The accuracy and utility of post conviction polygraph testing with sex offenders

Don Grubin
Lars Madsen
**Background:** Polygraphy is used increasingly in the treatment and supervision of sex offenders, but little research has addressed its accuracy in this setting, nor linked accuracy with utility.

**Aims:** To investigate the utility and accuracy of polygraphy in post conviction testing of community-based sex offenders.

**Method:** A self-report measure examined the experiences of offenders with polygraphy.

**Results:** Based on self report, the polygraph’s accuracy was approximately 85%. False negatives and false positives were not associated with demographic characteristics, personality variables or intelligence. The majority of offenders found the polygraph to be helpful in both treatment and supervision. Nine per cent of offenders claimed to have made false disclosures; these individuals had higher scores on ratings of Neuroticism and lower scores on ratings of Conscientiousness.

**Conclusions:** These results support the view that the polygraph is both accurate and useful in the treatment and supervision of sex offenders.

**Declaration of Interests:** None.
INTRODUCTION

The polygraph (often referred to as a ‘lie detector’) has been proposed as a useful tool in the treatment and supervision of sex offenders (Blasingame, 1998; English et al, 2000; Grubin et al, 2004). Proponents argue that it provides clinicians with more reliable sexual histories, more complete and accurate offence descriptions, and a greater likelihood of identifying high risk behaviours or breaches in probation conditions, enabling intervention to take place before reoffending occurs. Many American states now require sex offenders to undergo regular polygraphy examinations as a condition of probation or parole, or in order to be included on treatment programmes; similar measures are being considered in England. While research conducted in so-called post conviction settings is supportive, the focus has been on the utility of testing with a tendency to dismiss concerns regarding accuracy. However, if polygraphy is not particularly accurate then utility will be compromised as subjects come to believe that the polygraph does not ‘work’. In the study reported here, offenders’ self report is used to assess the accuracy and utility of post conviction polygraph testing.

METHOD

Participants

Three hundred and twenty-one sex offenders participating in community-based treatment programmes in the American state of Georgia were approached, of whom 176 (55%), including 3 females, agreed to take part. Age ranged from 18 to 82 years
(M = 40, SD = 12.6). One hundred and forty-four individuals were Caucasian (81%), 28 were African-American (16%), one was Hispanic, one was American-Indian, and two were from other backgrounds. With regards to education, 31 (18%) reported that their highest educational attainment was junior high school, 49 (28%) high school, 49 (28%) a technical or business course, 37 (21%) university, and 9 (5%) a post-graduate level of education; one subject did not report his educational status. Sixty-nine offenders (40%) reported having had previous psychological or psychiatric treatment.

One hundred and fifty (87%) of the offenders had been convicted of contact sexual offences, of whom 137 (80%) had offended against child victims, 12 (7%) against adult victims and 1 against both. Sixteen (9%) participants were convicted of non-contact sexual offences, including the manufacture or possession of child pornography, indecent exposure and voyeurism. Of the other 10 offenders, 8 (4%) participants had been charged with a sexual offence and were awaiting trial, and 2 (1%) had not been charged or convicted of any sexual offences (one had self-referred for ‘sexual addiction’ problems and the other had been referred due to professional misconduct issues. The mean length of time in sex offender treatment was 23.5 months (SD = 23), with a range of 1 to 120 months.

**Risk**

Static-99 (Hanson & Thornton, 2000) is a widely used actuarial risk instrument designed to estimate the probability of sexual and violent recidivism for adult males who have already been convicted of at least one sexual offence against a child or adult. The measure contains ten items that address static (historical) factors. Inter-rater reliability is reported to fall between .87 and .96 (Hanson & Thornton, 2000).
present study 10 cases (20%) were scored by two raters blind to each other’s results, with perfect agreement between them.

**Personality**

The Revised NEO-Personality Inventory (NEO-PI-R) is a self-report questionnaire developed by Costa and McCrae (1992) to assess normal personality dimensions based on a five-factor model. The five domains are: Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). The NEO-PI-R has been shown to have good reliability and validity (Trull, 1992; Clarkin et al., 1993; Wilberg et al., 1999). Internal consistencies for the facets range from .56 to .81 and .86 to .92 for the five broader domains (Costa, McCrae & Dye, 1991). Cronbach alpha coefficients in this sample for the five domains ranged from .6 (Openness) to .91 (Neuroticism).

**Intelligence**

The National Adult Reading Test (NART-2) was used to provide an estimate of intellectual ability (Nelson & Willison, 1991).

**Previous experiences of the polygraph**

A 12-item survey, the Previous Experiences of the Polygraph Questionnaire (PEPQ), was developed for the study to gather descriptive information about participants’ previous experiences and perceptions of the polygraph. The questionnaire is divided into three broad sections. *Section One* addresses false-positive and false negative rates, false admissions, and the use of countermeasures, *Section Two* the extent to which the participant considered the polygraph to be helpful in assisting him to avoid
risk behaviours and re-offending and to engage in treatment, and Section Three the participant’s perception of polygraph accuracy overall. The PEPQ took approximately 10 minutes to complete.

Procedure

All subjects were taking part in treatment programmes in which polygraphy was a requirement for participation, with individuals required to complete a minimum of two polygraph tests per year. Subjects were approached while attending their regular treatment groups. They were informed that the broad purpose of the research was to investigate the value of the polygraph in a post-conviction context. They were assured of confidentiality, and all gave their signed informed consent.

Participants were seen on a single occasion for up to 60 minutes. During this time they completed the NEO and the PEPQ, either by themselves or with other participants in a quite room at their treatment facilities. Subjects were then interviewed individually about their present circumstances and previous experiences of the polygraph; the NART-2 was administered at this time.

RESULTS

Characteristics of the sample

Based on Static-99 ratings, 93 (58%) subjects were assessed as being a low risk of a sexual or violent re-offence, 46 (29%) medium-low risk, 19 (12%) medium-high risk, and 3 (2%) high risk.
Thirteen participants did not complete the NART for reasons such as illiteracy or not having appropriate eyeglasses. For the remaining sample the mean IQ was 102 (SD = 11.9) with a range of 75 to 128.

Valid NEO profiles were obtained for 152 subjects (86%). Overall, scores were in the high range for Neuroticism (M=87, SD = 21), the average range for Extraversion (M = 101, SD = 16) and Agreeableness (M =120, SD = 15), and in the low range for Openness (M = 100, SD = 15) and Conscientiousness (M = 114, SD = 17).

**Self-reported accuracy**

One hundred and seventy four offenders provided information about previous polygraph tests. Of these 126 (73%) reported completing a total of 263 polygraph tests whilst on probation; the other remaining 48 individuals (27%) had not yet taken their first polygraph examinations, but were scheduled to do so.

The subjects reported that in 225 (86%) of their completed tests they had told the truth, and they reported being deceptive in 38 (14%) of the tests; according to them, the polygraph outcome on these tests was ‘no-deception indicated’ in 197 (75%), and ‘deception indicated’ in 66 (25%) (Table I), giving a false positive rate of 15%, a false negative rate of 16%, and an overall accuracy of 85%.

Thus, based on self-report the specificity of the tests (correctly detecting truthfulness) was 85%, while the sensitivity (correctly detecting deception) was 84%, with a false negative rate of 16%.
It can also be seen from Table I that in the 197 tests in which offenders reported the outcome as being ‘no deception indicated’, they said this was correct in 191 (97%) of cases (the positive predictive accuracy). However, in the 66 of tests that were reported to be ‘deception indicated’, they said this was correct in only 32 (48%) of cases (the negative predicative accuracy).

When the 126 individual subjects who had taken polygraph tests are considered rather than the number of tests they reported completing, 27 (21%) reported that they had been wrongly reported as being deceptive (false positive) when they had been telling the truth on at least one occasion, and 6 (5%) to have been wrongly reported as being truthful (false negative) when they had in fact been lying. There was no overlap been these individuals.

**False positives**

Individuals who reported telling the truth but were wrongly labelled as deceptive (false positives; n = 27) were compared with those who said they had be correctly classified as telling the truth (true negatives; n = 64), as well as with those who reported being correctly detected as being deceptive (true positives; n=29). Relevant variables were grouped into two broad categories: **historical** (age, ethnic origin, previous psychological and psychiatric history, education attainment, number of previous polygraph tests, and risk) and **psychological** (personality, intelligence). Univariate
analyses of these variables did not yield any significant differences between the groups in respect of any of these variables.

**False negatives**

Individuals who claimed that they had ever been deceptive but were classified as ‘no deception indicated’ (false negatives, n = 6) were compared with those who reported being deceptive and were accurately labelled as such (true positives, n = 29), and with those who said they had been correctly labelled as non-deceptive (true negative; n = 64). Univariate analyses did not yield any significant results.

**Utility**

One hundred fourteen of the 126 offenders who had been polygraph tested fully completed the PEPQ. Of these, 50 (44%) reported that they were more truthful with their probation officers and treatment providers than they otherwise would have been because of their experience of the polygraph; 39 subjects (34%) reported that it assisted them in being more truthful about their behaviours to families and friends. These findings are similar to those in the 45 subjects who had not yet been tested and fully completed the PEPQ, with 20 (44%) and 16 (36%) indicating that the expectation of a polygraph test increased their disclosures to probation officers and family and friends respectively.

Regarding behaviours associated with offending, 71 individuals (56%) who had previously been polygraphed reported that the polygraph was moderately to extremely assisting them to avoid reoffending, 81 (64%) that it was useful in assisting them to avoid engaging in risk behaviours, and 84 (67%) that it was generally helpful in
respect of treatment. Similar responses were given by those awaiting their first examinations. Responses for the whole sample in respect of the impact of polygraph on engaging in risk behaviours, actual reoffending, and engagement in treatment are shown in Figure 1.

FIGURE 1 HERE

Considering the group as a whole in regards to specific risk behaviours, 57 (33%) individuals reported they were less likely to masturbate to deviant (offence-related) fantasies, 53 (31%) that they were less likely to have contact with children and/or potential victims because of the polygraph, 47 (27%) that the polygraph contributed to less drug and alcohol use, and 44 (26%) that they were less likely to use or buy pornography. However, a significantly greater proportion of those who had undergone polygraph testing than those still awaiting their first test reported that they were less likely to visit places to view children (37 versus 5, \( \chi^2 = 5.9, df = 1, p = .01 \)) and to engage in other more general risk behaviours (18 versus 1, \( \chi^2 = 4.2, df = 1, p = .04 \)).

There was no difference in the perception of the polygraph’s accuracy between the subjects who had previously been polygraphed and the ones who had not. Overall, 16 subjects (10%) considered it to be no more accurate than chance, 15 (9%) slightly accurate, 73 (44%) ‘moderately’ accurate and 63 (38%) rated it as being ‘quite’ to ‘extremely’ accurate.
Sanctions

Twenty-seven (22%) out of 121 men who had completed a post-conviction polygraph exam reported experiencing a direct sanction due to a polygraph result or a disclosure made during a test, the most common of which involved having to address additional issues in treatment or supervision (78%). Four of these individuals (15%) claimed that there was a change to their supervision conditions: two that their treatment was terminated and two a change in living circumstances (e.g. reduced contact with family). There was no relationship between having experienced a sanction and claiming to have been a false positive ($\chi^2 = 3.07; \text{df} = 1; p = .08$).

To test whether having been sanctioned or erroneously classified (false-positive or negative) affected the participant’s perception of its utility, an overall ‘helpfulness’ variable was created by combining the scores for the three utility scales. There was, however, no difference in perceptions of utility between those participants were sanctioned and those who were not ($t (111) = .38, p = .7$), nor was there a difference between those who reported being false-positives and true negatives, or between the false-negatives and true positives.

Countermeasures and false admissions

Only two participants (1%) claimed to have used drugs to beat the polygraph. Both also claimed to have previously been deceptive without being detected.

Twelve participants (9%) reported making false admissions regarding their behaviour at some stage during a post-conviction polygraph test, only 5 of whom claimed to have been wrongly labelled as being deceptive. The main reasons given for making a false
admission were in 3 cases (25%) the fear of getting in trouble with their probation officers, and in another three cases (25%) feeling pressured by the polygraph examiner. In the remaining cases one man said he wanted to ‘make a good impression’, another was ‘confused’, one wanted to ensure that he passed the polygraph test and another wanted to demonstrate commitment to therapy. Two participants did not provide a reason.

A significant difference was found when a one-way between groups MANOVA was performed with the five NEO domain scores as the dependent variables and ‘having made a false admission’ as the independent variable (F (5, 96) = 2.46, p < .01). When results for the dependent variables were considered separately, two reached statistical significance using a Bonferoni adjusted alpha level of .01: Neuroticism [F (1,102) = 10.08, p < .01] and Conscientiousness [F (1, 102) = 7.85, p < .01], with the false confessors having higher levels of Neuroticism (104 compared to 84) and lower levels of Conscientiousness (101 compared with 116).

**DISCUSSION**

This study explores the experiences of community based sex offenders required to undergo regular post-conviction polygraph examinations. Broadly-speaking we found that the majority of sex offenders reported that the polygraph is helpful both in terms of treatment and in avoiding risk behaviours and re-offences. These findings, however, are based on the responses of the 55% of programme participants who
agreed to take part in the study, and it is possible that the other 45% may have had very different views on the value of polygraphy.

Our results are consistent with other research that has examined the utility of post-conviction polygraph testing. For example, a number of studies have found that sex offenders undergoing polygraph examinations report larger numbers of previously unknown offences and victims, younger ages at which they began to engage in deviant sexual behaviour, and a reduction in claims of their own sexual victimisation (Ahlmeyer et al, 2000; Emerick & Dutton, 1993; Hindman & Peters, 2001). Other studies have described increased reporting of high risk behaviours, and a consequent decrease in those behaviours as well as a reduction in reoffending (Edson, 1991; Harrison & Kirkpatrick, 2000; Grubin et al, 2004; Madsen et al, 2004).

In terms of the utility of post conviction polygraph testing, it has been argued that increased disclosure by offenders enables improved identification of treatment targets, encourages engagement by helping to overcome denial, and assists offenders in adhering to relapse prevention plans (Blasingame, 1998; English et al, 2000; Holden, 2000; Grubin et al, 2004; Madsen et al, 2004). Our results are supportive of these conclusions, and indicate that polygraphy can have a therapeutic role as well as the more usually perceived one of ‘detecting lies’. Indeed, confirmation that an individual is being honest in treatment and supervision, particularly in contexts where risk is a real issue, can be a critical element in the treatment process.
Accuracy

While the emphasis on utility in post conviction settings is understandable, polygraph accuracy can not be ignored. If subjects do not believe that polygraphy ‘works’ they will be less likely to disclose relevant information during a test, while a knowledge of accuracy rates is required to make sense of test results in cases where there is an indication of deception in the absence of disclosure. Thus, those tested as well as those who rely on test results must have confidence in the validity of the technique if it is to be viable clinically.

In terms of polygraph accuracy in general, the literature contains conflicting accounts, with many studies criticised for their methodological weaknesses (Furedy, 1996; Lykken, 1998; Cross & Saxe, 2001). However, a recent exhaustive review carried out by an expert panel appointed by the National Academies of Science concluded that the best estimate of polygraph accuracy falls between 81 and 91% (National Research Council, 2002). But because neither of the methodologies typically used to examine accuracy – laboratory studies that involve simulated exercises, and field studies where a retrospective assessment of polygraph outcome is carried out in real-life situations such as criminal investigations where ‘ground truth’ can be established – lend themselves easily to post conviction testing, none of the research reviewed in the National Academies report examined the accuracy of polygraphy when used in post-conviction or therapeutic contexts.

We are aware of only one unpublished study that has attempted to investigate the accuracy of polygraphy in a post-conviction setting (Kokish, personal communication). In this research sex offenders taking part in a treatment programme
in California and assured of anonymity were asked about the accuracy of the polygraph tests they had completed. According to the offenders’ self-report, they had been deceptive on 43 (15%) of 386 tests, with the polygraph correctly identifying truth-telling 93% of the time and deception in 74% of cases, producing an overall accuracy rate of 91%.

In our study, we made use of a method similar to that employed by Kokish and his colleagues. Our results indicating an accuracy rate of 85% in detecting truth-telling and 84% in detecting deception is similar to that found in the California offenders. Although this approach depends on the uncorroborated self reports of participants with no means of comparing their accounts with actual test outcomes, it is of note that the reported accuracy rates in both samples is consistent with the National Academies of Science estimate of polygraph accuracy in the range of 81 – 91%. The offenders themselves also perceived the accuracy of the polygraph to fall within this range, with the majority rating it as ‘moderately’ to ‘highly’ accurate.

Although overall accuracy appears good, interpreting this in respect of specific test outcomes is not straightforward. While the positive predicative rate (the likelihood that the subject is telling the truth when the examiner concludes ‘no deception indicated’) of 97% was very high, the negative predictive rate (the likelihood that the subject is lying when the examiner concludes ‘deception indicated’) of 48% is much less good. This outcome may partly reflect self-presentation biases (deceptive offenders for instance may be more likely to claim that the polygraph got it wrong when caught out and less likely to disclose having ‘beaten it’), but more relevant is the
relatively low base rate of reported deception in the sample, with deception admitted in only 38 of 263 tests (14%).

The importance the base rate of deception in the group of subjects being tested was highlighted by the National Academies of Science review (National Research Council, 2002). It is one of the primary reasons the review did not support the use of polygraphy in security contexts, where the base rate of deception is likely to be very low (one hopes that there are very few spies in federal agencies); the review suggested that polygraphy only becomes viable when the base rate of deception exceeds 10%. This is because where base rates of deception are low, even a highly accurate test will produce more false than true positives. For example, if 1000 employees are screened, and there is just one spy amongst them, even an instrument that is 99% accurate will wrongly identify 10 individuals as being deceptive (1% of 1000). It is only when the underlying rate of deception equals 50% that the positive and negative predicative values equal the overall accuracy of the test. Those arguing for its use in the intelligence community, however, point out that while the negative predictive rate may be low in this setting, polygraphy nonetheless serves as a useful screening tool to identify a small group of individuals who need to be looked at more closely.

Without going into the pros and cons of the various arguments regarding polygraphy in security settings, it should be recognised that the issues are very different in post-conviction contexts:

- the base rate of deception is likely to be well in excess of the 10% minimum recommended by the National Academies Review (although this does give rise
to a somewhat ironic situation in that the more honest examinees become, the less confident one can be with a ‘deception indicated’ result;

- in post conviction testing the emphasis is less on ‘passing’ or ‘failing’ the polygraph, and more on the facilitation of disclosures relevant to supervision and treatment;

- getting it ‘wrong’ in a post conviction test is of significantly less consequence then getting a wrong result in a criminal investigation or on a security screen, where much more reliance may be placed on the examination. In post conviction settings, polygraph is only one of a number of tools being used, and while a ‘failed’ test without disclosure may increase concerns and lead to more attention being paid to the individual, no punitive actions would normally be taken solely on the basis of a ‘failed’ polygraph examination.

It should also be noted that even based on the data provided by the offenders in our study, the 48% likelihood of an individual reported as being deceptive is not telling the truth does not mean that polygraph outcome is no better than chance. This figure relates to only a small proportion of the sample; the overall accuracy for the entire sample was 85%. An individual in the ‘deceptive’ group is much more likely to be deceptive than one in the ‘non-deception’ group, where the rate is 3%.

**False positives, false negatives, disclosures and false disclosures**

None of the variables we tested distinguished offenders more likely to be false positives or false negatives. Waid *et al* (1979) suggested that socialisation may be
associated with false-negative errors. Although socialisation has been related to the Neuroticism and Conscientiousness domains of the NEO (e.g. Costa & McCrae, 1992), neither of these characteristics distinguished false negatives from true negatives or true positives in our study.

As indicated above, much of the value of post conviction testing relates to the facilitation of disclosures made by subjects relevant to treatment and supervision. In a criminal milieu Gudjonsson (2001) reviewed the reasons suspects ‘confess’ to crimes, and concluded that the three most important factors are the strength of evidence against the individual, external pressure (e.g. fear of custody), and internal pressure (e.g. wanting to ‘clear their consciences’). While more than one factor may operate, it is the first – the perception of evidence against the individual – that seems to be the most powerful motivator to confess. In the context of post-conviction polygraph testing, it may be the case that a polygraph examination enhances the subject’s perception of the ‘evidence’ being against him, prompting disclosure. In support of this are studies which show that individuals are more likely to disclose undesirable attitudes if they believe they are being tested with an instrument that can detect deception (even if it cannot), presumably to avoid being ‘caught out’ (Jones & Sigall, 1971; Quigley-Fernandez & Tedeschi, 1978).

The converse to this, however, is that individuals may feel pressured to make untrue admissions in a polygraph examination. Nine per cent of the offenders in our study claimed to have done so, suggesting that while the incidence of this is not high, it is of relevance. We found that high Neuroticism and low Conscientiousness scores characterised those who reported making false admissions; ‘high Neuroticism’ is
associated with pervasive feelings of guilt, fear, disgust, anger and embarrassment as well as high impulsivity, and ‘low Conscientiousness’ with being less scrupulous and reliable. This suggests that “false disclosers” are possibly more emotionally disturbed in general, and more impulsive. In difficult interview situations, it may be that these individuals cope by ‘confessing’. Examining the reasons the 12 participants in our study gave for making false disclosures provides some support for this, with 6 of the 12 citing either a fear of getting into trouble with their probation officers or feeling pressured by the polygraph examiner. Interestingly, none of the 12 said they had received a sanction because of a polygraph test, indicating perhaps that the ‘false’ disclosures were either not particularly significant, or were viewed with scepticism. Regardless, it appears that some individuals may be more prone to making false disclosures, and polygraph examiners as well as those providing treatment and supervision need to be cautious when interpreting polygraph examination results.

In summary, our findings support the view that post conviction polygraph testing is a useful adjunct to the treatment and supervision of sex offenders in the community. Accuracy rates as reported by offenders who have undergone polygraph examination appear to of a sufficiently high level to maintain the utility value of the tests.

ACKNOWLEDGEMENTS

The research upon which this publication is based was supported by funding from the NHS National Programme on Forensic Mental Health R&D. However the views expressed in this publication are those of the authors and not of the Programme or the Department Health.
REFERENCES


Table I.

*Self-reported accuracy rates for post-conviction polygraph examinations.*

<table>
<thead>
<tr>
<th>Polygraph Test Result</th>
<th>Deceptive</th>
<th>Truthful</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deception Indicated</td>
<td>32</td>
<td>34</td>
<td>66</td>
</tr>
<tr>
<td>No Deceptive Indicated</td>
<td>6</td>
<td>191</td>
<td>197</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>225</td>
<td>263</td>
</tr>
</tbody>
</table>
Figure 1

Helpfulness of polygraphy with avoiding risk behaviours and re-offences and with overall treatment (%) (n = 169).
Clinical Implications

- Post conviction polygraph testing is a useful adjunct to the treatment and supervision of sex offenders in the community.

- The accuracy of polygraph testing as reported by offenders is similar to that found in studies carried out in other settings.

- A small proportion of individuals may make false disclosures.

Limitations

- The findings are based on the self report of offenders, with no means of comparing their accounts with actual test outcomes.

- Self-presentation biases may have influence self report of false positives and false negatives.

- Forty five percent of those approached to take part in the study declined to do so.
AUTHORS’ CONTRIBUTIONS

Don Grubin conceived the study, and was involved in the study design, data interpretation, redrafting of the article, and approval of the final article.

Lars Madsen was involved in the study design, data analysis and interpretation, preparing the first draft of the article, further redrafts, and approval of the final article.

Don Grubin MD FRC Psych
University of Newcastle/Newcastle, North Tyneside and Northumberland Mental Health Trust
St Nicholas Hospital
Gosforth
Newcastle upon Tyne NE3 3XT
don.grubin@ncl.ac.uk

Lars Madsen BBehSc (Hons) M Clin Psy
University of Newcastle
St Nicholas Hospital
Gosforth
Newcastle upon Tyne NE3 3XT

Correspondence to Professor Grubin