Polygraph-Based Testing of Deception and Truthfulness: An Evaluation and Commentary

COLIN TREDOUX* AND SUSAN POOLEY+

1 INTRODUCTION - THE POLYGRAPH IN SOUTH AFRICA

'The evidence of such a [polygraph] test is in reality little different from a police officer giving evidence that during an interview the accused shuffled, stammered or sweated profusely.'[1]

The charm of a machine or technology that can allegedly detect deception or "prove" innocence is obvious, and many such tests have been used for more than 2000 years in different parts of the world. The machine currently in vogue - since at least the 1890's - is a device that measures physiological activity and is known as the polygraph. In South Africa, an increasing number of companies have been seduced into making the polygraph a regular feature of the work environment. These companies include such giants of South African industry as Pick 'n Pay, Kentucky Fried Chicken, De Beers Marine, Platinum Group Minerals, First National Bank and ABSA. It is also used by a plethora of security and insurance companies like Gray Security Services and Multi Fund Insurance Brokers, as well as governmental concerns such as South African Airways and the South African Revenue Service.[2] In addition, the polygraph has been adopted with vigour by the South African Police Service - the SAPS recently paid for the dedicated training of 18 polygraphers by a foreign company,[3] and polygraph testing of deception was endorsed by the former SAPS commissioner George Fivaz in 1997.

We argue that this manifest and rapidly multiplying belief in polygraph tests of deception is wildly misplaced. Our argument derives from a comprehensive review of the scientific literature on polygraph tests of deception, and from an analysis of the South African labour cases in which polygraph evidence has been considered. We argue that the theory underlying polygraph tests of deception is fundamentally flawed, and empirical tests of its accuracy have found it to be little more accurate at detecting deception, or truthfulness, than tossing a coin. The use of the polygraph to detect deception in the workplace or elsewhere must result in unfairness and discrimination, and we suggest that it is time to consider a measure such as the USA Employee Polygraph Protection Act of 1988,

* Senior Lecturer, Department of Psychology, University of Cape Town.
+ Research student, Department of Psychology, University of Cape Town.
1 *Phipson on Evidence* 15ed by M N Howard (2000) 926.
2 List compiled from analysis of CCMA and other labour law reports as published on Jutastat Publications, as well as electronic versions of published newspaper articles from the period 1986 to 2000. Nearly all the legal cases cited in this article were sourced from the electronic databases of the South African Law Reports and its updated version on the Jutastat database.
3 See the Internet report at http://usaembassy.southafrica.net/YearBook1999/Safety/security.htm [12 December 2000].

which has served to protect American employees from the use of the polygraph in the workplace. Although Christianson[4,5,6] has written about the use of the polygraph in South Africa, her grasp of its scientific status has not been comprehensive or detailed, and we disagree with several of her major points. In particular, she argues that proper training of polygraphers can do a lot to improve the accuracy and reliability of the test,[7] but as we will show, the test is inherently unreliable and cannot be made reliable by any amount of examiner training.

We outline our argument against the use of the polygraph to detect deception in three sections. In the first, we provide a brief review of 76 reported cases from the CCMA, the Industrial and Labour Courts (from 1986 to 2000), and an analysis of relevant newspaper articles (from July 1999 to December 2000). In the second, we provide an up-to-date review of the scientific literature on the use of the polygraph as a means of detecting deception. In the third, we consider the scientific acceptability of using the polygraph in employee screening, a very common practice in South Africa.

2 Analysis of Labour Cases and Media Reports

In the private sector the polygraph is mainly used in staff disputes, frequently concerning theft, and is accorded an extraordinary amount of credibility by the companies that use it. In several disputes that have come before the CCMA, polygraph test results have provided the main evidence against the employee in the dismissal hearing leading up to the dispute.[8] Some companies go as far as to make passing a polygraph test a criterion of pre-employment selection[9] and some include clauses in their employment contracts that require employees to submit to polygraph testing on demand.[10]

Unsurprisingly, the rapid increase in polygraph testing in South Africa (approximately 20 000 tests are conducted per year),[11] accompanied by the highly lucrative business this represents for companies offering polygraph testing (in the region of R20m to R30m per year), has led to an increase in polygraph test results being offered as evidence in employment disputes heard by the CCMA and the Industrial and Labour Courts. Unfortunately, it has also become clear that this evidence is treated highly inconsistently in terms of admissibility, and commissioners' understanding of the scientific

[4] M Christianson 'Polygraph Testing in South African Workplaces: "Shield and Sword" in the Dishonesty Detection versus Compromising Privacy Debate' (2000) *ILJ* 21:1 16-38.
[5] M Christianson 'Truths, Lies and Polygraphs: Detecting Dishonesty in the Workplace' (1998) *Contemporary Labour Law* 8:1 1-10.
[6] M Christianson 'The Testing of Employees' (1999) *Contemporary Labour Law* 9:2 11-16.
[7] Christianson 'Polygraph Testing in South African Workplaces'.
[8] For example, see *Mahlangu v Cim Deltak, Gallant v Cim Deltak* (1986) 7 *ILJ* 346 (IC); *Malgas v Stadium Security Management* (1999) GA21495; *Eckstein v Queens Club Casino* (1999) KN33861; *Geni v Ceramic Tile Market* (1999) EC10121; *Govender v Containerlink* (2000) KN32917; *NUMSA obo Qhusheka v Alloy Wheels International* (2000) EC17665; *Singh v First National Bank* (1999) KN33299; *MTWUSA obo Qabalatsane v Fidelity Cash Management* (2000) NW14877; *Sibongile v New York Clothing* (2000) GA64451.
[9] For example, Galleon, Platinum Group Minerals and the elite police unit the Scorpions.
[10] For example, see *Boonzair v HICOR Trading Ltd* (1999) E18745.
[11] This estimate was offered by the South African Polygraph Association in a seminar held at the 2000 National Conference of the Psychological Society of South Africa, August 2000.

status, validity, reliability and research findings regarding polygraph tests of deception is, with due respect, often confused.[12] Some authors have pointed to the widespread coercive use of the polygraph,[13] calling it a "psychological rubber hose",[14] and it appears from analysis of CCMA cases that some employers may be treating the polygraph as a means of intimidating staff and extracting confessions.

Sometimes polygraph evidence is discarded as inadmissible,[15] but frequently it is allowed as evidence, even if it is not given much weight.[16] In our analysis of the 76 cases reviewed, it appears that polygraph test results have made their way into commissioners' judgements as a form of corroborative evidence. In some cases this has meant that it is used to support other, often highly suspect, evidence. The grave danger here is that this may result in situations where employers submit two or more unreliable pieces of evidence and combine them to make a "reliable" conclusion against an unfortunate employee. This is of particular concern in cases in which it appears that polygraph evidence has been used to tip the balance of probabilities.[17]

One of the reasons underlying commissioners' inconsistent handling of polygraph evidence stems from the testimony of polygraphers themselves. As expert witnesses, in most cases for the employer or respondent party, polygraphers tend to make extravagant and highly inaccurate claims as to the power of the polygraph in detecting deception, often exceeding 99% accuracy,[18] and they show a cursory understanding of the scientific literature relating to polygraph tests of deception. They rarely offer any citations to studies which report these ostensibly high accuracy rates, and tend to rely to a great extent

---

[12] For example, in *SACCAWU obo Sydney Fongo v Pick 'n Pay Supermarket* (2000) FS15555 the commissioner is prepared to accept that polygraph tests are 96% fool-proof, but in *Mahlangu v Cim Deltak, Gallant v Cim Deltak* (1986) 7 *ILJ* 346 (IC), it is said: 'The court cannot ignore the preponderance of expert opinion which holds the view that the use of a lie detector machine … is on scientific, psychological and ethical grounds, reprehensible.' Similarly, in *Jacob v Unitrans Engineering* (1999) KN21921, the commissioner says: 'It is absurd to assume that a man is guilty purely because he exercises a legitimate right to refuse to submit to a test or answer a questionnaire.' But in *Govender and Chetty v Cargo & Container Services* (1997) KN4881, the commissioner states that: 'There is no direct evidence linking the applicants to the theft. However, the inferences to be drawn are overwhelming. There is no doubt that the two employees were responsible. Their failure in two polygraph tests lends some support to this finding on the facts. However, their failure of the two polygraph tests strengthens the belief that they were involved.'

[13] For example, G Ben-Shakhar and J J Furedy *Theories and Applications in the Detection of Deception* (1990).

[14] The allusion here is to the practice of some police officers, in which suspected criminals are beaten with a rubber hosepipe in order to extract a confession, but without causing traceable physical damage. See repeated references in Ben-Shakhar & Furedy.

[15] For example, see *Mahlangu v Cim Deltak, Gallant v Cim Deltak* (1986) 7 *ILJ* 346 (IC); *Chauke and Jacobs v Gray Security* (1998) GA23984; *Rix v Ryder Security (Pty) Ltd* (1994) 3 *LCD* 68 (IC); *Mbele and Cash Payment Services (Pty) Ltd* (1997) KN4084; *Mnguni v Oryx Security* (1998) KN17731; *Malgas v Stadium Security Management* (1999) GA21495; *Govender v Containerlink* (2000) KN32917; *Maqanda, Wiseman and 2 Others v Fashion Fair Elegant* (2000) EC5025; *NUMSA v Glacier Van Der Vell (Pty) Ltd* (1997) KN3151; *NUMSA obo Qhusheka v Alloy Wheels International* (2000) EC17665; *Singh v First National Bank* (1999) KN33299; *Themba and Luthuli v National Trading Company* (1998) KN16887; *Gray Security Services (Western Cape) (Pty) Ltd v Cloeto No and Another* (2000) 21 ILJ 940 (LC), *SACCAWU obo Vereen and Valela v MAKRO SA (Pty) Ltd* (1999) EC11555.

[16] For example, see *Malgas v Stadium Security Management* (1999) GA21495; *Maqanda, Wiseman and 2 Others v Fashion Fair Elegant* (2000) EC5025; *CEPPWAWU and Francis v Thermopac* (2000) WE33153; *Amos v NBS* (2000) KN42888.

[17] See, for example, *Govender and Chetty v Cargo & Container Services* (1997) KN4881, and the commissioner's comments regarding polygraph test results as evidence quoted above in footnote 12.

[18] For example, see *Zoned and One Other v Floccotan (Pty) Ltd* (1997) KN2845; *Mbele v Cash Payment Services (Pty) Ltd* (1997) KN4084; *Singh v First National Bank* (1999) KN33299; *NUMSA v Glacier Van Der Vell (Pty) Ltd* (1997) KN3151.

on their own untested personal experience. They also completely ignore the very large body of scientific publications which argue that polygraph testing of deception is deeply flawed and highly unreliable. Employee applicants are often unable to offer counter testimony, probably due to the cost of acquiring expert opinion, leaving the naïvely enthusiastic testimony of polygraphers uncontested. Faced with such a situation, some commissioners have admitted polygraph test results and misleading expert testimony without the benefit of opposing testimony.[19]

We would like to call attention to three particular ways in which polygraph testing is presently used in South Africa, and which we believe to be highly problematic. The first is the way in which polygraph tests are used to place the onus of proof on the employee. In many cases, employees who are suspected of theft have been forced to take polygraph tests to prove their innocence. If they refuse, they are likely to be dismissed, and if they fail the test, they are summarily dismissed.[20] The second is the way in which the polygraph is used to blanket-test employees when a theft has occurred in an organization. In other words, an entire group of employees – all of whom may be completely innocent - will be tested in order to determine a guilty party, often in the absence of other evidence.[21] Finally, we point to a disturbing new trend in the construction of employment contracts. Although commissioners have held that employees are within their rights to refuse a polygraph test,[22] employers appear to have responded to this by including clauses in employment contracts that oblige the employee to undergo polygraph tests on demand. In so doing, they give their employees no option but to submit to a scientifically dubious procedure. In fact, some employees have objected to taking the test on the (correct) precept that polygraph testing is scientifically unsound, and have been dismissed as a result.[23]

Just as employer-respondent cases supporting the use of the polygraph in employment settings are frequently based on the self-serving testimony of polygraphers, so too is a large proportion of the press coverage. We have reviewed the national media between 1999 and 2000,[24] and posit that the regular appearance of uncritical and laudatory newspaper reports on the polygraph has created an atmosphere

---

[19] For example, see *Govender and Chetty v Cargo & Container Services* (1997) KN4881; *Themba and Luthuli v National Trading Company* (1998) KN16887.

[20] See, for examples, *Armoed v Gray Security Services* (1999) EC9809; *Boonzaier v HICOR Trading Ltd* (1999) E18745; *Malgas v Stadium Security Management* (1999) GA21495; *Zoned and One Other v Floccotan (Pty) Ltd* (1997) KN2845.

[21] For example, see *Kgapare v Robridge CC* (1998) GA34288; *Mack v Hire Anything* CC (1999) KN33581; *Mquanda & 2 Others v Fashion Fair Elegant* (1998) EC5025; and *Singh v First National Bank* (1999) KN33299.

[22] Cases, and also note that it is against the American Polygraph Association's code of ethics to conduct a test under such circumstance. See Clause 3.7.2 of the code of ethics, at http://www.polygraph.org/apa1.htm#code [19 January 2001].

[23] This was the case in *Cunningham v Benguela Operations (Pty) Ltd* (1999) C542/98, before the Labour Court, which was settled by the respondent. Cases in which employees were dismissed for refusing to take a polygraph test to "prove" their innocence include: *Jacob v Unitrans Engineering* (1999) KN21921; *Armoed v Gray Security* (1999) EC9809; *Boonzaier v HICOR Trading Ltd* (1999) E18745; *Chemaly v Super Rent* (1998) KN15402; *PTWU obo Matshivha v Fidelity Guards Services* (2000) GA72762; and *T & G Workers Union obo Cloete v Gray Security Services* (1999) NC2396.

[24] Between 07/99 and 04/12/00, 36 separate articles were published by newspapers in the Independent Newspaper Group (accessed through Independent Online).

of acceptability for polygraph tests of deception, and is often little more than advertorial copy. For example, a news story in *Finance Week* titled 'Catching the Liars',[25] opens as follows:

'The use of polygraph or lie-detector testing in insurance claims could save the short term industry millions of rands every year while also reducing the level of premiums.'

The article goes on to advertise the services of a particular polygraph-testing company and fails to offer a single comment as to the reliability or validity of the polygraph or the ethics of seducing clients with lower premiums based on a highly suspect and scientifically-unsound device.

A second very strong theme in the press is the glorification of the polygraph as a general-purpose crime-fighting device. A cursory glance at headlines reveals this: 'Polygraphs now Weed out the Weasels',[26] 'Poligraaf: Bestuur se Vriend',[27] and 'Catching the Liars'.[28] This leads in extreme cases to completely mistaken reports about the efficacy of the polygraph - in an article in *The Star* titled 'Man Lied about Wife' s Killing Gets Jail', the journalist mistakenly reported that polygraph evidence formed the basis for a conviction, when in fact it was used as a coercive technique to extract a confession.

It would seem that the national press and South African business currently holds polygraph testing in great reverence. This is not surprising, given the long history of fascination with "truth" technologies in many parts of the world. The very idea of a machine that could make one of the most difficult human tasks perfectly reliable by simply automating it, is a naturally appealing one for a country that is wracked by crime and talk of crime. It is equally unsurprising that this reverence is misplaced. The following review of the scientific literature on polygraph tests of deception makes this clear.

3 The Scientific Status of the Polygraph in Detecting Deception

Despite polygraphers' claims to the contrary, the technology of the polygraph has not changed significantly since the early 1900s. Although advances have been made in the technology for making permanent records of polygraph data, such as digitisation and magnetic recording on computer, the theory and methodology is virtually unchanged.

The polygraph is merely a device that measures and records electro-physiological activity. Electrical signals are transduced from the body through conducting electrodes and are then filtered and amplified through electronic circuitry so that accurate measurement and recording of physiological activity may be made. In older polygraph models the recording is made by transmitting the electric signal into deflections of an ink pen on paper, and in newer models the signal is digitally transformed

[25] *Finance Week* 24-30 October 1996 at 12.
[26] *Financial Mail* 11 September 1998 at 102.
[27] *Finansies & Tegniek* 7 February 1992 at 56.
[28] *Finance Week* 24-30 October 1996 at 12.

and recorded onto a permanent medium such as a computer's disk drive. Typically, several "channels" of physiological activity are recorded (hence the term "*poly*graph"), including differential blood pressure, heart rate, respiration rate and skin conductance (subcutaneous sweating). These four channels are usually recorded when the polygraph is used in an attempt to detect deception or, alternatively, "prove" innocence.

The accuracy of the physiological measurements made by the polygraph is not in question. Provided the apparatus is in good order, and the polygraph operator follows generally accepted procedures, the physiological measurements are highly accurate.[29] However, the polygraph cannot and does not measure deception or lying, or the absence of deception or lying. It merely records physiological activity, and any attempt to use it to detect deception involves drawing an *inference* from the physiological activity that it records. In fact, there is no known physiological or psychological measuring instrument that directly records deception or lying. Therefore, the central question regarding the use of the polygraph in the physiological detection of deception does not concern the reliability of the polygraph physiological measuring instrument, but the theory and method underlying the way inferences are drawn from the polygraph's physiological recordings.

3.1 *The status of the scientific theory underlying the use of the polygraph in the physiological detection of deception*

The theory underlying the use of the polygraph in the physiological detection of deception is simply that physiological activity increases when a person is lying. There is no specific pattern to this increase - the increase occurs in all of the four channels typically recorded. In order to detect deception, then, one needs to compare the rate of physiological activity at a particular point in time to the activity at another point in time so that any changes can be observed. This can be done by posing a critical question to a polygraph examinee (for example, 'Did you steal the money?') and observing the contemporaneous physiological activation, and comparing this activation to that made when the subject is posed an irrelevant question (for example, 'Is your name Brian?').

Many critics of polygraph tests have pointed out that there is no necessary reason why physiological activation should increase when a person tells a lie, nor is a mechanism or process offered that would explain this. It is *assumed* rather than demonstrated.

Since the assumption underlying detecting deception relies on observing increased physiological activity or arousal, and since a question regarding a particular event might be more arousing than a neutral comparison question, it is clear that the comparison question needs to be formulated very carefully. It is generally and widely acknowledged that a person may easily show more arousal to a

[29] See D T Lykken *A Tremor in the Blood: Uses and Abuses of the Lie Detector* (1981).

critical question than to an irrelevant question, and yet be quite *truthful* in his or her response to the critical question.[30]

The matter of finding the right kind of comparison question is often formulated as an issue of scientific control. That is, in order to draw an inference that the increased physiological arousal is due to an act of attempted deception, one must ensure that the only possible difference between the responses to the two types of question is the presence of deception. The difference should not be ascribable to inherent differences between the questions, such as their "arousability".

There are several different kinds of methods or tests that use the polygraph in the detection of deception, each attempting to provide a solution to the problem of the control question. The two most widely used tests are the Relevant-Irrelevant Test (RIT) and the Control Question Test (CQT).

The RIT is the oldest and most widely used in the physiological detection of deception. In it the critical questions are relevant to the purpose of the assessment (for example, to find out who stole the money), and the control questions are irrelevant. If physiological activity is greater for relevant than for irrelevant questions, then the examiner concludes that these questions have greater significance for the examinee, or that the examinee is being deceptive.

The RIT has fallen into disrepute in recent years, particularly in the scientific literature, because it is thought to institute a very weak form of control, and virtually all commentators now view it as unacceptable.[31] For instance, the question 'Did you shoot the deceased?' is more likely to result in physiological arousal than the question 'Do you live in Cape Town?', regardless of whether the examinee is guilty or innocent. Results from research studies indeed bear out the contention that innocent subjects are just as likely to be falsely labelled deceptive as correctly labelled innocent.[32]

The polygraph test most frequently used in criminal and intra-organizational security investigations is the CQT. This test compares the physiological arousal of examinees in response to critical questions to their level of arousal when telling an unrelated lie. For example, an examinee may be asked 'Before today, did you ever take something that didn't belong to you?' (The exact nature and wording of the questions put to the examinee are agreed by the examinee and polygrapher in a pre-test interview.) It is assumed that most people will have taken something that didn't belong to them at some stage in their lives but are reluctant to admit to it. The examinee's level of arousal when "lying" in regard to this question (everyone is expected to lie) is recorded and used as a benchmark against which to compare critical questions relating to the matter under investigation. In other words, a situation is set up where the polygrapher gets the examinee to lie, records the physical reactivity, and then compares this with

[30] For example, see Ben-Shakhar & Furedy and Lykken *A Tremor in the Blood.*
[31] See D Raskin, C Honts and J Kircher 'The Scientific Status of Research on Polygraph Techniques: The Case for Polygraph Techniques' in D Faigman, D Kaye, M Saks and J Sanders (eds) *Modern Scientific Evidence: The Law and the Science of Expert Testimony* (1997) 565-582 at 568; and W Iacono and D Lykken 'The Scientific Status of Research on Polygraph Techniques: The Case Against Polygraph Techniques' in D Faigman, D Kaye, M Saks and J Sanders (eds) *Modern Scientific Evidence: The Law and the Science of Expert Testimony* (1997) 582-618 at 585.
[32] Raskin 'Does Science Support Polygraph Testing?'

the reactivity to critical questions. If the level of activation recorded by the polygraph is greater for the critical question than the so-called control question, the examiner will take this as evidence of deceptive behaviour. Conversely, if the level of activation is greater for the control question than the critical question, the examiner will take this as evidence of innocent behaviour.

In some versions of the CQT a number of control- and critical-question pairs are put to the examinee, and the physiological responses are then compared for each pair. A quasi-numerical method of scoring devised by David Raskin, formerly of the University of Utah, is often used in this so-called "Zone of Comparison" variant of the CQT. A score ranging from -3 to 3 is assigned to each critical-control pair of questions, depending on whether and to what extent the physiological activation differs across control and critical questions. Total scores greater than 0 are judged to indicate deception and scores smaller than 0, truthfulness, although in practice many examiners take scores between -5 and +5 to be inconclusive.[33]

It should be noted that there is more to the CQT than just the procedure outlined above. The "successful implementation" of the CQT is critically dependent on the pre-test interaction between the polygrapher and examinee. For instance, the examinee must be induced to tell a lie, falsely believing that the examiner is not aware that a lie is being told. Also, a suitable so-called "stimulation test" must be set up in order to induce confidence in the examinee that the polygraph accurately measures deception. This stimulation test usually involves actively *deceiving* the examinee into believing that the polygraph will determine the identity of a playing card ostensibly chosen at random by the examinee, but in fact marked by the polygraph operator prior to distribution. In other words, the reliability and validity of CQT results largely depend on how good the examinee (falsely) believes the polygraph is at detecting deception. One does not have to be overly cynical to notice the similarity of this procedure to the stage tricks of magicians: a confidence trick is embedded deep within the structure of polygraph testing.

Partly for these reasons the methodology of the CQT has attracted a great deal of criticism in the scientific and professional literature. In particular it is criticised on ethical grounds, as the examinee must be intentionally misled for the test to be "successfully" conducted. It has also been criticised on scientific standardisation grounds, in that the administration of the CQT is highly dependent on the skill of the examiner to find suitable critical and control questions, and is therefore impossible to standardise across different examiners, even when they have had the same training.

A particular methodological problem has to do with the confusion between the role of polygraph examiner as both the operator of the polygraph machine and the interrogator of the examinee. In his/her role as the interrogator, the polygrapher may be privy to a great deal of information about the examinee, including non-polygraphic information such as alleged incriminating evidence offered by the employer, which is most often the party paying the polygrapher. This information has been shown

[33] See Iacono & Lykken 'The Scientific Status of Research on Polygraph Techniques: The Case Against' at 586.

to influence the polygrapher, even without the polygrapher being consciously aware of this influence, resulting in the decision as to the truthfulness of the examinee being based on the non-polygraphic information and not on the physiological evidence.[34] There is also abundant evidence that the polygraph is used in many settings as a coercive way of extracting confessions.[35] For these reasons the American Polygraphic Association suggests that polygraph charts be independently scored by a second "blind" examiner who is not privy to this information.

In addition to the above problems in the administration of the CQT polygraph test, the major theoretical problem with the CQT from a scientific point of view is that the control questions used in the test are not true or effective control questions. In order for the questions to constitute a form of scientific control, they would have to demonstrate that any increase in physiological activation is *solely* due to attempts at deception. However, there may be many equally viable explanations. It is quite conceivable that an innocent person will show higher levels of physiological arousal when posed a critical question simply due to the content of the critical question, especially if the examinee believes they are under suspicion for a crime such as theft. This is very often the case in labour disputes.[36]

It is also well established that a guilty or deceptive person can "beat" the polygraph by adopting appropriate physical and mental countermeasures. These countermeasures, including counting in 13s back from 1000 and pinching oneself, raise physiological activity levels when answering the control questions. This results in the pattern of activation being difficult to judge, and may well lead to a judgement of truthfulness or an undecided judgement, as it certainly has in many laboratory studies.[37]

In summary, the theoretical rationale of polygraph tests of deception is widely acknowledged as highly questionable. A series of improbable assumptions underlie the commonly used polygraph-deception tests, and the methodology employed in the administration of these tests does not satisfy some basic canons of scientific practice. The most important of these is that the tests do not use true or reasonable forms of control, and therefore are unable to distinguish between observed increases in physiological activation which are due to attempted deception or some other reasonable cause, such as anxiety or self-inflicted pain.

---

[34] D Carroll 'Lie Detection: Lies and Truths' in R Cochrane and D Carroll (eds) *Psychology and Social Issues* (1991) 160-170.

[35] Lykken *A Tremor in the Blood*; and Ben-Shakhar & Furedy. See also the following cases: *Thulisie v Game* (2000) KN33380; *Msomi v Beacon Sweets & Chocolate* (1999) KN27774; *Maqanda and 2 Others v Fashion Fair Elegant* (1998) EC5025; *ECCAWUSA obo Rosy Nhlapo and Sipho Busakwe v Miladys* (1999) GA30036; and *Van Zyl v Game* (1999) FS6652.

[36] In many of the CCMA cases in which the polygraph had been used to test deception, employers had rounded up all possible suspects and polygraphed them, and then dismissed as the culprits those who failed the test. See note 21 for examples of cases in which this occurred. It is possible that under such conditions stress could be the reason for the observed increased arousal. The British Psychological Society in its 'Report of the Working Group on the Use of the Polygraph in Criminal and Personnel Screening' (1986) *Bulletin of the British Psychological Society* 39 81-94 at 88, gives several examples of alternative explanations for observed arousal other than lying, including stress or anxiety.

[37] For example, see C Honts and J Kircher 'Mental and Physical Countermeasures Reduce the Accuracy of Polygraph Tests' (1994) *Journal of Applied Psychology* 79:2 252-259.

3.2 *Empirical research on polygraph accuracy*

Many adherents of polygraph testing do not concern themselves with the theoretical problems that beset polygraph testing, and address themselves instead to the empirical matter of whether physiological detection of deception tests are able to accurately distinguish between deceptive and truthful examinees. A great many empirical research studies have been conducted in order to answer this question.

Most of the research has focused on the two most widely used such tests, the Relevant-Irrelevant Test and the Control Question Test. In the case of the RIT described above, the results have been overwhelmingly negative: it is unable to correctly identify truthful responses at an accuracy level greater than 50%, or chance.[38] It should be noted that an accuracy level of 50% is, from a scientific perspective, essentially zero accuracy. The accuracy of any polygraph test of deception is dependent on how much *higher* a rate than 50% it achieves, since one could achieve 50% accuracy by merely flipping a coin. The very poor accuracy rate of the RIT has convinced almost all researchers in the field that it is of no use in the detection of deception.

In the case of the CQT, essentially two kinds of empirical research have examined its accuracy in the detection of deception: laboratory-based research and field studies. Traditionally, scientists have conducted laboratory research because it enables them to control the environment, in this case to find out with certainty who is telling the truth and who is lying, which is done by constructing the research design in such a way as to make deception randomly determined. Laboratory research has typically made university students into deceivers by having them commit a mock crime, such as stealing a watch from an office, and then instructing them to lie about it during a subsequent test.

However, laboratory research in general, and laboratory research into deception measurement in particular, is often severely criticised for a lack of realism. The contrived conditions under which the polygraph test is conducted are very unlike those in real life. In particular, the examinee in a real situation may face severe consequences if identified as deceptive, and this can be expected to induce a considerable amount of anxiety, even in innocent examinees. Since increased physiological arousal is known to accompany increased anxiety, it is likely to be more difficult to distinguish truthful and untruthful examinees in a real situation than a laboratory situation. In other words, results from laboratory studies are likely to be *overestimates* of accuracy.

Notwithstanding these difficulties, in excess of 40 high-quality laboratory studies have been conducted over the past 30 years. These studies provide ambiguous results regarding the accuracy of the CQT: scientists are divided on how to interpret the results and on what weight to attach to particular studies. The percentage of deceptive subjects which the CQT correctly classifies as deceptive ranges from approximately 53% to 90%, with a weighted average percentage of

[38] See S Horowitz, J Kircher, C Honts and D. Raskin 'The Role of Comparison Questions in Physiological Detection of Deception' (1997) *Psychophysiology* 34:1 108-115.

approximately 78%.[39] This corresponds to a rate of 78 in 100, and is 28 in 100 better than mere guessing would have achieved.

The percentage of truthful subjects which the test correctly classifies as truthful ranges from approximately 70% to 90%, with a weighted average percentage of approximately 84%.[40] This corresponds to a rate of 84 in 100, and is 34 in 100 better than mere guessing would have achieved. However, in the opinion of many critics of the polygraph tests of deception, these figures are highly inflated, mainly due to the unrealistic nature of laboratory experiments. This lack of realism means that one should not generalise the results obtained in the laboratory to real-world settings.

Whilst scientists are divided about the results of laboratory-based studies, they are virtually unanimously in agreement that field or "real-life" studies provide the most useful information about the identification accuracy of polygraph tests of deception.

In order to conduct a field study of polygraph tests of deception, a researcher first has to identify a number of cases in which the polygraph was used to test deception, and then whether the "ground-truth" status of the polygraph examinee is now known (ie whether the examinee was truthful or not). It is, of course, a problem to unambiguously determine ground truth: some method independent of the outcome of the test is needed for determining who is in fact telling the truth. The most commonly used criterion of ground truth in these studies is the confession of a polygraph examinee, usually made after being confronted with the results of a failed polygraph test, or of an alternate person who confesses to a crime or action the polygraph examinee was suspected of.

Several authors set out a number of other methodological requirements for acceptable field studies, which will not be detailed here except to note that scientific opinion is not in unanimous agreement about exactly what these should be. As a result, different authors accept different sets of studies as adequate, and there is a lack of agreement about the accuracy rates of polygraph CQT administrations in the field.

One influential set of polygraph critics, William Iacono and David Lykken of the University of Minnesota, argue that most field studies suffer from problems of sampling bias. This is because researchers only tend to select cases for study where the polygraph has been used and a confession has been extracted. The only cases thus selected for analysis are those involving a guilty suspect who failed a CQT and subsequently confessed, which constitutes sampling bias. All the cases in which an innocent person failed the test are omitted from study. Similarly, cases where the suspect was guilty, but in which the CQT failed to identify deception, are also excluded. The only cases that typically remain in the samples of field studies are those where a confession occurred, and where the polygraph either indicated deception or truthfulness. In other words, such studies are based on cases where the polygraph was subsequently and independently found to be "accurate".

[39] See Raskin, Honts & Kircher 'The Scientific Status of Research on Polygraph Techniques: The Case For'.
[40] ibid.

This problem of sampling bias has never been successfully dealt with in a field study, and the existing data from field studies is almost certainly unrepresentative of real-life polygraph applications. Despite this lack of agreement regarding desirable methodological features, a number of authors have recently presented accuracy rates for field studies. David Raskin, Charles Honts and John Kircher, who are all well known as advocates of polygraph testing, in an article in a 1997 legal treatise on scientific evidence, identify the field studies they find acceptable and summarise the results regarding accuracy.[41] By their reckoning, the percentage of deceptive subjects which the CQT correctly classifies as deceptive ranges between approximately 73% to 100%, with a weighted average percentage of approximately 86%. This is 36% more than could be expected from mere random guessing and at first may appear very impressive. However, and most importantly, the percentage of *truthful* subjects the test correctly classifies as truthful ranges from approximately 30% to 83%, with a weighted average percentage of only about 50%. This is no more than could be expected from mere random guessing or flipping a coin!

Raskin, Honts and Kircher suggest that the accuracy rates may be deflated by the presence of what they call an 'outlier point', a result that has a disproportionately large influence on the overall accuracy rate because it is extreme, which they remedy by simply omitting it.[42] Omitting inconvenient data is, needless to say, a highly questionable practice. If this is done then the weighted average percentage of correctly classified truthful subjects is 64%, which is 14% more than could be expected from guessing.

Raskin, Honts and Kircher make further arguments, pointing out that if one looked only at the original examiners without taking the "blind" or independent examiners into account, the accuracy for detecting either truthful or deceptive people would be over 80% in each case. However, they point out earlier in the same research article that it is widely agreed that independent examiners *must* be used to blind-score polygraph charts to avoid the undue influence of extraneous factors.

William Iacono and David Lykken, who are both well known as opponents of CQT polygraph testing but not of polygraph testing *per se* (Lykken invented the Guilty Knowledge Test for the polygraph), in an article in the same 1997 legal treatise on scientific evidence,[43] identify the field studies they find acceptable and then summarise the results regarding accuracy.

The percentage of deceptive subjects which the CQT correctly classifies as deceptive ranges from approximately 77% to 98%, with a weighted average percentage of approximately 84%, ie 34% more than could be expected from random guessing. The percentage of truthful subjects the test correctly classifies as truthful ranges from approximately 51% to 63%, with a weighted average percentage of approximately 56%. This is not much more than could be expected from guessing.

---

[41] ibid.

[42] The study in question was conducted in 1991 by Patrick and Iacono. See C Patrick and W Iacono 'Validity of the Control Question Polygraph Test: The Problem of Sampling Bias' (1991) *Journal of Applied Psychology* 229-238. They found only a 30% accuracy prediction of truthful subjects.

[43] Iacono & Lykken 'The Scientific Status of Research on Polygraph Techniques: The Case Against' at 575.

In summary, researchers who have examined the accuracy rates of the RIT and CQT polygraph tests agree in some respects, but disagree in many others. They agree that the RIT gives very poor results, particularly when it comes to identifying truthful subjects, which it fails to do with any more accuracy than random guessing. Although researchers disagree about the overall accuracy rates of the CQT in laboratory studies, it is fair to say that there is evidence to suggest that accuracy rates are about 80% in respect of detecting either truthful or untruthful subjects. However, researchers agree that laboratory studies are likely to be unrepresentative of the accuracy rates of polygraph tests in practice, and further agree that field studies will give more reliable information in this respect. They disagree in a profound manner regarding the methodology field studies should use. At best, the two most prominent groups of polygraph researchers provide evidence that places the accuracy rate of the CQT in the field no higher than 86% for untruthful subjects and only 64% for truthful subjects. In fact, the accuracy rate for truthful subjects may be as low as 50%, which is no better than random guessing.

Regardless of which accuracy rates are accepted, it should be clear from the above discussion that researchers in the field of polygraphy and the physiological detection of deception disagree about the scientific status of the polygraph in detecting deception. Although there is agreement to a certain extent about the theoretical weaknesses,[44] and although there is near-unanimity regarding the scientific status of the RIT (ie it is highly unreliable), there is sharply divided opinion between two sets of scholars concerning the CQT and its overall accuracy and usefulness in detecting truthfulness or deception.

On the one hand, the group of researchers consisting of David Raskin and his former first- or second-generation students (Charles Honts, Susan Amato, John Kircher) have published many papers in which they argue that the accuracy rates from suitably conducted studies are sufficiently high to justify using the CQT as a means of detecting deception. On the other, David Lykken and William Iacono have published many papers in which they argue that the accuracy rates from suitably-conducted studies show that the CQT has a very high probability of classifying truthful people as deceptive (ie false positives), and that it should *not* be used as a means of detecting deception. A major part of their disagreement stems from the fact that the two groups cannot agree on which studies should be included and excluded.

Disagreements of this sort are not unusual in scientific fields of study, and one way of deciding the issue is to refer the matter to the community of scientists who work in the relevant area, in this case the community of scientists who are psychophysiologists. But there is a case to be made for extending the relevant community to a broader group of psychologists, since the polygraph's use of

[44] See Raskin, Honts & Kircher 'The Scientific Status of Research on Polygraph Techniques: The Case For'; and Iacono & Lykken 'The Scientific Status of Research on Polygraph Techniques: The Case Against'.

physiological measures is trivial and not at issue, and relies to a far greater extent on principles of research design and method.[45]

There have indeed been a number of surveys over the past 20 years to establish if there is a consensus position regarding the accuracy of the polygraph in detecting deception. Three surveys are commonly cited in the literature, all conducted in the USA, but full information in the form of a published report in a peer-reviewed journal is only available for one of these, the Iacono and Lykken survey.[46] In 1982, the Gallup organization conducted a telephone survey poll of 155 members of the Society for Psychophysiological Research. Details regarding the methodology of this survey are unclear and the study has not been published in a recognised, peer-reviewed scientific journal (see results table below). In 1993, Susan Amato, a student at the University of North Dakota, conducted a postal survey of members of the Society for Psychophysiological Research. Of the 450 members targeted in the survey, 135 or 34% responded. This study has only been published in summary form (approximately 200 words), and it is accordingly difficult to evaluate. In 1995, William Iacono and David Lykken conducted a survey in which they randomly sampled 214 members of the Society for Psychophysiological Research in a postal survey. They had a high response rate of 91%. Iacono and Lykken also surveyed 226 Fellows of the American Psychological Association. Fellows are all elected and are considered elite, high quality scientists. Of these, 226 or 74% returned the questionnaires and the results of this study were published in the *Journal of Applied Psychology*.[47] In terms of the standard scientific requirements for conducting surveys, it is clear that the third survey should be accorded the most weight. It attempted to draw a representative sample of the population in question and it ensured high response rates, both of which are typically a problem in survey research.

The results of the above three surveys differ to some extent in the amount of support the community of relevant scientists accords the CQT as a method of detecting deception. They can specifically be compared on a question all three of the surveys utilised for comparability. In this question, respondents were asked to choose one of four statements best describing their opinion regarding the CQT method of interpreting whether a subject is or isn't telling the truth. The results for the three surveys, in respect of this four- or five-part question, are set out below:

[45] ibid.
[46] W Iacono and D Lykken 'The Validity of the Lie Detector: Two Surveys of Scientific Opinion' (1997) *Journal of Applied Psychology* 82:3 426-433.
[47] ibid.

Response options:	Percent agreeing		
	Gallup (1982)	Amato & Honts (1993)	Iacono & Lykken (1995)
A. Sufficiently reliable method to be the sole determinant	1	1	0
B. Useful diagnostic tool when considered with other available information	61	60	44
Between "B" and "C"	2	Not asked	Not asked
C. Questionable usefulness, entitled to little weight against other available information	32	37	53
D. No usefulness	3	2	2

Table 1: Comparison results of three surveys of polygraph experts

It is clear from the results of the surveys that almost no respondents endorse the opinion that CQT polygraph tests are sufficiently reliable on their own to judge deceptiveness. Iacono and Lykken went further than the question outlined above and asked specific questions about the CQT. The responses to these are summarised below, for each of the samples (SPR members and APA Fellows).

Question:	Percent agreeing	
	SPR members	APA Fellows
1. The CQT is scientifically sound	36	30
2. I would admit failed tests as evidence in court	24	20
3. I would admit passed tests as evidence in court	27	24
4. The CQT can be beaten	99	75
5. The CQT is at least 85% accurate for:		
a) guilty	27	Not asked
b) innocent	22	Not asked
6. It is reasonable to use laboratory studies to estimate CQT validity	17	Not asked

Table 2: Results of the polygraph CQT survey conducted by Iacono and Lykken in 1997

It seems that the majority of both the psychophysiological research community and the community of APA Research Fellows consider the CQT to be scientifically unsound; that it can easily be beaten by countermeasures, and do not think the outcomes of these tests should be presented as evidence in court. In addition, they do not think that the tests are 'at least 85% accurate', as is often claimed in court by polygraphers using the CQT method – in front of the CCMA, polygraphers frequently assert that the CQT accuracy is higher than 99%!

In addition to these survey measures of scientific opinion regarding the use of the polygraph, it should also be noted that in the 1980's there were two major scientific committees of inquiry into lie-detector testing both in the USA and the UK. In 1983 in the USA, the Office of Technology Assessment was requested by Congress to evaluate the polygraph and its use in the physiological

detection of deception.[48] The request followed the suggestion by President Reagan that the polygraph be used more extensively in government settings. A summary of the report, published in 1985, concludes: 'Neither available data nor theoretical analysis indicates that polygraph tests function as claimed by their proponents. Substantial numbers of both untruthful and deceptive individuals may be misidentified through the use of polygraph tests, and the test can be "beaten". For most common uses of polygraph testing there is not even rudimentary evidence to support such use.'[49]

In the UK, the British Psychological Society (BPS) – an organization that registers and regulates psychologists - appointed a special commission of enquiry to investigate the scientific validity of the polygraph. The special committee of the BPS delivered its report in 1986, and said the following regarding polygraph-testing procedures: 'Our conclusion is that while polygraph-based techniques have some limited application in forensic investigations, they are unlikely to be acceptable in the British context of employment and staff screening. Even in the context of criminal investigation, there is controversy over the theoretical rationale behind lie-detection procedures, and their accuracy and efficacy.'[50] The BPS also found that: '...[polygraph tests are] contrary to the spirit of the Society' s Code of Conduct.'[51]

Finally, although it cannot be claimed that there is unanimity amongst scientists regarding the question of whether polygraph tests of deception are scientifically sound and are accurate in identifying truthfulness or deception, the preponderance of opinion in the relevant scientific communities is that they are neither scientifically sound nor accurate at identifying deception or truthfulness.

4 THE SCIENTIFIC STATUS OF THE POLYGRAPH IN EMPLOYMENT SCREENING

Of particular importance in the South African context regarding the use of the polygraph is its increasing use in pre-employment and employment screening. Whilst the polygraph is used in much the same way in employment, criminal justice and employment settings, in employment settings the polygraph is also used to make judgements about the supposedly enduring traits of honesty and truthfulness. The critical difference is that most criminal justice and security applications of polygraph detection of deception technology are *event related*. In employment settings, the applications are frequently *event free*. That is, the polygraph is used to make judgements as to whether an employee is lying about any of a number of issues posed to him or her. For example in an article in 1992, Johan

[48] See: Office of Technology Assessment, US Congress 'Scientific Validity of Polygraph Testing: A Research Review and Evaluation' (1983).
[49] L Saxe, D Dougherty and J Cross 'The Validity of Polygraph Testing' (1985) *American Psychologist* 40:3 355-366.
[50] British Psychological Society 'Report of the Working Group on the Use of the Polygraph' at 81.
[51] ibid.

Rossouw, an economist writing on behalf of the Afrikaanse Handelsinstituut in *Finansies & Tegniek*, makes the following claim in article titled 'Poligraaf: Bestuur se Vriend':[52]

'Die gebruik van die poligraaf om die lojalitiet en eerlikheid van werknemers te bepaal, bied die bykomende voordeel dat dit as ' n soort meganiese gewete optree wat die individu ontmoedig om aan oneerlikheid toe te gee.' ['The use of the polygraph to determine the loyalty and honesty of employees offers the additional advantage that it can function as a sort of "mechanical conscience" which dissuades the individual from succumbing to dishonesty.'][53]

Therefore, two kinds of polygraph test are predominantly used in employment screening, and these follow from the distinction between event-related and event-free situations. In cases where a particular event is at issue, the CQT is frequently used, as this test was specifically devised to deal with event-related situations. However, despite being slightly better at detecting deception than the RIT, the CQT, as was stated above, is not generally accepted by researchers in the field, or by professional psychologists and psychophysiologists to be scientifically sound or accurate at detecting deception.

In cases where no particular event is at issue and the aim is to identify general truthfulness or deceptiveness in the examinee's responses, the RIT is most frequently used. Indeed, the RIT is overwhelmingly the most widely used test in organizational settings.[54] This test is used in employment and pre-employment screening because it allows the examiner to ask questions about a number of different potential sources of dishonesty, whereas the CQT is geared to asking questions about specific issues.

As mentioned above, empirical research on RIT has proven it to be seriously flawed and inaccurate. Although there is some disagreement within scientific circles concerning the CQT, in the case of the RIT it is virtually unanimously agreed that the test is of very little value. For example, David Raskin, a known advocate of polygraph testing, says in a 1988 review: 'Pre-employment polygraph screening in the private sector represents the worst case in terms of accuracy and the problems inherent in that type of testing'.[55]

An additional, and very serious, problem with the use of either the CQT or RIT in employment screening relates to what is known as the "base rate" problem. It is likely that the kinds of behaviours that employment screening wishes to protect against occur with relatively low frequency in the relevant population.[56] For example, let us assume that the true proportion of applicant employees who have stolen from any of their previous employers is 0.2, or 1 in 5. Let us assume further that the accuracy rate of the polygraph test used to detect deception is 0.8 for both guilty (deceptive) and innocent (truthful) applicants (even though it is likely to be much lower than this). This means that the likelihood of correctly classifying an innocent or truthful person in this situation is 0.5. That is, only

[52] Rossouw 'Poligraaf: Bestuur se Vriend' at 56.
[53] Our translation.
[54] Raskin 'Does Science Support Polygraph Testing?'
[55] ibid.
[56] ibid.

50 of every 100 innocent people will be correctly classified by the test, even though the test is allegedly 80% or more accurate! This follows from a Bayesian probability formulation, and its troublesome conclusions are well recognised in many fields where the accuracy of tests is at issue.

The problem of the base-rate should not be underestimated. It leads David Raskin to comment as follows (in 1986, before the Employee Polygraph Protection Act):

'It is estimated that approximately two-million people each year are given employment polygraph tests in the United States ... even with the generous assumption of 85 percent average accuracy for [polygraph] tests, approximately 320 000 honest people would be denied employment because of false positive errors.'[57]

There is overwhelming consensus amongst scientists in the USA that polygraph tests cannot be used with any acceptable degree of confidence in employment screening situations. It is noteworthy that this conclusion is shared by researchers who are known for advocating the use of the polygraph as a lie detector in other settings, and by researchers who are known for their opposition to almost all uses of the polygraph in the physiological detection of deception.

The report to the US Congress by the Office of Technology Assessment - referred to above - was shortly followed by the drafting of the Employee Polygraph Protection Act of 1988, and this has virtually eliminated the use of the polygraph in employment screening in the USA.[58] We believe that due to the problematic scientific status, admissibility as evidence in labour and other non-criminal disputes, and use in employment of the polygraph in the physiological detection of deception that such an act is also essential in South Africa.

5 CONCLUSIONS

We have argued in this paper that the current wave of enthusiasm in South Africa for polygraph-based tests of deception is fundamentally misplaced. Polygraph tests of deception are offered as a general panacea for fighting crime in South Africa, in both private and public sectors, where their application includes disingenuities like blanket-screening employees, extracting confessions, dismissing people who refuse to take the test, and selling insurance policies at reduced premiums.

We have gone to some lengths to show that the scientific basis of the polygraph test is extremely flimsy - it is based on an implausible theory and methodology, and the accuracy of the machine is in considerable doubt. The test may literally be no more accurate than tossing a coin. The preponderance of opinion in the research community, certainly, is that the test is not of an acceptable scientific standard.

[57] ibid. at 104. Raskin assumes a base dishonesty rate of 0.2 in this instance.
[58] See Raskin, Honts & Kircher 'The Scientific Status of Research on Polygraph Techniques: The Case For'.

This paper is by no means the first publication to draw attention to the lack of scientific credibility of the entire polygraph enterprise. We have cited many research papers and experiments in this article where this has been demonstrated emphatically. It has also been noted by a number of CCMA commissioners, and by the court in *Mahlangu v CIM Deltak,* but the message has gone unheeded. In fact, the use of the polygraph in employment settings has increased, and it seems that employers are unlikely to voluntarily cease using the polygraph, just as they were loathe to in the USA in the 1970's and 1980's. It also seems to us from an examination of 76 CCMA reports that it is difficult to regulate the abuse of the test through regular channels - while some CCMA commissioners approach polygraph-based evidence with great circumspection and thoroughness, many do not, and on several occasions some have allowed admission of evidence from so-called expert-witness polygraphers that is decidedly inaccurate and misleading.

The Employee Polygraph Protection Act was introduced in 1988 to provide statutory protection in the United States against the abuses of polygraph testing. It is not within the ambit of this review to provide a thorough legal argument in favour of similar legislation in South Africa, but we believe that this may be the remedy of choice. Also, the introduction of a similar act in South Africa may be the most efficient way to counter the misuse of polygraph testing. Scientific research has shown that the polygraph test has a near 50% false-positive rate, and this means that many employees have been - and will be - falsely accused and dismissed on the basis of polygraph testing.