offence Against children Internet Other (missing)	166 (48%) 111 (32%) 70 (20%) (0)	94 (53%) 60 (32%) 59 (15%) (26)	NS
Past sex offence Yes No (missing)	130 (30%) 213 (70%) (4)	24 (18%) 111 (82%) (78)	$\chi^2 = 17.05$ (df=1) p < .0001
Risk Matrix Low Medium High Very High (missing)	127 (37%) 117 (34%) 71 (21%) 28 (8%) (4)	77 (43%) 64 (35%) 38 (21%) 2 (1%) (32)	$\chi^2 = 11.29$ (df = 3) p = .01
Deviance Low High	(n=130) 87 (67%) 43 (33%)	(n=104) 76 (73%) 28 (27%)	NS

Table 4: Comparison between offenders who agreed to take part in polygraph testing and non-polygraph offenders from the comparison areas. Chi square used to test for significant differences except for age, where a t test was used.

	PARTICIPANTS (n=347)	NON-POLYGRAH (n=308)	Significance
Age (mean)	43 (sd 13.7) range: 19-90	42 (sd 13.3) range: 18-81	NS
Legal status Comm. Rehab. Order Parole (missing)	210 (61%) 137 (39%) (0)	150 (56%) 119 (44%) (39)	NS
Ethnicity white black minority ethnic (missing)	331 (95%) 16 (5%) (0)	186 (83%) 38 (17%) (84)	$\chi^2 = 22.84$ (df = 1) p <. 0001
Offence Against children Internet Other (missing)	166 (48%) 111 (32%) 70 (20%) (0)	155 (57%) 73 (27%) 45 (16%) (35)	NS
Past sex offence Yes No (missing)	130 (30%) 213 (70%) (4)	58 (22%) 206 (78%) (44)	$\chi^2 = 16.97$ (df = 1) p <.0001
Risk Matrix Low Medium High Very High (missing)	127 (37%) 117 (34%) 71 (21%) 28 (8%) (4)	77 (35%) 95 (43%) 43 (19%) 7 (3%) (86)	$\chi^2 = 8.36$ (df=3) p <.05
Deviance	(n=130)	(n=81)	NS
Low High	87 (67%) 43 (33%)	62 (77%) 19 (23%)	

Demographic summary

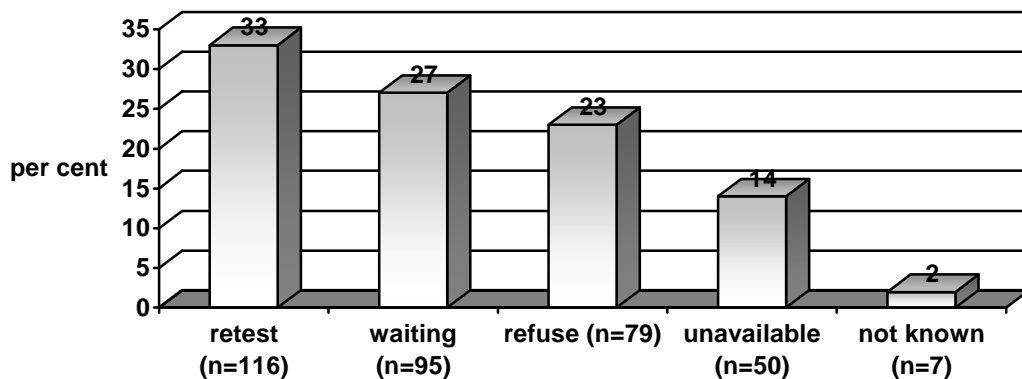
Based on the background data described above, the polygraphed offenders differed from the other two sex offender groups in that sex offence histories were found more commonly in the former. The polygraph group also contained more offenders rated as 'very high risk' on Risk Matrix 2000; this was caused by a greater number of offenders who were convicted of indecent exposure, the relevance of which is unclear. However, the groups could only be compared on a limited number of factors, and the findings may also have been influenced by missing data. Account will need to be taken of these potential biases, as well as ethnicity, in any randomised controlled trial.

Retests

As of 15 September 2005, 33% of the tested offenders (n=116) attended for retests, of whom 17 were tested on three occasions and 3 on four occasions. A further 27% (n=95) agreed to be retested, and were waiting for their retests. Of the remaining 136 offenders (Figure 2):

- 23% (n=79) refused to attend for a retest. Systematic data on the reasons for their refusal are not available as men were not required to provide any explanation given the voluntary nature of the pilot, but anecdotal feedback indicated a range of reasons being provided, including advice from solicitors and the 'stress' associated with the polygraph examination.
- 14% (n=50) were not available for retesting for a variety of reasons, such as having finished the programme, being removed from it, breach, recall to prison, medical problems, or moving from the area.
- in 2% (n=7) the situation regarding retesting was not known as examiners were unable to establish the position regarding them from case managers.

Figure 2: Retest uptake by the 347 offenders who attended for a first polygraph examination.



Excluding those awaiting a retest (some of whom could of course refuse to be tested on the day) and those for whom the situation is unclear, the retest rate amongst the remaining 245 offenders was 47%.

Whether or not an offender attended for a retest was not dependent on test outcome (taking into account NDI, DI and inconclusive results only), having made a sex offending relevant disclosure during the first examination, having a sex offending history, Risk Matrix 2000 score, deviance score, or CRO/parole status (Table 5).

It is not clear from our data what factors influence whether or not an offender will agree to be retested. This would require further research, with offenders interviewed regarding their reasons for not wanting to continue with testing.

Table 5: Factors associated with retest uptake, comparing offenders who were retested with those who refused (n = 116 and 79); none of the differences are statistically significant (chi square).

	Retested (n = 116)	Refused (n = 79)
Sex offending history	28 (24%)	23 (30%)
Risk Matrix 2000 score		
low	46 (40%)	33 (42%)
medium	40 (35%)	22 (28%)
high	21 (18%)	16 (21%)
very high	8 (7%)	6 (8%)
High Deviance score (n = 59 retest, 30 refusal)	20 (36%)	6 (20%)
Parole	39 (34%)	28 (35%)
Outcome of first test		
NDI	21 (20%)	9 (14%)
DI	52 (49%)	44 (68%)
inconclusive	33 (31%)	12 (18%)
Sex disclosure on first test	86 (80%)	51 (70%)

4. POLYGRAPH PROCEDURES

The polygraph examination

All polygraph examinations consist of a pre-test interview, the examination itself, and a post-test interview. Depending on the examinee and the nature of the test, the whole procedure typically lasts from two to four hours.

Pre-test interview: the pre-test interview usually takes from one and a half to three hours. During it the procedure is explained to the offender and the polygraph is demonstrated with a so-called acquaintance test in which he is asked to lie about the identity of a number he has chosen. In the pre-test interview the examiner reviews the offender's background with him, identifies areas of concern, and agrees the questions that will be asked in the actual polygraph test. Many offenders make relevant disclosures during the pre-test interview. The offender is not attached to the instrument at this time (except for the acquaintance test).

The polygraph examination: the polygraph test itself follows on immediately from the pre-test interview, and takes about 20 minutes to conduct. The examiner records physiological data relating to respiration, the cardiovascular system, and skin conductance while the offender is asked a set of 10 to 12 questions that were agreed with him during the pre-test interview. Only some of these questions (which are referred to as relevant questions) relate directly to the issues of interest. All questions are very specific, and the offender is instructed to answer 'yes' or 'no' to each of them. No discussion takes place during the test. The question set is repeated three times to exclude 'rogue' responses.

Following the test the examiner evaluates and scores the chart using standardised scoring rules to reach an opinion on whether or not the test is indicative of deception; this usually takes from ten to twenty minutes. A number of different scoring systems are in use. The scoring procedures used in the pilot were based on guidelines issued by the Department of Defence Polygraph Institute (described in Krapohl, 2005). In brief, physiological reactions are given scores ranging from -3 to +3 depending on how they differ from other, comparison questions. If the summation of all the scores from the 3 charts is -6 or lower, the outcome is referred to as '*Deception Indicated*' (DI), while if it is 6 or over and all the relevant questions are positive a '*No Deception Indicated*' (NDI) outcome is given; tests scoring between -6 and +6 are said to be inconclusive. Although computerised scoring algorithms are available, they have not been validated in respect to sex offender testing, which differs in a number of technical ways from the single issue tests on which the algorithms were developed.

Post-test interview: Once the examiner has reached an opinion, the outcome is discussed with the offender during a post-test interview. If the finding from the examination is 'deception indicated', the offender is given an opportunity to account for it; even when the result is NDI or inconclusive, however, some offenders will disclose additional information at this stage.

In the pilot the entire examination was videotaped so that any discrepancies between an offender's and an examiner's account of what took place could be checked if necessary. The videotapes were also used by examiners when writing their reports, and for supervision purposes.

Types of test

Four different types of test were used in the pilot:

- **Sex history disclosure:** this test is designed to explore an offender's sexual background, in particular his history of sexually deviant behaviour. Its aim is to produce a more accurate understanding of an offender's sexual behaviour that will be of use in risk assessment and treatment. A questionnaire covering a range of behaviours forms the core of the pre-test interview.
- **Maintenance:** this test asks about behaviours associated with an offender's licence conditions or relapse prevention plan; in American programmes questions regarding reoffending would also be asked, but such questions were excluded from the pilot because of a concern that it would discourage offenders from taking part. The aim of maintenance testing is to uncover risky behaviours so that intervention can take place before they result in a reoffence.
- **Specific issue-offence:** this test asks about circumstances associated with the index offence. It may relate to whether or not an offence was committed at all, as in the case of an offender who absolutely denies the charge on which he was convicted, or it might focus on a limited aspect of the offence about which there appears to be significant denial. The aim is to assist the offender to provide a more reliable offence description, and to help overcome denial.
- **Specific issue-other** (also referred to as a monitoring exam): this test is focuses on a single issue about which there is currently concern, for example, whether an offender is in contact with a former victim.

Initially, the protocol was for sex history disclosure tests to be offered at the start of treatment, so that treatment facilitators would have the maximum amount of information available to them before the start of the group. Maintenance tests would then be undertaken at six monthly intervals, but sooner if outstanding issues remained from a previous test. However, because some offenders were on a lengthy treatment waiting list, the protocol was changed so that a maintenance test could be carried out in the first instance.

A specific issue-offence exam was typically considered if issues remained after the 'offence account' module of the treatment programme. In some cases, however, the test was offered to an offender who was precluded from treatment because of a complete denial of his offence. The aim in such cases was to test his account in the context of the polygraph examination, possibly leading to an acknowledgment of at least some aspects of the offence that would make it possible for him to take part in treatment.

Other specific issue exams were considered when probation officers had particular concerns about an offender.

Although in sex history exams offenders were asked about past victims in general terms, specific details that would enable an offence to be identified were not sought. Similarly, in maintenance tests offenders were asked about behaviours believed to increase their risk of reoffending, but not about actual reoffences. The protocols were designed in this way both because of a concern that offenders might not agree to be tested otherwise, and also because of uncertainty about the legal status of any self-incriminating disclosures.

Retests

Following a first polygraph test, the aim was to retest offenders in about six months, although the retest would be scheduled for sooner if issues needed to be clarified from the initial examination or there were concerns about the offender generally. Subsequent retests were similarly brought forward where appropriate. A number of retests took place at intervals longer than six months – according to information provided by the examiners this was usually because of difficulties in agreeing or arranging an appointment time with the offender, other scheduling problems, or the offender having defaulted from an initial appointment. Some of these difficulties were associated with the voluntary nature of testing; if testing were mandatory problems regarding offender cooperation would be discouraged by the potential imposition of sanctions.

Interaction between examiners and probation officers and the use of information

Prior to every polygraph examination, examiners contacted the offender's case manager or treatment facilitator to clarify what was known about him, establish whether there were any existing concerns, and discuss specific topics to be explored during the examination. Following the test examiners provided verbal feedback to case managers within a few days (immediately if there were issues of public safety), with a written report provided a week later. Although in practice the written reports were sometimes delayed, the verbal feedback occurred in nearly every case.

In briefings that took place in every area, probation offices were advised of the importance of distinguishing between test outcome and disclosures made during a test, along with the different implications that arise from each:

Disclosures: Probation officers were told that disclosures made during a polygraph examination were to be treated in the same manner as disclosures made in any other setting, and efforts should be made to clarify any disclosure with the offender in a subsequent interview. Except if followed by a 'No Deception Indicated' outcome, it was emphasised that disclosures should be considered to be partial only. Probation officers were able to view the videotape of an examination if they requested to do so.

Test outcome: Probation officers were informed that no action should be taken solely on the basis of a 'failed' polygraph, although such an outcome might indicate the need for further exploration or investigation. In this respect examiners indicated which questions appeared to give the offender particular difficulty. Similarly, a 'passed' test should to be put in context of all other information known about the offender and not relied upon on its own.

If offenders disclosed information regarding previously unknown offences, any decision about informing the police was made by the probation service, not the examiner.

Where offenders were managed within Multi-Agency Public Protection Arrangements (MAPPA) as being high risk (Level 3 in MAPPA terminology), test results could be shared by the case manager at meetings of the Multi-Agency Public Protection Panel (MAPPP), comprising representatives from the police, probation, and other agencies where appropriate. If asked to do so, examiners attended these meetings to elaborate on test results or to discuss strategies for further testing if asked to do so.

5. IMPLEMENTATION

Examiners and examiner training

Initially 4 examiners were trained at a specially run American Polygraph Association accredited polygraph school held at Newcastle University between 6 May and 4 July 2003; the school was set up as a satellite of the International Academy of Polygraph of Ft. Lauderdale Florida. The nine week (360 hour) training programme consisted of an 8 week basic polygraph course and a specialist one week module on post conviction sex offender testing required by the American Polygraph Association for examiners conducting tests with sex offenders. All of the examiners subsequently became members of the American Polygraph Association.

Three of the examiners had backgrounds in psychology, the other was a former policeman with experience in child protection. After they were trained, two examiners were employed full time on the project, the other two part-time (one working two days and the other three days a week), amounting to 3 full-time equivalent examiners in total; examiners were based in the Thames Valley, West Midlands and Northumbria regions, but each tested in three areas. Because of the geographic spread of the areas and a higher take-up rate than expected, a fifth part-time examiner, also from a psychology background, was recruited and trained in 2004. In addition, a sixth examiner (also a psychologist) who had been trained to work on another project assisted with tests in Northumbria and Lancashire from mid-2005. Both of these examiners trained at the International Academy of Polygraph, the same school that delivered the training in Newcastle, which meant that the training they received was similar to that provided to the original examiners.

Examiners used standardised testing and scoring procedures recommended by the Department of Defence Polygraph Institute, as described on page 16.

Examiners received supervision throughout the project from two of the original instructors from the polygraph school that organised their initial training. Three booster training sessions were held in each year of the pilot to review technique, deal with questions and problems, and ensure that protocols were being following in a similar manner by each examiner. As referred to in the previous chapter, all tests were videotaped, and some of these were reviewed in supervision. In addition, examiners were able to contact the supervisors by email or phone to deal with specific questions as they arose.

The design of the pilot was intended to ensure that all of the examiners received the same training, followed identical protocols and scoring guidelines, received appropriate monitoring and guidance, and participated in continuing professional development. Even at four monthly intervals, however, supervision reviews with the polygraph trainers uncovered occasional 'drift' in procedures as well as errors in technique, although these declined as the examiners gained in experience. In addition, modifications to recommended scoring procedures were made during the course of the pilot, and the examiners appear to have adapted to the new system at differing rates.

A degree of variability between examiners is inevitable given differences in personality and interpersonal style, but it is important to ensure that this variability does not reach a stage where it impacts on the quality of testing, or results in significant inconsistency between sites. In any future implementation, attention will need to be given during training to identify examiners unable to function within established protocols (although this was not an issue in the pilot). Based on the experience of the pilot, supervision will need to be of greater intensity in the first months of testing, and include direct observation as well as reviews of videotapes; by about six months examiners seem to be confident and able enough for supervision to become more intermittent.

The main impact of examiner inexperience in the early stages of the pilot related to the number of inconclusive tests, which decreased during the course of the project (see below). In addition, however, it was noted that case managers (but not treatment facilitators), reported significantly fewer disclosures during the first six months of an examiner's testing compared with later; the extent of this difference as reported by case managers is shown in Table 6. No differences were found in respect to other variables. This suggests that an evaluation in the context of a randomised controlled trial will need to take into account examiner learning curves.

Table 6: Case manager reports of new disclosures from first polygraph tests carried out during the initial six months of an examiner's practice compared with tests carried out after 6 months.

	First 6 months	After 6 months	Total	Significance
No disclosures	26 (39%)	28 (24%)	54	$\chi^2 = 4.24$ (df=1) P < .05
Disclosures	41 (61%)	87 (76%)	128	
Total	67	115	182	

Inconclusive tests

After the first year of the pilot, 32% of test outcomes were reported as 'inconclusive'. This rate was similar for all of the examiners. Although inconclusive tests can be of value, either because of disclosures made by the offender or because a response to one of the questions may still be suggestive, the lack of a definitive conclusion makes this a generally unsatisfactory result: offenders are often disappointed not to be given a clear result, and probation officers may find inconclusive outcomes difficult to interpret.

There appears to have been two contributing factors to the high inconclusive rate. The first relates to examiner inexperience, as inconclusive findings can be the result of problems in question formulation, interview technique, or chart scoring. This was addressed in booster training that took place in autumn 2004, and in examinations carried out after January 2005 the inconclusive rate fell to 15% (reducing the rate since the start of the pilot to 23%). However, this reduction in inconclusive findings did not occur for all of the examiners: in three the inconclusive rate since January has been in the range of 4 to 14%, while for the other two regular examiners it remains at about a

third. It is not clear whether the difference between examiners relates to examiner characteristics, examination technique, or other factors, but the inconclusive rate has now also reduced for these two examiners following more recent supervision at which question formulation and scoring parameters were further reviewed.

A second contributing factor may relate to the lack of a well defined system for tests to be followed-up by probation officers. If probation officers do not appear to respond to test findings, offenders in some instances may conclude that there is little 'consequence' associated with being deceptive. This could have the effect of reducing their responsiveness generally, making a clear outcome during polygraphy less likely.

Guidelines do not exist regarding 'acceptable' rates of inconclusive findings in sex offender testing. The American Society for Testing and Materials (ATSM, 2005; Krapohl, in press) recommends that polygraph techniques should have an inconclusive rate under 20%, which on clinical grounds would also seem to be a reasonable target.

Inter-rater reliability

Inter-rater reliability in terms of chart scoring was variable when checked at the end of the first year of testing, but improved when the charts were reviewed together as a group. A further assessment of inter-rater reliability took place after the second year of testing, with the 6 examiners each scoring the same 12 charts. The results of this exercise were as follows:

- in 3 of the charts all 6 examiners were in agreement
- in 5 of the charts 5 of the 6 examiners were in agreement
(*in each case the odd one out was an inconclusive finding*)
- in 1 chart 4 examiners were in agreement while the other 2 were inconclusive
- in 1 chart 3 examiners were in agreement and the other 3 inconclusive

Thus, in 10 of the 12 charts examiners were either in agreement with each other, or were inconclusive. In the other two charts agreement was poor:

- in 1 chart 2 examiners concluded DI, 2 NDI, and 2 were inconclusive
- in 1 chart 2 examiners concluded DI, 1 NDI and 3 were inconclusive

The results for these latter two charts were clearly of more concern given the marked variability in outcome. When the disagreements in scoring were reviewed, a consensus was usually reached, and the reasons for the discrepancy between scores understood. One characteristic of the two charts for which there was poor agreement was their poor quality, which meant that greater interpretation than normal was required. These findings, however, demonstrate the need for a more regular review of examiner scoring

than the annual exercise conducted in the pilot. Ideally, only minimal disagreement should be seen when ‘clean’ charts are scored by more than one examiner.

Test administration

Frequencies for the different test types carried out over the course of the pilot are listed in Table 7 (descriptions of the test types can be found on page 18). It can be seen that sex history examinations were the most common type of first test, while maintenance testing accounted for three quarters of retests.

Table 7: Frequency of polygraph test types.

	First tests	Retests	All tests
Sex history disclosure	185 (53%)	12 (9%)	197 (41%)
Maintenance	103 (30%)	104 (76%)	207 (43%)
Specific issue – offence	57 (16%)	8 (6%)	65 (13%)
Specific issue – other	2 (0.6%)	12 (9%)	14 (3%)
Total	347	136	483

In a number of cases, the offender did not remain for the post-examination feedback, which meant that a proper debrief with the examiner did not happen – because the test was voluntary there was nothing that could be done about this, but as this is part of an offender’s co-operation with testing this is an issue that could be addressed in the context of a mandatory testing programme.

Two issues arose in relation to the sex history exam. First, some offenders found it difficult to discuss their sexual histories prior to starting treatment as they had not yet become accustomed to such focused attention on their sexual behaviour (for this reason, many American programmes do not conduct this test until the offender has been in treatment for a number of months). On the whole, however, treatment facilitators preferred having this information prior to the commencement of treatment. Second, the questionnaire used in the exam covers such a vast array of behaviours that it can be difficult for an offender to be confident he is responding accurately, leading to a ‘wrongly’ failed test. During the polygraph examination there are in any case limits to the depth that this information can be explored. In many American programmes the questionnaire is completed by offenders with their probation officers as part of the treatment regime before the examination takes place, which both shortens the length of the test and allows the examiner to focus on specific areas of interest.

As referred to on page 18, offenders were not asked about actual reoffences or unknown past offences, except in general terms. However, boundaries sometimes became blurred, for example when a probation officer suspected that an offender may have acted in breach of his licence conditions or actually reoffended, and the

polygraph examiner was requested to question the offender about relevant behaviours in a specific issue monitoring exam. In such cases offenders were asked to give additional, explicit consent to this type of questioning, which could lead to an admission of a reoffence. In general, more work needs to be done regarding how such information can be sought without compromising the cooperation of offenders, or resulting in legal complications.

Integration with programmes

Polygraph testing appeared to become firmly rooted in three of the areas from early on in the project: Northumbria, the Eastern Region, and Leicestershire. In these areas it integrated well with programme management, and became seen as part of the routine provision for sex offenders. Manchester, Lancashire, and Northamptonshire took longer to achieve this integration, but it seems to have occurred in Year 2, resulting in increased take up and better utilisation of test results. In these areas there appeared to be a high awareness of polygraph testing, with offenders being approached without the need to remind probation officers of the availability of testing, and regular discussions taking place between treatment managers and polygraph examiners. In these 6 areas the key factor underlying successful integration appeared to be the commitment of the programme and treatment managers to testing, and their view that the findings were contributing in important ways to the treatment and supervision of offenders.

The remaining four areas faced significant internal difficulties in keeping their programmes functioning because of sickness, high staff turnover, and unfilled posts, quite apart from polygraphy. Given other pressures, polygraph provision remained more on the periphery in these areas, dependent on a small number of treatment facilitators and case managers putting forward offenders for testing, although the enthusiasm of these individuals meant that a reasonable number of men were nonetheless tested (with the exception of Devon and Cornwall, which had the additional limitation of being able to provide treatment for only a small number of offenders during much of the testing period).

In the course of the pilot, a number of other issues effecting implementation emerged:

Case manager involvement: Because of a high turnover of case managers (the probation officers responsible for the supervision of offenders) in many of the sites, relying on case managers to discuss polygraph testing with offenders and to obtain their consent was of variable consistency, even in those areas where testing integrated most successfully. In spite of extensive briefings in each area at the start of the pilot, midway through the testing period there were significant numbers of new case managers who had only a limited understanding of the project and of their role in it. In a mandatory programme this would represent less of an issue as case managers would have little choice but to be well briefed in the procedure and ensure that offenders attend for examinations, but it would also make it imperative for systems to be in place enabling case managers to acquire the necessary information and management tools.

The use of test findings by case managers: Although test findings were fed back verbally to case managers and reports were sent to them soon after, no system was in place to ensure that the results fed directly into case management. While many case managers were able to make good use of polygraph findings, illustrated by case examples in this report, the outcome of polygraph tests was not always discussed by case managers with offenders, nor by case managers with their seniors during supervision; the inexperience of some case managers in supervising sex offenders generally was also a limiting factor in their ability to make full use of test results. Not only did this mean that opportunities to intervene were sometimes missed, but as described above it also potentially weakened the responsiveness of offenders in further testing.

The following cases are examples of these difficulties. The first an instance where no use was made of an important relevant disclosure, the second where a case manager was unsure of how to use information arising from a polygraph test:

A case manager was told both verbally by the examiner and in the written report that an offender had admitted to masturbating to fantasies of children during the polygraph examination, something he had consistently denied in spite of his conviction for a sex offence against a child. Six months later, when the examiner contacted the case manager to discuss the offender's upcoming retest, the case manager reported that in the previous week he had admitted for the first time his sexual arousal to children, clearly being unaware of his disclosures from the first polygraph examination; it emerged that she had not discussed the outcome of the test with the offender at all.

An offender disclosed in a polygraph examination that he had had unsupervised contact with young girls, he was using his mobile phone to photograph girls on public transport, as well as other breaches of his licence. No action was taken at that time, however, as it was felt that these disclosures needed to be "confirmed" from other sources; as clarified at a subsequent review, had the offender made these disclosures in a context outside polygraphy, immediate action would have been taken.

In contrast, the case below illustrates the effective use of polygraph information effectively by a case manager:

An offender who had been convicted on two occasions of downloading indecent images of children on his computer was thought to have made good progress in treatment. He had been living apart from his family since his arrest, but at the time of his third polygraph examination he had recently returned home. During the polygraph interview, however, he disclosed being sexually aroused when the daughter of a neighbour sat on his lap while visiting his children, and also further downloading of indecent images on a family computer. He subsequently discussed this further in treatment. Social Services were informed and he again moved out of the family home; the police were also contacted by probation and the computer was seized for forensic evaluation. He has, however, continued to participate in testing. He

had a fourth, maintenance test, which he passed, providing reassurance to those involved in his supervision and treatment.

As the first two examples above demonstrate, attention needs to be given to ensuring that case managers are aware of relevant information that emerges from a polygraph examination, that they incorporate it into their supervision, and that they are clear about its status (in other words, because information arises from a polygraph examination does not make it any less, or any more, reliable than disclosures made by offenders in other settings).

In response to this, in the Eastern Region the treatment manager now reviews all polygraph reports, following which a summary form is sent to the case manager that highlights the key issues from the polygraph examination and recommends actions to be taken, with a timeframe for their completion. This system is being considered elsewhere.

Following a review of the pilot at the end of the two year study period, 4 of the areas planned to issue instructions and guidance to the effect that polygraph results should be discussed routinely in the next supervision meeting that takes place between case manager and offender following the examination. In 2 other areas the plan was for senior officers to be advised to review polygraph reports in their supervision sessions with case managers. The aim was for these changes to be in place for the additional year of testing that followed on from the pilot..

The use of test findings in treatment: Because the treatment programmes were developed prior to the introduction of polygraphy, there is no place within them for polygraph results to be used explicitly unless raised by the offender himself; this of course is not an issue in the context of one-to-one treatment where probation officers have been able to make immediate use of the findings. This is unlike American programmes, where the offender will report back to the group following his polygraph test. In the review meetings held in each area, treatment facilitators (the probation officers responsible for delivering treatment on the SOTP) indicated that at present their use of the results is confined to gaining a fuller understanding of the offender, particularly in respect to sexual background, with more specific discussion of test results not taking place until the three-way meetings that occur between group modules between them, the offender and the case manager.

Given the manualised nature of the treatment programmes this issue can only be addressed through revision of the manuals, which would need to be considered should a further trial of polygraphy be carried out.

Administrative difficulties: Administrative difficulties of one sort or another arose at all of the sites at some stage, but where an administrative assistant was assigned to coordinate appointments, manage the paperwork and organise room bookings testing was generally observed to run more smoothly (although a high turnover of administrators in some sites has been a further complicating factor). In some areas examiners had to take much of this on themselves, resulting in the loss of time that could have been devoted to testing or report writing, and on occasions resulted in

offenders being unaware of appointments or a delay in organising retests. In addition, finding rooms in which to carry out tests was at times been problematic in some areas given the heavy use made of probation facilities.

Some of the sites are developing data bases to assist in tracking offenders and their polygraph status. Other issues were addressed on an ad hoc basis as they arise. There is a clear need, however, for a small amount of dedicated administrative time supplied by probation areas to complement the systems maintained by examiners if polygraphy is to be delivered in an efficient manner.

Examiner competencies and characteristics: Experience from the pilot indicated that two different types of competency are needed by examiners, one relating to their interview skills and ability to formulate appropriate questions, the other to their ability to master polygraph techniques (running the equipment, understanding the software, scoring the charts, etc.). The latter is more easily taught than the former. In selecting potential examiners, therefore, more attention is probably best given to a candidate's experience and aptitude in interviewing, but both must be rigorously assessed at the end of training.

Examiner characteristics also appeared to be an important component in achieving successful integration of polygraphy with existing programmes. As referred to above, in the current pilot some examiners are employed full time on the project, while others are only part time. Because the schedules of the former are more flexible and they do not have the competing work demands from outside the project that are faced by the part-timers, it has been easier for them to establish a 'presence' in their areas. In addition, the more assertive examiners, who chased-up referrals and actively followed-up cases, also appear to have become better integrated with the work of the sex offender teams.

In summary, based on feedback from examiners and probation officers and on observations of practice in the different areas, the efficient implementation of polygraph testing appears to be dependent on the following factors:

1. The active involvement of the sex offender team programme and treatment managers, and in turn the support of their higher managers.
2. Keeping case managers up to date with the nature of polygraph provision and their role in it.
3. Establishing systems to support case managers in their understanding and use of polygraph findings.
4. Developing protocols to integrate polygraph testing with treatment.
5. Good administrative arrangements.
6. Examiners with good competencies both in interviewing and in polygraph technique.

7. Full time examiners who interact assertively with case managers and members of the treatment team.

Defaults

A significant problem for examiners was the number of men who either cancelled appointments at short notice, or who simply did not turn up for them. Sometimes this represented a change of mind regarding testing, but a range of other excuses were given, including sickness, transport problems, forgetting, or other commitments, and many of those who defaulted from a polygraph examination were in fact tested at a later date. In total, approximately 250 of the 733 testing sessions were lost because the offender did not attend, amounting to about a third of the examiners' testing time.

This high default rate is probably inherent to the voluntary basis of the pilot, and the lack of any consequence for not attending. Discussion with practitioners in North America, where offenders typically have to pay for their examinations, and where they may be ejected from treatment programmes or be in breach of their licences if they are not tested, suggest that default rates there are low, but there is no published data in relation to this and it is difficult to know the accuracy of their impressions. The poor involvement of some case managers referred to above was also a factor influencing the commitment of offenders to testing: whereas a number of case managers would immediately seek out reasons why an offender did not attend and encourage him to keep a future appointment, on occasions organising transport for him, others took no action at all (which of course would not be an option if testing were mandatory).

A high default rate has clear implications for the efficient and cost effective provision of polygraphy, and defaults need to be factored into any programme where polygraph testing is voluntary. Even in a mandatory scheme, however, clear guidelines need to be in place for how to deal with defaults when they occur.

Capacity analysis

In order to carry out a polygraph examination an examiner must first obtain relevant information about the offender from probation files and through discussion with case managers or treatment facilitators. Depending on the case, this requires one to two hours of the examiner's time, although incomplete files and difficulties in contacting probation officers can prolong the process.

The polygraph examination itself typically lasts from two to four hours, with sex history disclosure exams being on the long side and retests generally taking less time. Following the test, the examiner will review his or her notes, the charts, and the recording of what took place in order to prepare a report. This requires a further one to two hours on average. In addition, one would almost certainly want examiners to discuss the examination with case managers or treatment facilitators as soon as possible after the test is completed, as was the case during the pilot.

In total, therefore, a typical test will require 4 to 8 hours of an examiner's time. It might be possible to reduce this, for example by case managers completing referral summaries, or probation officers completing the sex history disclosure interview with offenders as referred to above. However, it is difficult to see how proper sex offender testing could involve less than 3 hours of an examiner's time, with 4 hours being a more realistic target. Thus, while in some circumstances three tests a day might be feasible (for example, if they were all relatively straightforward retests), two tests a day is a more realistic norm; carrying out more than 3 tests a day on a regular basis should give rise to questions about the standard of testing.

On this basis, over a week an examiner could be expected to conduct from 10 to a maximum of 15 tests, although more than 10 tests a week is probably not sustainable. It would be reasonable to set a target of 8 to 10 tests a week per examiner (assuming that it may not be possible to test every day), amounting to 320 to 420 tests a year, taking into account holiday and continuing professional development, but not the time needed for administration or travel. The latter is a particular issue if a small number of examiners are required to work into a large number of sites at some distance from one another.

The target of 320 to 420 tests a year does not take into account offenders cancelling or failing to show up for testing without giving sufficient notice to allow another offender to be tested in his place. As found in the pilot, in a voluntary scheme defaults may take place at a rate of one for every three tests, which would reduce the number of tests actually completed in a year to between 210 and 280 per examiner. It is difficult to estimate the number of defaults that would occur in the context of a mandatory programme, although it is likely to be considerably less; a sense of what the default rate might be could be obtained from the experience of American programmes, although this would need to be obtained from discussion with them as we are not aware of any relevant published data.

Implementation conclusions

In summary, observations from the pilot indicate that introducing polygraph testing to probation sex offender treatment programmes is feasible and practical, and can be implemented effectively. The most important lesson learned from the pilot, however, is that successful implementation of polygraph testing in the probation service will depend on the process being seen as an integral component in the treatment and supervision of offenders, with the allocation of sufficient management time to ensure that this takes place.

In any future trial it will be important to ensure that examiners are appropriately monitored and supervised, with regular checks on inter-rater reliability. Although this is particularly necessary for inexperienced examiners, a means of ensuring ongoing supervision will be critical to prevent drift in scoring procedures and the potential development of poor practice; the lack of this in the United States has been an important contributor to criticisms there regarding variations in practice and failure to adhere to standards (for example, Cross & Saxe, 2001).

In the long term, ensuring that monitoring and supervision takes place outside the confines of a trial will be difficult unless it forms part of a regulatory regime designed to ensure good practice in polygraphers.

6. FINDINGS

Examination outcome

Test outcome for each of the different examination types is shown in Tables 8.

Table 8: Test outcome in relation to the type of polygraph examination carried out: sex history disclosure (in which an account of the offender's sexual background is obtained), maintenance (in which questions are asked about current behaviour), specific issue – offence (in which the offender is asked about specific aspects of the index offence), and specific issue – other (in which a particular issue is explored).

	<u>Sex History</u>	<u>Maintenance</u>	<u>SI–Offence</u>	<u>SI–Other</u>	<u>Total</u>
No decep. indicated	46 (23%)	48 (23%)	6 (9%)	3 (21%)	103
Deception indicated	91 (46%)	84 (41%)	31 (48%)	8 (57%)	214
Inconclusive	38 (19%)	59 (29%)	9 (14%)	2 (14%)	108
Pre-test admissions/ no test	0	3 (1%)	11 (17%) ^{***}	0	14
Invalid – any reason	15 (8%)	8 (4%)	5 (8%)	0	28
Exam terminated before complete	7 (4%)	5 (2%)	3 (5%)	1 (7%)	16
Total	197	207	65	14	483

p < .001

There were no significant differences in outcome between the different test types, apart from a greater number of pre-test admissions in specific issue–offence examinations making testing while attached to the polygraph redundant ($\chi^2 = 47.29$, $p < .001$). In these cases, the test was typically intended to obtain an offence account from an offender who had been denying what he had done – when this objective was met in the pre-test interview, there was no further need to question him while attached to the polygraph.

Pre-test admissions in specific issue–offence examinations were particularly observed with offenders who had been convicted of internet related crimes, occurring in 6 of 17 such men (in contrast, 8 of the 10 rapists undergoing specific issue–offence examinations were found to be deceptive). The pre-test interview in these cases focussed on the reason indecent images of children were found on their computers, with disclosures frequently describing a sexual interest in children dating back many years that had been denied in both treatment and supervision.

Retests

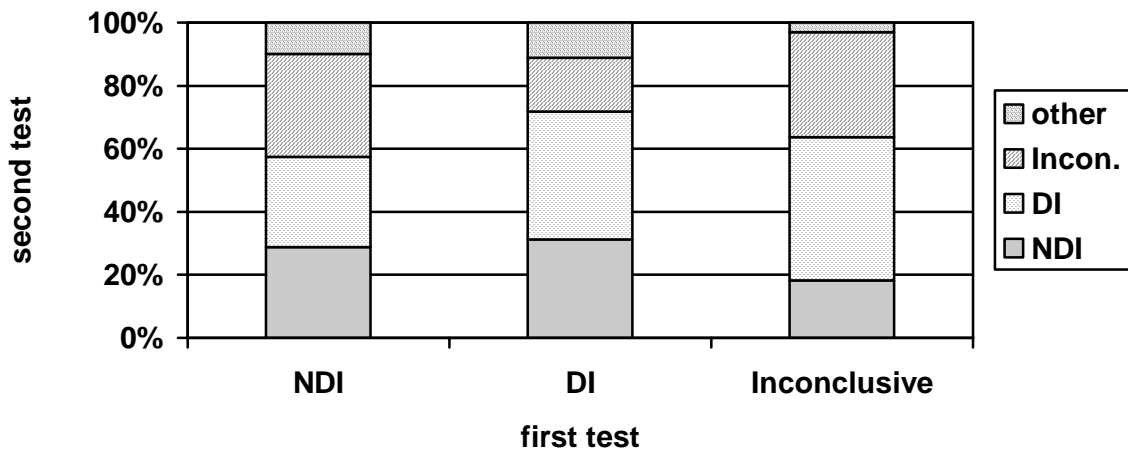
Looking at first tests and retests as a whole, outcomes were similar, with no significant differences observed in the proportion of tests reported as 'no deception indicated', 'deception indicated' and 'inconclusive' (Table 9).

Table 9: Test outcome for first polygraph examinations and retests.

	First tests	Retests	All tests
No deception indicated	68 (20%)	35 (26%)	103 (21%)
Deception indicated	161 (46%)	53 (40%)	214 (44%)
Inconclusive	73 (21%)	35 (26%)	108 (23%)
Test not administered due to pre-test disclosure	12 (4%)	2 (1%)	14 (3%)
Invalid – suspected us of countermeasures	7 (2%)	4 (3%)	11 (2%)
Invalid – other reasons	15 (4%)	2 (1%)	17 (4%)
Exam terminated	11 (2%)	5 (4%)	16 (3%)
Total	347	136	483

The 136 retests were carried out in 116 offenders, with 17 tested three times and 3 on four occasions. No clear pattern emerged regarding outcome on first and second tests. It can be seen in Figure 3, for instance, that of the offenders who were NDI on their first test, 29% were also NDI on their second test, but a similar proportion (29%) were DI; likewise, of those who were DI on their first test, 31% were NDI when tested a second time, but 40% were again DI.

Figure 3: Results of second polygraph examination compared with outcome on first test (n = 21 NDI on first test, 52 DI, 33 Inconclusive).



Follow-up data

Actions Taken forms

Polygraphed offenders

As described on pages 8 and 9, 'Actions Taken' forms were sent to case managers and treatment facilitators to obtain feedback regarding the impact of polygraph testing. Completed forms were received in relation to 68% of tests undertaken. Return of a feedback form was not dependent on whether offenders 'passed', 'failed', or received an 'inconclusive' finding, but significantly fewer forms were returned in relation to the 33 first tests where the examination was invalid (for example, because of recording or equipment problems), or had to be terminated ($\chi^2 = 13.19, p < .001$); as this represented only 17 forms, the impact of this on the research findings was minimal.

Over three quarters of forms were returned for sex history exams, about two thirds for maintenance exams, just over half for specific issue–offence exams, and just over a third for other specific issue exams, giving an under-representation in relation to the specific issue type tests ($\chi^2 = 9.94, p < .01$), which means that conclusions regarding this type of test based on probation officer feedback alone need to be made cautiously.

There was no evidence that age, ethnicity, Risk Matrix category, offence type, and whether the victim of the index offence was a child or adult were related to the likelihood of a form being returned.

Non-polygraph offenders

Actions taken forms were completed in respect of 58% (n=180) of the non-polygraph offenders. There did not appear to be any marked differences on key variables between those for whom we obtained forms and those for whom we did not: there

were no significant differences in terms of age, Risk Matrix category, offence type, or whether the victim of the index offence was a child or adult. Details of ethnic background were available for only 62% (n=112) of the offenders for whom we had forms, making comparisons regarding this variable unreliable.

The 64 non-polygraph offenders for whom we received second forms in relation to a second follow-up period had higher Risk Matrix scores compared with the 115 offenders from the comparison sites for whom we had one form only: amongst the former, 33% of the offenders (18 of 55) for whom Risk Matrix data was available were rated as high or very high, compared with 17% (12 of 71) for whom we had just the one form ($\chi^2 = 6.51, p = .01$).

The variable and uneven return of Actions Taken forms by the non-polygraph case managers means that the extent to which findings in respect to the comparison areas are representative of all relevant sex offenders in these areas is uncertain.

Pilot v non-polygraph samples

Polygraphed offenders for whom Actions Taken forms were completed did not differ significantly in terms of their age, Risk Matrix scores or offence characteristics from offenders in the non-polygraph sample for whom we had forms. Because of the lack of information regarding ethnic background for many in the non-polygraph group for whom forms were completed, it is unclear whether there were differences between the two groups in this respect.

Although as indicated above Risk Matrix scores were higher in non-polygraph offenders for whom we had second forms than in those for whom we had one form only, their Risk Matrix scores were similar to retested polygraphed offenders for whom we had forms.

Disclosures

Based on examiners' reports

Part way through the pilot examiners were told to highlight in their reports information disclosed during the polygraph interview that they believed to be new, differentiating between whether this occurred in the pre- or post-test. They were also asked to review their previous tests to identify new disclosures based on documentation and notes from the time. Such retrospective assessments were carried out in respect to about a quarter of tests; examiners were unable to determine whether new information had been disclosed in 4% of cases (n=18). Ratings made retrospectively may have been biased in either direction by the need to make judgements post hoc, but there was no difference between retrospectively and contemporaneously rated tests in terms of the number considered to contain new disclosures.

New disclosures were classed as being either sexual or non-sexual (for example, relating to alcohol or drug use, or travel away from the area) in nature. Excluding the

18 first tests that could not be rated, examiners indicated that in their opinion nearly 80% of examinations resulted in at least one new disclosure relevant to treatment or supervision. Most of these new disclosures were of a sexual nature (Table 10).

Table 10: Clinically relevant new disclosures made during polygraph examinations according to examiners' ratings. Some offenders made both pre-test and post-test disclosures, which means that the number making 'any disclosure' is less than the total of pre- and post-test disclosures.

	First Test	Retest	Total
Pre-test disclosure, non-sexual	56 (17%)	25 (19%)	81
Pre-test disclosure, sexual	241 (73%)	87 (66%)	328
Post-test disclosure, non-sexual	7 (2%)	4 (3%)	11
Post-test disclosure, sexual	74 (22%)	35 (27%)	109
Any disclosure	261 (79%)	102 (78%)	363
No disclosure (missing)	68 (21%) (18)	29 (22%) 0	97
Total	329	131	460

It can be seen from Table 10 that 33% of the disclosures (120 of 363) occurred in the post-test interview. Twenty per cent (n=24) of these post-test disclosures took place in tests where no disclosures at all had been made in the pre-test interview.

Disclosures more typically took place in the context of sex history and maintenance examinations compared to specific issue ones. They were not associated with test outcome overall, with a similar proportion taking place regardless of whether tests were passed, failed, or inconclusive (Table 11).

Post-test disclosures occurred significantly more frequently after 'Deception Indicated' tests than after tests with 'No Deception Indicated' or 'inconclusive' outcomes; they also occurred significantly more frequently after inconclusive than NDI tests (Table 12).

Table 11: Disclosures as recorded by examiners made in the different test types, and according to test outcome (No Deception Indicated, Deception Indicated and inconclusive outcomes only).

	No Disclosure	Disclosure	Total	Significance
Exam Type				
Sex History	29 (15%)	168 (85%)	197	$\chi^2 = 29.09$ (df=2) p < .0001
Maintenance	37 (19%)	159 (81%)	196	
Specific Issue	31 (46%)	36 (54%)	67	
Total	97	363	460	
Test Outcome				
NDI	10 (10%)	88 (90%)	98	NS
DI	77 (38%)	127 (62%)	204	
Inconclusive (Missing)	24 (23%)	79 (77%)	103 (17)	
Total	111	294	405	

Table 12: Disclosures as recorded by examiners made during the post-test interview following different test outcomes (missing in for 17 tests).

	No Post-Test Disclosure	Post-Test Disclosure	Total	Significance
NDI	17 (17%)	81 (83%)	98	$\chi^2 = 26.41$ (df = 2) p < .0001
DI	44 (22%)	160 (78%)	204	
Inconclusive	23 (22%)	80 (78%)	103	
Total	84	321	405	

As illustrated in Table 13, disclosures were not associated with Risk Matrix category or Deviance scores.

Table 13: Relation between disclosures and offender Risk Matrix and deviance scores.

	No Disclosure	Disclosure	Total	Significance
Risk Matrix score				
Low	20 (17%)	97 (83%)	117	NS
Medium	28 (25%)	85 (75%)	113	
High	12 (18%)	56 (82%)	68	
Very High	6 (22%)	21 (78%)	27	
Total	66	259	325	
Deviance score				
Low	11 (14%)	70 (86%)	81	NS
High	10 (26%)	29 (74%)	39	
Total	21	99	120	

Based on probation officer feedback

As described in Chapter 3, Actions Taken forms were completed by case managers in relation to 53% of first tests and 40% of retests. Actions taken forms were also completed by 58% of case managers for non-polygraph offenders, of whom 36% completed forms for a second time period.

Table 14 shows that case managers for the polygraphed offenders reported that new disclosures relevant to supervision or treatment were made by offenders in 70% of the first tests for which we had feedback. This compared with case managers for non-polygraphed offenders reporting disclosures in respect to 14% of offenders when initially surveyed (Time 1). The odds of a polygraphed offender making a disclosure relevant to his treatment or supervision were 14 times greater than a non-polygraphed offender making such a disclosure². The confidence interval relating to this odds ratio is wide (ranging from about 8 to over 24), but it is well above 1, suggesting a meaningful clinical impact. The incomplete and uneven return of Actions Taken forms, however, means that care must be taken in generalising from these findings.

As shown in Table 14, a similarly large difference was found in respect to case manager reports of the number of offenders making disclosures during retests (67%) compared with the number of non-polygraphed offenders (13%) doing so during a second period of time (Time 2), although feedback was available on fewer offenders.

² The odds ratio is defined as the ratio of the probability of an event occurring in one group to the probability of the event occurring in a comparison group. It is a way of comparing whether the probability of a certain event is the same for two groups. An odds ratio of 1 implies that the event is equally likely in both groups. An odds ratio greater than one implies that the event is more likely in the first group, while an odds ratio less than one implies that the event is less likely in the first group.

Table 14: Clinically relevant new disclosures reported by case managers of polygraphed offenders after their first tests and retests, and by case managers of non-polygraphed offenders at comparable times. Because of the incomplete return of feedback forms and possible differences between the groups caused by the voluntary nature of the study and the absence of randomisation, these findings are indicative.

	Case Mangers Polygraph	Case Managers Non-polygraph	Significance
New Disclosure	First test	Time 1	$\chi^2 = 114.65$ (df=1) $p < .0001$ odds ratio = 14.4 (C.I. 8.5 – 24.4) ¹
Yes	128 (70%)	25 (14%)	
No	55 (30%)	155 (86%)	
Total	183	180	
New Disclosure	Retest	Time 2	$\chi^2 = 34.47$ (df=1) $p < .0001$ odds ratio = 14.0, (C.I. 5.6 – 35.0) ¹
Yes	36 (67%)	8 (13%)	
No	18 (33%)	56 (87%)	
Total	54	64	

¹confidence interval 95%

In the non-polygraph cases, new information from a source besides the offender was reported in a further 7% (n=12) of cases. Even when this additional information is combined with offender disclosure, the number of new disclosures reported in polygraph cases remains significantly greater ($\chi^2 = 87.30$, $p < .0001$; odds ratio 9.0, 95% C.I. 5.6 – 14.5), although due to differences between the two groups as discussed in Chapter 7 these findings are indicative.

According to treatment facilitators of polygraphed offenders, new disclosures were made in 71% (76 of 107) of first tests for which they completed forms, and in 72% (23 of 32) retests (information was not available from treatment facilitators of non-polygraphed offenders for comparison). Overall, case managers or treatment facilitators reported new disclosures in 73% (179 of 245) first tests, and in 69% (55 of 80) retests.

Forms were completed for the same offender by both case manager and treatment facilitator in relation to the first tests for 45 of the polygraphed offenders. There was agreement between them in terms of new disclosures in 78% of them (n=35); in 6 the treatment facilitator indicated that new information was disclosed while the case manager did not, with the reverse being the case in 4 tests. Whether disagreement in respect of these 10 tests is the result of case managers and treatment facilitators having different degrees of knowledge about the offender, or whether it is a reflection of one or the other missing new information contained in the examiner's report, is unclear, and could not be tested within the design of the study.

Similar to what was found in relation to the examiners recording of new disclosures, neither test type nor test outcome were associated with probation officer reports of new information having been disclosed during a test.

Agreement between examiners and probation officers

When examiners' reports and Actions Taken forms were compared in the 360 tests for which both were available, there was agreement between them regarding new disclosures in nearly 80% of cases (Table 15). The kappa for agreement between probation officers and examiners in relation to the presence or absence of new disclosures was 0.4 (95% confidence interval .29 -.50), falling within the 'fair' to 'moderate' range (Landis & Koch, 1977).

Table 15: Agreement between examiners and probation officers regarding tests in which new disclosures were made.

	Examiner and Prob. Officer Agree	Examiner Yes Prob. Officer No	Examiner No Prob. Officer Yes	Total
Case Manager/ Examiner	175 (76%)	42 (18%)	11 (5%)	228
Treatment Facilitator/ Examiner	105 (80%)	20 (15%)	7 (5%)	132
Total	280 (78%)	62 (17%)	18 (5%)	360

Based on the information available, it could not be determined whether discrepancies regarding new disclosures were the result of examiners believing something to have been previously unknown that was not, a delay in probation officers completing the forms and thus being unclear about what was known at the time of the test, or probation officers failing to note new information in examiners' reports.

Seriousness of disclosures as reported by probation officers

The new disclosures reported by probation officers were rated in terms of their 'seriousness' based on the following broad definitions:

- low: passive types of behaviour, such as masturbating to deviant fantasies
- medium: potential preludes to reoffending, such as going to places where there are potential victims
- high: specific breaches or actual offending

The distinction between 'medium' and 'high' can at times be blurred as there are breaches (for example, for drinking, or breaking a curfew) that might seem less serious than some 'medium risk' behaviours (such as having unsupervised contact

with a child). For present purposes this is not a major issue as both relate to *behaviours* associated with risk, rather than the thoughts or urges included within the ‘low’ category.

Table 16 shows the seriousness of disclosures as reported by case managers of polygraphed and non-polygraphed offenders, and by treatment facilitators of polygraphed offenders. The degree of seriousness of disclosures was not significantly different between the groups after the first test (Time 1 for the non-polygraph group). However, significantly more of the disclosures reported by polygraph case managers after retests were of high seriousness. Caution must be taken in interpreting this latter finding, however, because of the low numbers.

Table 16: Degree of seriousness of new disclosures reported by case managers of polygraphed offenders after their first tests and retests, and by case managers of non-polygraphed offenders at comparable times. For definitions of severity ratings see text.

	LOW	MEDIUM	HIGH	TOTAL	Significance
Case manager, first test	80 (63%)	35 (27%)	13 (10%)	128	NS
Treatment fac., first test	44 (58%)	23 (30%)	9 (12%)	76	
Non-polygraph time 1	15 (60%)	5 (20%)	5 (20%)	25	
Case manager, retest	23 (40%)	11 (19%)	23 (40%)	57	$\chi^2 = 17.78$ (df = 4) p = .01
Treatment fac., retest	19 (83%)	4 (17%)	0	23	
Non-polygraph Time 2	4 (40%)	4 (40%)	2 (20%)	10	

In the polygraphed offenders, disclosures that took place in sexual history or specific issue–offence examinations tended to relate to past actions rather than to current behaviours of concern. Maintenance and monitoring examinations, however, relate to current behaviour, and because of this disclosures of medium or high risk behaviours during them are likely to be of more immediate relevance, possibly being indicative of preludes to offending (or actual offending). According to the case manager Actions Taken forms, over a quarter of maintenance or monitoring examinations resulted in disclosures that were rated as being of ‘medium’ seriousness, and 7% were in the high category, amounting to about a third of the disclosures made in these types of test. This is shown in Table 17. A similar finding was not observed in the treatment facilitator feedback.

Table 17: Seriousness of new disclosures made during different test types, based on case manager report.

	NO DISCLOSURE	LOW	MED.	HIGH	TOTAL
Sex History	26 (22%)	63 (54%)	21 (23%)	6 (5%)	116
Specific Issue-Offence	9 (47%)	8 (42%)	0	2 (11%)	19
Maint./Mon.	29 (31%)	34 (36%)	25 (26%)	7 (7%)	95
TOTAL	64 (28%)	105 (46%)	46 (20%)	15 (7%)	230

Probation officers were not asked about the extent to which they were able to confirm what was said by offenders, but none referred to any offender later denying what he is reported to have disclosed in a polygraph examination. This would in any case be difficult given that video recordings of the tests could be checked.

Risk assessment and intervention

Polygraph and non-polygraph case managers revised their risk assessments with similar frequency. However, it can be seen in Figure 4 that the polygraph case managers more typically *increased* risk – 81% (29 of 36) of their risk modifications were increases – while non-polygraph case managers more often *reduced* risk – 55% (17 of 31) of their changes in risk assessment being decreases ($\chi^2 = 7.60$, $df=1$, $p < .01$; odds ratio 5.0, 95% C.I. 1.7 – 14.6), although as discussed in Chapter 7 differences between offenders mean that these findings are indicative. Similar to the polygraph case managers, the polygraph treatment facilitators also tended to revise their determination of risk upwards when modifying their risk assessments, doing so in over 90% of their reassessments.

Figure 4: Percentage of cases for which there was a change in risk assessment, and the proportion of changes which resulted in an increase in risk, as reported by polygraph offender case managers (n=177), polygraph offender treatment facilitators (n=114), and non-polygraph case managers (n=180).

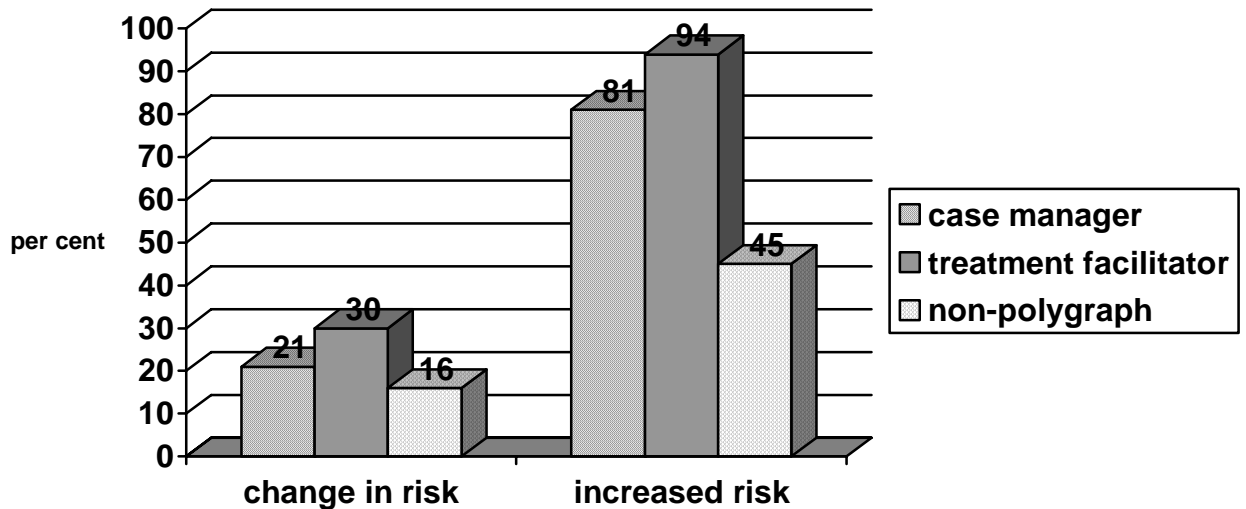


Table 18 shows that case managers of polygraphed offenders reported increasing their risk assessment in a higher proportion of offenders than did the case managers of non-polygraphed offenders, although again differences between the two groups mean that these findings are indicative. They also reported more offenders in whom treatment changes were made. No differences were reported in terms of changes in supervision arrangements or the initiation of other interventions. The odds of a polygraph case manager reporting at least one effect in terms of increased risk assessment, a change in supervision, a change in treatment, or the initiation of another intervention were twice that of a non-polygraph case manager reporting one of these changes.

Findings in relation to treatment facilitators are shown separately in Table 19 as comparable information from treatment facilitators of non-polygraphed offenders is not available.

Table 18: Changes in risk assessment, supervision, treatment and the introduction of other interventions reported by polygraph and non-polygraph case managers. Due to differences between the two groups as discussed in Chapter 7, these findings are indicative.

	First tests	Non-polygraph	Significance (chi square tests; 95% confidence intervals for odds ratios)
Increased risk	29 (16%)	14 (8%)	p = .01 (df=1) OR = 2.3 (CI = 1.2 – 2.5)
Decreased risk	7 (4%)	17 (9%)	p < .05 (df=1) OR = 0.40 (CI = 0.16 – 0.95)
Change in supervision	31 (18%)	24 (13%)	NS
Change in treatment	34 (19%)	13 (7%)	p < .001 (df=1) OR = 3.1 (CI = 1.6 – 6.0)
Other interventions	24 (14%)	20 (11%)	NS
At least one effect *	73 (41%)	49 (27%)	p < .01 (df=1) OR = 1.9 (CI = 1.2 – 2.9)
No effect*	104 (59%)	131 (73%)	p < .01 (df=1) OR = 0.53 (CI = 0.34 – 0.83)
Total	177	180	

* defined as an increase in risk, change in supervision, change in treatment, or introduction of other intervention

Table 19: Changes in risk, treatment and the introduction of other interventions as reported by treatment facilitators for polygraph offenders following first tests.

	All first tests
Increased risk	32 (28%)
Decreased risk	2 (2%)
Change in treatment	23 (20%)
Other interventions	10 (9%)
At least one effect *	50 (44%)
No effect*	64 (56%)
Total	114

* defined as an increase in risk, change in supervision, change in treatment, or introduction of other intervention

The following is an illustration of changes in risk assessment, supervision arrangements, and a referral to another agency that took place following a polygraph test:

A 58 year old man was on a Community Rehabilitation Order having been convicted of internet offences. At the time of his polygraph examination he was on a relapse prevention treatment programme, and was considered to be low risk. He had a granddaughter with whom his contact was always meant to be supervised. The case manager was concerned that he was still using the internet, which the offender denied. During his polygraph exam he disclosed regular use of internet chat rooms, unsupervised contact with his granddaughter, and frequent interaction with two girls in his street aged 8 and 9, one of whom had been into his home a number of times. Based on this information, his case manager increased his risk level, stepped up her supervisory contact with him, and informed social services of the situation regarding his granddaughter.

Agreement between case managers and treatment facilitators regarding risk

There were 50 first tests for which case managers and treatment facilitators provided feedback for the same offender in respect to changes in risk assessment:

- in 63% (n=30) of cases their responses were similar
- in 23% (n=11) only the treatment manager modified the risk assessment
- in 19% (n=9) only the case manager modified the risk assessment

The reason for the disparity is unclear, but it may relate to the different perspective case managers and treatment facilitators have of an offender.

Police and MAPPA

Case managers shared information with the police following 21% (37 of 177) first examinations compared with 19% (35 of 180) of the non-polygraph case managers. In the case of polygraph offenders, the feedback form made clear that this information related to new disclosures made during the polygraph test, but it was unclear from the responses of the non-polygraph case managers how much of the information passed to the police was new material as opposed to general discussion about an offender.

Polygraph case managers reported sharing information with the police in relation to 35% (18 of 51) of maintenance and monitoring tests (i.e., those that deal with current behaviours). This was significantly more frequent than reported in the non-polygraph group, even taking into account the uncertain status of the information passed to the police by the non-polygraph case managers referred to above ($\chi^2 = 5.65$, $df=1$, $p < .05$; odds ratio 2.3, 95% C.I. 1.2 – 4.5), although due to the differences between the two groups discussed in Chapter 7 these findings are indicative.

Case Managers in both groups were also asked about new MAPPA referrals, but as the majority of offenders were already subject to Level 2 and Level 3 meetings few new referrals were made. Less than 5% of offenders in each group were breached and recalled to prison.

Retests

In relation to retests, for present purposes only second tests (that is, the first retest) are considered here. Table 20 shows that significant differences were found in respect to changes in supervision and treatment as reported by polygraph and non-polygraph case managers. The number of offenders involved is limited, however, which means that these findings must be viewed cautiously. In addition, due to differences between the two groups as discussed in Chapter 7 these findings are indicative.

As with first tests, it appeared that non-polygraph case managers tended to modify their risk assessments downwards, with 6 of their 8 risk changes being reductions in risk, in contrast with polygraph case managers where 5 of 8 changes were upwards, but the numbers are too small to demonstrate significance.

The effect of the second polygraph examination as reported by treatment facilitators is shown separately in Table 21 for the 27 tests for which they completed feedback, as comparable information from treatment facilitators of non-polygraphed offenders is not available. It can be seen that these tests continued to result in upward reassessments of risk and had an impact on management and treatment.

Table 20: Changes in risk assessment, supervision, treatment and the introduction of other interventions reported by polygraph case managers following a second polygraph examination, and by non-polygraph case managers at Time 2 (chi square; 95% confidence intervals); due to differences between the two groups as discussed in Chapter 7, these findings are indicative.

	Retest	Non-polygraph Time 2	Significance (chi square tests; 95% confidence intervals for odds ratios).
Increased risk	5 (12%)	2 (3%)	NS
Decreased risk	3 (7%)	6 (9%)	NS
Change in supervision	6 (14%)	1 (2%)	p < .05 (df=1) OR = 9.9 (CI = 1.5 – 64.8)
Change in treatment	9 (21%)	3 (5%)	p < .05 (df=1) OR = 5.2 (CI = 1.4 – 19.0)
Other interventions	4 (9%)	3 (5%)	NS
At least one effect *	16 (36%)	7 (11%)	p < .01 (df=1) OR = 5.0 (CI = 1.9 – 13.3)
No effect*	28 (64%)	57 (89%)	p < .01 (df=1) OR = 0.21 (CI = 0.08 - 0.58)
Total	44	64	

* defined as an increase in risk, change in supervision, change in treatment, or introduction of other intervention

Table 21: Changes in risk, treatment and other interventions as reported by treatment facilitators following a second polygraph examination.

	First Retest
Increased risk	6 (22%)
Decreased risk	1 (4%)
Change in treatment	3 (11%)
Other interventions	2 (7%)
At least one effect *	16 (59%)
No effect*	11 (41%)
Total	27

* defined as an increase in risk, change in supervision, change in treatment, or introduction of other intervention

Other effects

Apart from explicit changes in risk assessment, supervision arrangements, new interventions and treatment targets, case managers indicated in 46% (84 of 181) first tests and 53% (29 of 54) retests that the results of the polygraph examination had additional 'other effects' in relation to their management of offenders. Similarly, treatment facilitators reported 'other effects' following 42% (46 of 110) first tests and 35% (11 of 32) retests. Because this information related specifically to the outcome of a polygraph test, no similar question was asked of the non-polygraph case managers.

These 'other effects' are difficult to quantify as they relate to qualitative aspects of the work probation officers carry out with sex offenders. They include such things as improved co-operation with supervision, increased engagement in treatment, confirmation of risk assessment, or obtaining information that is passed on to child protection agencies or MAPPP meetings. Below are examples of specific comments made by probation officers in relation to these 'other effects':

- changed his attitude as he feels he is more likely to be believed (*this was following an NDI test*)
- the successful outcome of the polygraph tests has . . . given me the confidence to apply for early revocation [of his Order] . . . knowing that the polygraph test reinforces the assessment that he is low risk (*also after an NDI test*)
- enabled challenges to be made to his 'evangelical' assertions of 'total honesty'
- identified new areas that need to be addressed in supervision
- helped focus supervision on specific targets
- the polygraph test . . . substantiated his explanation of the offence, which has informed the one-to-one work currently being completed
- he is taking treatment more seriously since participating
- he was able to form a more honest relapse prevention plan
- more information for MAPPP
- enabled police and hostel staff to monitor behaviour more closely
- [the test provided] 'evidence' to support clinical assessment of risk previously discounted by some agencies

The case examples below provide a fuller illustration of some of these 'other effects':

Overcoming denial, and increasing motivation

A 72 year old man with no sex offending history was convicted of the indecent assault of an 11 year old girl whom he and his wife had informally 'adopted'. He was making little progress in treatment, being adept at avoiding attempts to address sexual matters. In the course of a 'deception indicated' test, he disclosed a large amount of information about his sexual history that was contrary to what he had

been reporting in the treatment group, which he subsequently spoke about in the group. According to the treatment facilitator, this was, “the first new information about himself for some months, and he was able to acknowledge his sexual fantasies and fully engage with the module”, adding, “A small thing, but definitely in his case proved some benefit in moving him on with the treatment programme”.

A 30 year old man with no previous offending history was convicted of the indecent assault of his 10 year old stepdaughter. He maintained that he had only ever touched her on one occasion, on her breasts and over her pyjamas, in spite of the girl’s statement indicating digital penetration of her vagina. After a ‘deception indicated’ polygraph test, he admitted to sexual arousal to 13 and 14 year old girls, and to having masturbated to fantasies of his stepdaughter prior to the offence. The treatment facilitator commented, “He is now able to view his offending behaviour in the context of sexual abuse as a result of disclosing inappropriate sexual fantasies towards his victim. Motivation has improved, denial and minimisation has decreased”.

Improved co-operation with supervision

A 60 year old man convicted of incest, who was on parole following a short prison sentence, disclosed during his polygraph maintenance examination that he had met up with another offender for coffee after the treatment group, contrary to group rules. This led directly to him telling his case manager that the other offender had borrowed £500 from him. According to the treatment manager, “The other offender has a life long record of offending against children and is very experienced at ‘grooming’. We felt that the other man was ‘grooming’ J and that perhaps borrowing the money was just a way of ensuring continued contact. J was warned of the consequences of continuing to associate with this man. The other man’s case manager was informed”. Although the assessment of risk had not changed, the treatment facilitator wrote, “assessment of risk would have changed if he had maintained regular contact with the other group member”. The treatment facilitator commented further, “The polygraph test was very useful in prompting J to disclose his contact with a much more sophisticated offender before it developed further. As a result more attention was given to the other man . . . with the result that he eventually had to be recalled. J is now working well on the Treatment Programme”.

A 45 year old man was on parole following a conviction for the indecent assault of his stepdaughter. On his sex history examination he disclosed a large amount of previously unknown sexual behaviour associated with the use of pornography and cross dressing. Subsequent to the test he has started to discuss this and his sexual fantasies more generally with his case manager, who wrote, “he has allowed himself to open up and discuss personal issues”.

Confirmation of risk assessment

A 50 year old man was on a CRO having been convicted of the indecent assault of his 13 year old stepdaughter. He had no previous convictions. During his first

polygraph examination he did not disclose any information of concern, and the outcome of this test was reported as 'No Deception Indicated'. The case manager wrote that this, "reassured me in my own assessment of risk. It reinforced the assessment I had made". Following a retest 6 months later that was again NDI, the case manager commented that he was confident in assessing this offender's risk as low, "based on OASys, my professional assessment and polygraph".

A 63 year old man was on a CRO having been convicted of internet offences but having no other history of offending. He had been a foster parent for many years during which no allegations of a sexual nature had been made against him. His account was a common one of accessing the internet in search of adult pornography, and he claimed to have been unsure of how the child images came to be on his computer. 'No Deception Indicated' was reported following a specific issue-offence type polygraph examination. Based on this, the lack of any other concerns about him, and his low Deviance score on psychometric testing it was decided that there was little to be gained in placing him on the treatment programme, and he was instead managed on a one to one basis by his case manager, which meant that a place on the group was freed up for another individual.

Contribution to treatment

A 56 year old man with no previous sex offending history was convicted of internet related offences. In a sex history examination he admitted to stealing knickers from his sister's house a few years previously, to sexual fantasies regarding schoolgirls, and to sitting in a cinema car park in order to watch young girls. Based on this and other information he disclosed about his sexual fantasies, the treatment facilitator identified "a new target around dealing with inappropriate fantasies and possible fantasy modification work", and noted that he would be required to repeat the 'victim empathy' block.

A 40 year old man committed an indecent assault on a child in a supermarket when intoxicated. He admitted the offence, but denied any memory of having pushed his groin into the girl's back as reported by her mother even though he accepted that this could have happened. Much time had been spent in the group trying to overcome his 'denial'. On a specific issue-offence type examination he was questioned about his lack of recall, and passed the test. The consistency of his account taken together with the test result led to a decision to accept his account of partial amnesia, allowing treatment to progress beyond this issue.

Passing on information

A 50 year old man with previous sex offence convictions against boys disclosed on his first polygraph examination that he had been going into public toilets on a regular basis, and though he denied offending he admitted to watching boys urinate. As he had offended in public toilets in the past, this information was passed to the police, who used it as part of their application for a Sex Offending

Prevention Order. He later breached this Order and received a prison sentence. On his release from prison the polygraph examiner was invited to a MAPPP meeting. It was agreed that he would be asked to undertake another polygraph examination, the questions for which were determined at the meeting. When tested on this second occasion the offender denied any further entry into public toilets, but admitted having contact with another sex offender about whom there were concerns. The outcome of the test was 'Deception Indicated', and at his own request he attended for another test soon after. On this occasion he admitted to chatting to boys in supermarket toilets, but again the test outcome was 'Deception Indicated'. This information was brought back to the MAPPP, and it was provisionally agreed to place him under surveillance.

A 24 year old man was on parole having been convicted of unlawful sexual intercourse with a 14 year old girl. His case manager believed he was maintaining a sexual relationship with his victim, but this was persistently denied by the offender. His monitoring included a tag and a curfew at night, with which he was compliant. He denied any wrongdoing during the pre-test interview, but the outcome of the test was 'Deception Indicated'. In the post-test interview he admitted to ongoing regular contact with his former victim, and to some sexual activity with her. The probation officer passed this information to the police and Social Services, and the offender was arrested. When interviewed by the police the girl reported spending one night a week in the offender's home, where in addition to the sexual activity described by him she said they also engaged in sexual intercourse. In his feedback the probation officer commented, "Hugely beneficial to my assessment and intervention. Developed a closer relationship with other agencies [in managing the case]".

A 39 year old man with a long history of sex offences against teenage girls had discussed with probation officers his strong sexual urges towards girls. He was living in a probation hostel, and was on a Sex Offender Prevention Order which prohibited him from having any contact with children. During his polygraph examination he told the examiner that the newspapers were delivered to the hostel by an 11 year old girl, and that in the morning he would meet her at the door to collect the paper. The probation officer passed this information on to the police, who considered prosecuting him for breaching his Order. The police also contacted the newsagents to inform them that it was probably not sensible for young girls to be delivering papers to the hostel, and this is now undertaken by an adult male.

Utility

On the Actions Taken forms distributed after each test, case managers and treatment facilitators were asked:

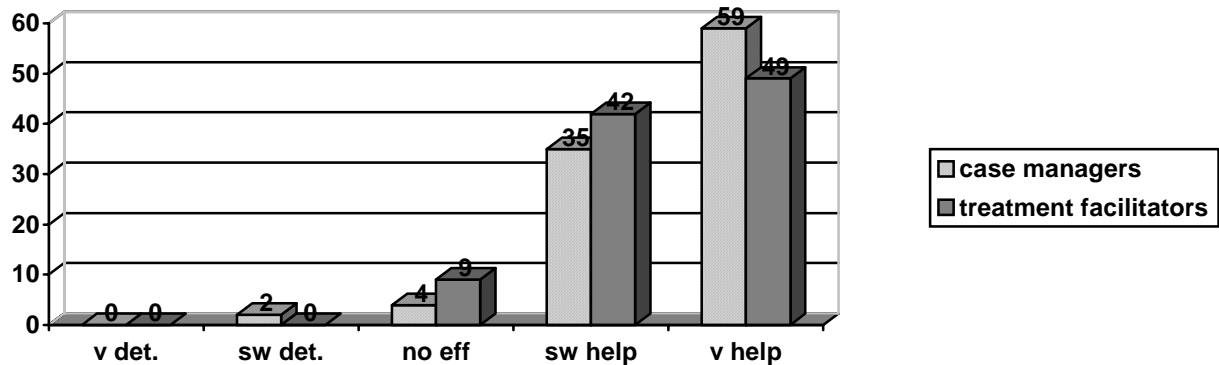
"How helpful was the polygraph in your dealings with the offender?".

Responses to this question were received from case managers in relation to 234 tests, and from treatment facilitators in relation to 145. Probation officers rated polygraphy as being Somewhat or Very Helpful in over 90% of cases (Table 22 and Figure 6).

Table 22: Global rating of probation officers of the usefulness of polygraph tests in their interactions with offenders.

	Case managers	Treatment Facilitators	Total
Very detrimental	0	0	0
Somewhat detrimental	4 (2%)	0	4 (1%)
No effect	10 (4%)	13 (9%)	23 (6%)
Somewhat helpful	81 (35%)	61 (42%)	142 (37%)
Very helpful	139 (59%)	71 (49%)	210 (55%)
Total	234	145	379

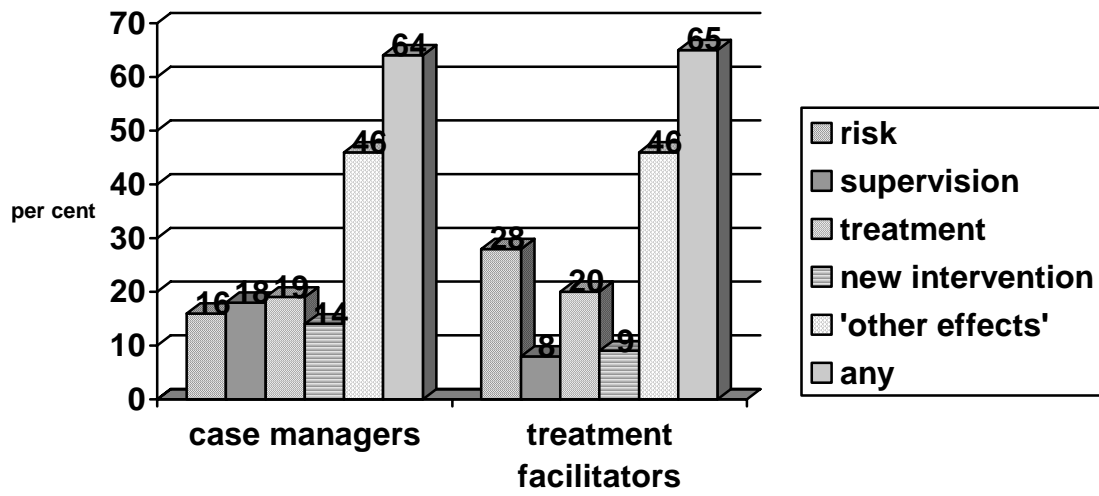
Figure 6: Global rating of probation officers of the usefulness of polygraph tests in their interactions with offenders.



Overall impact of polygraphy on supervision and treatment

Case managers and treatment facilitators reported at least one impact in terms of an increase in their assessment of risk, a supervisory change, a new intervention, the identification of a new treatment target, or an ‘other’ effect following nearly two thirds of the first tests for which they completed forms (Figure 5). Case managers also reported at least one of these impacts following 68% of retests, while treatment facilitators did so after 50% of retests.

Figure 5: Overall impact of polygraph examinations following a first polygraph examination as reported by probation officers.



7. DISCUSSION

Impact of polygraph testing

The results from this exploratory study are supportive of the potential value of polygraph testing within sex offender treatment programmes run by the probation service, although caution in their interpretation is necessary given the caveats referred to below. They suggest that polygraphy is associated with marked increases in the quantity, and an enhancement in the quality, of new disclosures made by offenders – the odds were 14 times greater that a polygraphed offender as opposed to a non-polygraphed one would make disclosures relevant to their treatment or supervision, indicative of a meaningful clinical impact. These disclosures are reported by probation officers to lead directly to changes in their risk assessment and supervision, and to the identification of new treatment targets, which occurred about 3 times more frequently than was the case in the non-polygraph offenders. In addition, case managers and treatment facilitators reported a number of qualitative ‘other effects’ associated with polygraph testing related to confirmation of their risk assessments, improved co-operation of offenders with supervision, increased engagement in treatment, and the obtaining of information relevant to child protection agencies or MAPPA meetings. Overall, testing was reported to have had an impact on offender management in about two thirds of cases for which the information was available, and was rated as somewhat or very helpful in over 90% of cases.

In the context of voluntary testing, it was found that less than half of sex offenders agreed to participate, with about a quarter of those tested subsequently refusing to have further tests. Polygraph is therefore not something with which all sex offenders will cooperate voluntarily, although at present one can only guess at the reasons underlying the reluctance of some to undergo testing. Given the impact polygraphy may have had on the assessment and management of those offenders who were tested, it might be argued that polygraph testing should be a mandatory part of sex offender treatment and supervision.

While the data and observations from the pilot are supportive of the value of polygraphy and mandatory testing, they are not conclusive. Three caveats in particular need to be noted in interpreting the findings reported here.

First, because an independent formal research evaluation was not commissioned as originally planned, data collection in relation to those who refused testing and offenders in the non-polygraph sample is incomplete, while feedback forms were not available for about 30% of polygraphed offenders and 40% of offenders from the comparison sites. Both these factors mean it is possible that the two groups may have differed in important ways which could have influenced the findings. In addition, follow-up interviews with case managers, treatment facilitators and offenders in both the polygraph and non-polygraph sites would have allowed a fuller examination of the value and impact of the test results.

Second, as this was not a randomised, controlled study, it is possible that the polygraph volunteers differed in some important manner from the non-

polygraph offenders. For example, differences in stage of treatment may be a biasing factor.

Third, as the pilot involved only volunteers, it is not known whether polygraphy would result in similar numbers of disclosures by offenders who did not have a choice about whether or not to undergo testing (although it should be borne in mind that useful information may also be obtained from ‘passed’ and ‘failed’ tests even when no disclosures are made). There were some differences between offenders who volunteered for testing and those who refused in terms of risk levels and sex offence history, and between them and sex offenders who were being managed by the probation service generally regarding risk levels and ethnicity, and it may be that some of these factors influenced the findings.

Further research designed to resolve these issues should not require large numbers of offenders. If large scale implementation of polygraphy is to be worth the cost, decisions will need to be made about the type and degree of impact sought from it, but we recommend that this should focus on the extent to which the test *influences* an offender’s treatment or supervision, either in the sense that it brings about some action by the probation officer, or it results in information directly relevant to treatment or supervision. In our opinion, unless effects of this type are at least 10 to 20% more frequent in polygraphed as opposed to non-polygraphed offenders the introduction of polygraphy would probably not be worth the effort (longer term one would be interested in its contribution to changes in reoffences rates, but it will be difficult to tease out the impact of polygraphy as opposed to other interventions). The findings from this exploratory study suggest that such an effect size is a reasonable expectation. Based on a conservative estimate of a 10% effect size, and starting from a baseline of change in 50% of non-polygraphed offenders, power calculations indicate that 300 to 400 offenders would need to be tested, and compared with a similar number of offenders in a comparison group.

Seriousness of disclosures and changes in risk assessment

New disclosures relevant to treatment and supervision were reported by probation officers in relation to 70% of polygraph tests. Although the majority of these disclosures were of ‘low’ seriousness, over a quarter of the disclosures made were rated by probation officers as being ‘medium’ in their level of severity (representing behaviours considered to be preludes to reoffending), and about 15% as ‘high’ (specific breaches or actual offending). In maintenance and monitoring tests, which are concerned with current behaviours, 25% of tests were associated with disclosures rated by case managers as being of medium severity, and 7% as high, indicative of problematic activity and an increase in the risk of reoffending. These disclosures were made even given the terms of the pilot which prevent examiners from questioning offenders specifically about new offences.

The quantity and the degree of seriousness of the disclosures made by polygraphed offenders may account for the finding that when probation officers of polygraphed offenders reassessed risk they typically judged it to have *increased*, but case managers of non-polygraph offenders more typically modified their risk assessments downwards; the odds ratio of nearly 3 to 1 is indicative of an important clinical effect

(although given the absence of randomisation, as well as incomplete feedback, it is possible that this finding was the result of genuine differences in risk rather than the effects of polygraph testing). It is not possible to determine from the information available whether the polygraph case managers were over-reacting to information that emerged from the polygraph examinations, or whether non-polygraph case managers were making their risk assessments and subsequent management decisions based on an incomplete knowledge base.

Regarding the disclosures themselves, we had no accounts of any offender denying what he is reported to have said in a polygraph examination, although probation officers were not asked about this directly; if it did occur videotape recording of the examination could have been checked. However, the ‘corroboration’ of polygraph disclosures was in most cases neither feasible nor an immediate issue – for example, reports regarding previous sexual history, ongoing fantasies, or index offence related issues do not lend themselves to verification in the normal sense, except by an offender repeating in treatment or supervision what he said in the polygraph examination. While we did not routinely receive information regarding follow-up, and so firm conclusions cannot be reached, in those cases where we did receive it the offender’s disclosure was confirmed in every instance – two illustrations of which are given below:

An offender disclosed in his polygraph examination a continuing sexual relationship with his previous 15 year old victim, but when his case manager passed the information on to the police he claimed to have fabricated his account because of the ‘pressure’ of the examination. When the girl was located and questioned by the police, she both confirmed and expanded on his admissions of sexual activity between them.

An offender made disclosures in his polygraph examination regarding his contact with a woman neighbour and her children. When the case manager spoke to this woman it transpired that he had been assisting them decorate their flat while in his underwear; as his index offence had occurred in similar circumstances, he was breached and recalled to prison.

Is the polygraph the cause of the disclosures?

Some might question whether it was polygraph testing that was responsible for the substantial increase in disclosures made by offenders, and that simply engaging them in a similar interview might achieve the same results. Although the current study cannot resolve this issue, the suggestion seems unlikely for a number of reasons. First, while offenders were in different stages of treatment when they received their first polygraph test, many had been involved in treatment programmes for months without disclosing the information they did in the polygraph examination (illustrated by the case example on bottom of page 43). Second, so-called ‘bogus pipeline’ studies have shown that if individuals believe they are attached to a ‘lie detector’, even if they are not, they will be more honest in their disclosures, hence, a ‘bogus pipeline to the truth’ (Quigley-Fernandez & Tedeschi, 1978); this works, of course, only so long as the individual continues to believe in the validity of the ‘lie detector’, which in practice means that it must be genuine if the effect is to be maintained. And third, it

was found in the pilot that a third of disclosures were made in the post-test interview *after* the offender had been tested, with 20% occurring in tests where no pre-test disclosure had taken place at all, which would be difficult to explain if the test was not related to these further disclosures.

Why does the polygraph appear to facilitate disclosure in sex offenders? In addition to the 'bogus pipeline' influence referred to above, and the systematic manner in which the tests are conducted, it may be that polygraph examination provides offenders with an 'excuse' to reveal information that they had previously denied or kept hidden, and which they would have found difficult to disclose in ordinary circumstances.

For some offenders, there may also have been a perceived need to explain 'Deception Indicated' outcomes, raising the possibility of false disclosures in these cases. There was nothing to indicate that this was a major factor in the current pilot, but this was not something that could be tested. Two studies which asked polygraphed sex offenders on anonymous questionnaire whether they had ever made up disclosures during a polygraph examination found that 5 to 10% admitted to doing so on at least one occasion (Grubin & Madsen, 2006; Kokish et al, 2005), suggesting that this is a real but probably not a major problem in sex offender testing. Nonetheless, the possibility of false admissions does need to be taken into account if consequences are to be associated with disclosures made to polygraph examiners.

The use of polygraph results

Although a number of briefings for probation officers and other staff took place in each of the areas prior to the start of the pilot, examiners reported that probation officers were often uncertain about what to ask for in a test, or how to make use of the information they provided following a test, until they had some experience of their offenders having been tested. It was also their impression that it was those case managers with greater numbers of sex offenders on their case loads who made best use of test results, suggesting that their greater level of familiarity in the management of sex offenders may make them more able to evaluate and make use of the new information coming from the polygraph examinations.

Similarly, it was the impression of the examiners that those probation officers who viewed polygraph testing as part of an overall package of measures, integrating test results with their treatment and management more generally, that made best use of the procedure. This was supported by the data collected on the feedback forms, where reports of test findings being fed into supervision arrangements, treatment plans and where appropriate MAPPA were also associated with particularly positive comments. On the other hand, if testing exists in isolation, it is less likely that offenders themselves will engage with it, diminishing its efficacy. Thus, when considering future evaluations, consideration should be given not only to measuring changes that occur in supervision and treatment as a result of polygraph testing, but also changes that ought to have occurred judged by an independent review of the case.

In this context it is worth emphasising the need to differentiate between disclosures that are made during a test, and test outcome itself. Disclosures that are followed by a

'Deception Indicated' or 'Inconclusive' finding, or that are made during a post-test interview, must be considered as partial disclosures only. Nevertheless, they provide an opening for further work that can be carried out with offenders by probation officers. While test outcome gives an indication of how confident one can be with the offender's account, it is the disclosures themselves that are most pertinent in post-conviction settings such as sex offender testing. Viewed in this way, polygraphy is perhaps better seen as a truth facilitator rather than a lie detector, with the information obtained from it used to improve offenders' engagement in treatment and supervision, rather than to catch them out telling lies.

Utility

Feedback was obtained from probation officers in respect to 327 tests, representing two thirds of polygraph examinations that were carried out during the pilot. In over 90% of these cases the probation officers reported that polygraphy was either somewhat or very helpful in their work with individual offenders. Indeed, it was difficult to find any who were critical of the intervention, although a small number referred to offenders being upset or distressed by the test (typically after having disclosed previously unknown information); in a future study it might be worth enquiring specifically about negative consequences such as this, asking also about the 'stress' associated with psychometric testing and with sex offender treatment generally. It may be, of course, that it was only probation officers who were supportive of polygraphy who responded, but one might have expected at least some of those with a complaint to have made their views known.

Conclusion

This pilot has demonstrated that the use of polygraph testing in probation sex offender treatment programmes is feasible, practical, and can be implemented effectively. It provides evidence to show that polygraphy can make an important contribution to the treatment and supervision of sex offenders on probation, potentially making a valuable contribution to public safety by enabling probation officers to better monitor risk and to bring about more effective and timely interventions. As an exploratory study, however, the findings should be seen as indicative rather than definitive. Further research in the context of a randomised controlled trial is needed to confirm the results reported here, and to explore their implications in more depth. Apart from any policy considerations, this would best be carried as part of a mandatory testing programme in order to remove the self selection bias inherent in voluntary testing.

8. RECOMMENDATIONS

The results of the pilot support the need for further research to evaluate the impact of polygraph testing in the supervision and treatment of sex offenders. A more definitive, randomised controlled trial, ideally within the context of mandatory testing, would be an appropriate next step. Such a trial, however, would bring with it a number of challenges, which are addressed in the following recommendations, which are based on the experiences of this pilot study:

1. If testing were made mandatory for only some offenders, for example just those on parole, then thought would need to be given about how to manage treatment groups where some offenders have to undergo polygraph testing, and others can choose whether or not to do so.
2. There will need to be a clear policy on how to respond if an offender either defaults from his polygraph examination, or does not cooperate with test procedures.
3. Systems will need to be developed to ensure that case managers remain up to date with the nature of polygraph provision and their role in it, taking into account a high turnover rate of case managers. This should include support in relation to understanding and applying the results of polygraph examinations.
4. There needs to be clear advice for probation officers about how to deal with 'Deception Indicated' outcomes in the absence of disclosures. In this respect, in line with a polygraph error rate of 10 to 20%, such DI results are probably best treated as a positive screen, indicative of the need for further enquiry or investigation, rather than as 'evidence' of deception.
5. There will be situations where offenders will need to be asked about previous offences or current reoffending. Clarification is required as to the extent to which such questions are permitted, taking into account both the maintenance of offender cooperation and remaining within legal boundaries regarding self incrimination.
6. Consideration will need to be given to the development of protocols enabling polygraphy to be incorporated into treatment programmes.
7. Good administrative arrangements will need to be established, taking into account differences in policy and practice between probation areas.
8. A competency framework for examiner selection and 'certification' needs to be established.
9. Provision must be made to ensure adequate training and supervision of examiners. In the long term, monitoring and supervision can probably best be addressed through formal regulation, but attention could usefully be given to this now.

10. Consideration must be given as to how such a trial is best evaluated. The main challenge will be in ensuring adequate data collection. Large numbers of offenders are unlikely to be required; based on a conservative estimate of a 10% effect size, and starting from a baseline of change in 50% of non-polygraphed offenders, power calculations indicate that 300 to 400 offenders 300 to 400 offenders in a polygraph group, compared with a similarly sized control group, should be adequate. In addition, in the current pilot it is arguable that in some cases risk assessment and supervision should have changed more in response to polygraph outcomes than they in fact did, and the evaluation must therefore be capable not only of measuring changes that occur in supervision and treatment, but also changes that ought to have occurred.

9. REFERENCES

- ASTA (2000). American Society for Testing and Materials. *Standard Guide for the Conduct of PDD Screening Examinations*. WWW.ASTM.org
- British Psychological Society (1986). Report of the working group on the use of the polygraph in criminal investigation and personnel screening. *Bulletin of the British Psychological Society*, 39, 81-94.
- Cross, T.P., Saxe, L. (2001). Polygraph testing and sexual abuse: The lure of the magic lasso. *Child Maltreatment*, 6, 195-206.
- English, K., Jones, L., Pasini-Hill, D., Patrick, D., Cooley-Towell, S. (2000a). The value of polygraph testing in sex offender management (Research Report Submitted to the National Institute of Justice D97LBVX0034). Colorado Department of Public Safety, Division of Criminal Justice, Office of Research and Statistics.
- English, K., Jones, L., Patrick, D., Pasini-Hill, D., Gonzalez, S. (2000b). We need you to become experts in the post-conviction polygraph. *Polygraph*, 29, 44-62.
- Fiedler, K., Schmid, J., Stahl, T. (2002). What is the current truth about polygraph lie detection? *Basic and Applied Social Psychology*, 24, 313-324.
- Grubin, D., L. Madsen, S. Parsons, D. Sosnowski, B. Warberg (2004). A prospective study of the impact of polygraphy on high risk behaviours in adult sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16, 209-222.
- Grubin, D., Madsen, L. (2005). Lie Detection and the Polygraph: A historical review. *British Journal of Forensic Psychiatry and Psychology*, 16, 357-369.
- Grubin, D., Madsen, L. (2006). Accuracy and utility of post conviction polygraph testing of sex offenders. *British Journal of Psychiatry*, 188, 479-483.
- Holden, E.J. (2000). Pre- and post-conviction polygraph: Building blocks for the future – procedures, principles and practices. *Polygraph*, 29:69-115.
- Iacono, W.G., Patrick, C.J. (2005). Polygraph (“Lie Detector”) Testing: Current Status and Emerging Trends. In, *The Handbook of Forensic Psychology*, 3rd Edition, I. B. Weiner and A. K. Hess, Eds. Hoboken, NJ: John Wiley and Sons.
- Kokish, R., Levenson, J.S., Blasingame, G.D. (2005). Post-conviction Sex Offender Polygraph Examination: Client-Reported Perceptions of Utility and Accuracy. *Sexual Abuse: A Journal of Research and Treatment*, 17, 211-221.
- Krapohl, D.J. (2005). Polygraph decision rules for evidentiary and paired-testing (Marin Protocol) applications. *Polygraph*, 34:184-192.
- Krapohl, D.J. (in press). Validated polygraph techniques. *Polygraph*.

Landis, R. J., Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

Madsen, L., Parsons, S., Grubin, D. (2004). A preliminary study of the contribution of periodic polygraph testing to the treatment and supervision of sex offenders. *British Journal of Forensic Psychiatry and Psychology*, 15, 682-695.

Medical Research Council (2000). http://www.mrc.ac.uk/pdf-mrc_cpr.pdf.

National Research Council (2002). *The Polygraph and Lie Detection*. Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Quigley-Fernandez, B. & Tedeschi, J. T. (1978). The bogus pipeline as lie detector: Two validity studies. *Journal of Personality and Social Psychology*, 36:247-256.

United States v. Antelope, 395 F.3d 1128 (2005).