Forensic Psychophysiology:
Detecting espionage with the Guilty Knowledge Test

Ashcroft, K. R.[1], Rill, H.-G.[2], Gamer, M.[2], Gödert, H. W.[2], Vossel, G.[2], & Busuttil, A.[1]

[1] Forensic Medicine Section, Department of Pathology, University of Edinburgh, Scotland
[2] Interdisciplinary Research Group Forensic Psychophysiology, Department of Psychology, University of Mainz, Germany

Introduction

Forensic Psychophysiology
- **Classification of individuals** as either deceptive (guilty) or truthful (innocent) on the basis of differential autonomic responses to crime-related and comparison questions.

Detection of Deception: Test for Espionage and Sabotage (TES)
1. **Relevant Questions (RQs):** directly address the crime under investigation (e.g. "Have you provided secret information to an unauthorised person?").
2. **Directed Lie Comparison Questions (DLCQs):** Require the negation of behaviors that everyone has done in their life (e.g. "Have you ever told a lie?").
   Individuals are instructed to lie and to think about particular situations when they committed the acts embodied in the DLCQs.
3. **Irrelevant Questions:** neutral (e.g., "Is your name …?").

➤ **Assumptions**
   - **Guilty:** stronger responses to RQs than to CQs (RQs > DLCQs).
   - **Innocent:** stronger responses to CQs than to RQs (DLCQs > RQs).

➤ **Problem**
   - **Emotional significance of DLCQs is doubtful:** Can DLCQs reliably produce larger responses in innocent subjects?

Detection of Concealed Information: Guilty Knowledge Test (GKT)
➤ **Multiple-choice questions:** asking crime-specific details of the investigation
   (e.g. "The murder took place in ...").
   1. **Relevant Items (RIs):** crime-related (e.g. "a hotel?").
   2. **Irrelevant Items (IIs):** similar and plausible, but not crime-related alternatives (e.g. "a service station?, a store?, a house? etc.").

➤ **Assumptions**
   - **Subjects who possess "guilty knowledge":** recognise crime-relevant information and react more strongly to RIs than IIs.
   - **Subjects without "guilty knowledge":** no differential reactions to RIs or IIs.

➤ **Advantage**
   - **Low risk for innocents** (without "guilty knowledge") to react systemically more strongly to relevant item and to misclassified as "guilty".

➤ **Standard Parameters of the GKT**
   - **Skin conductance responses** (SCR): Overall, SCR-magnitudes are larger to RIs than to IIs for guilty subjects; This difference has been replicated by many studies (cf. MacLaren, 2001).
   - **Respiration**: Respiration Line Length (RLL) tends to be smaller to RIs than to IIs for guilty subjects.

➤ **Typical Application of the GKT (e.g. in Japan and Israel)**
   - Crimes like **theft** or **violence** against persons.

Aims of this Study
- Examination of the utility of the **GKT** in a mock **espionage crime** scenario.
- **Comparison** between **laboratory equipment** measuring skin conductance responses (SCRs) and **Stoelting Computerized Polygraph System** (CPS) measuring SCRs as well as thoracic and abdominal respiration.

Method

Participants
- $N = 56$ **students** (31 female; age: $M = 23.5$, $SD = 3.42$) were randomly assigned to one of following groups:

1. **Guilty subjects** ($n = 28$)
   - In order to make the **mock espionage** ecologically valid for the students, they had to gather information about an examination paper from a professor's office.
   - **Guilty subjects** were instructed to: go to the **professor**'s office, take the **key no. 8** out of a **leather jacket** pocket, unlock and open a desk drawer that was covered by an **executive case**, open a **yellow file** inside the drawer, read aloud and record the examination questions with a **dictation machine**.

2. **Innocent subjects** ($n = 28$)
   - Carried out a specific instruction in the same building, but oblivious to the relevant details of the mock espionage scene.

Guilty Knowledge Test (GKT)
- **6 multiple choice questions** seeking knowledge of: last name of professor, key number, kind of jacket, colour of the file, dictation machine and executive case (cf. Table 1).
- **1 buffer item** (the first in each sequence), **1 relevant item (RI)** and **5 irrelevant items (IIs)**.
- **Stimuli** were presented as **pre-recorded audio samples**, the inter-stimulus-interval was 22 s.
- Subjects were instructed to **deny** any knowledge of the items.

Table 1. Example of a GKT question

What is the professor's last name?					
a) Kressel	b) Rieder	c) Strobel	d) Kiefer	e) Weber	f) **Bender**

Measurement

1. **Laboratory equipment:**
   - **SCR:** Amplitude of highest SCR (in µS) within a latency window of 1-10 s following question onset.
2. **Stoelting Computerized Polygraph System:**
   - **SCR:** Difference between the lowest and the highest value (in µS) within a time window of 0.5-15 s following question onset.
   - **Respiration:** Respiration line length for thoracic and abdominal respiration within a time window of 0-10 s following question onset.

Table 2. Lykken-Scoring (cf. MacLaren, 2001) to classify each subject as guilty or innocent

2 Pts.	Response to the **RI** is the **largest** in the inspected multiple-choice block.
1 Pt.	Response to the **RI** is the **second largest** in the inspected multiple-choice block.
0 Pts.	At least two of the responses to the IIs are higher then the response to the RI.

Note: Classification as guilty if the sum is greater or equal to 6 points.

Results

SCR-Magnitudes
- **Significant interaction between** subjects' guilt status and item type (cf. Figure 1).
- **Guilty subjects** showed **stronger reactions to RIs than to IIs**.
- **No significant differences** were found **for the innocents**.
- **Hit rates:** see Table 3.

Table 3. Hit rates based on Lykken-Scoring of SCR-amplitudes

Sensitivity[1]:	75.0%
Specificity[2]:	85.7%
Overall:	80.4%

[1] hit rate for guilty subjects
[2] hit rate for innocent subjects

Figure 1. SCR-magnitudes: interaction between guilt status and item type.

CPS: Optimised scoring algorithm
- Research indicated, that the diagnostic value of the Lykken-Scoring is limited due to the neglect of the **absolute differences in the physiological responses**. Therefore an optimised scoring algorithm using the response differences between RIs and IIs was implemented.
- Several **analyses of discrimiance** were computed to estimate the contribution of every psychophysiological parameters to the diagnosis (see Table 4).

	SCR	Thoracic RLL	Abdominal RLL	Total RLL	SCR and total RLL
Sensitivity:	78.6%	67.9%	64.3%	78.6%	89.3%
Specificity:	96.4%	64.3%	78.6%	71.4%	96.4%
Overall:	87.5%	66.1%	71.4%	75.0%	92.9%

- Overall, SCR and total RLL (mean of thoracic and abdominal RLL) discriminated best between guilty and innocent subjects.
- **Cross validation** of this discriminant function (leave one out method) yielded an overall hit rate of 91.1% (sensitivity: 89.3%, specificity: 92.9%).

Discussion

✓ Both, SCR and CPS, showed a **significant differentiation between guilty and innocent subjects**.
✓ Even the simple technique of the Lykken-Scoring on the SCRs achieved an overall hit rate above 80%.
✓ An optimised scoring algorithm based on differences in the raw scores of SCR and RLL yielded an impressive overall hit rate above 90%.
✓ Results confirm the status of the **GKT** as **useful diagnostic tool** in different situations. The findings support the utility of the GKT for identifying espionage suspects.

Tasks of future research
- **Cross validation** of the computed discriminant function.
- **Field studies** in espionage crime scenarios using the GKT.

Reference

MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology, 86,* 674-683.