

# Evaluating Security

## Dublin Airport Behaviour Detection Trial

**Aoife O'Grady**  
**Economic and Financial**  
**Evaluation Unit**



**An Roinn Iompair**  
**Turasóireachta agus Spóirt**

Department of Transport,  
Tourism and Sport



Irish Government Economic & Evaluation Service

# Aviation Security Evaluation

---

**Positives:** recognition from policy team that formal, scientific evaluation of new initiative was needed

---

to demonstrate success and get buy-in

---

**Negatives:** programme had initiated 12 months previously

---

analysts therefore not in from beginning

---

**Luckily...** team developed excellent monitoring data as pilot operation and therefore aware of need for evidence of success

---

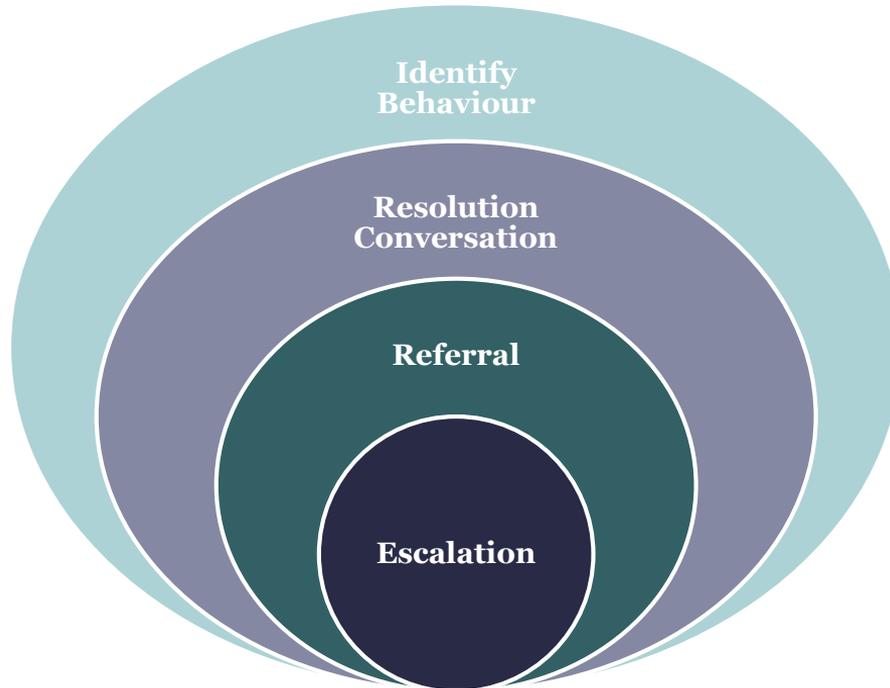
App developed after initial months to monitor activity and outcomes

---

What is

Behaviour Detection?

# BD Operation Methodology



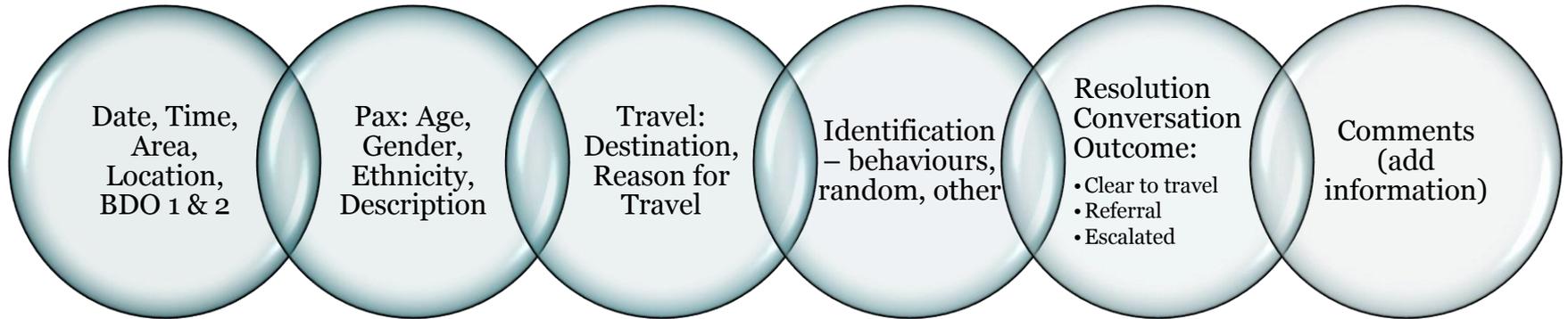
# BD Programme Contents



# Programme Objectives

Above Baseline	Human Factors	Complementary
Deterrent Effect	BD for aviation Security	Risk Basis
Detection?	Unpredictable	Wider Security Culture

# data gathering: app



# Programme Validation & Evaluation

The title is positioned above a series of horizontal lines that span the width of the slide. These lines include a thick teal bar, followed by several thinner white and teal lines, creating a modern, layered effect.

# Evaluation Questions

Does BDP deliver risk based security?

- Is programme comprehensive?
- Programme design
- Desk based analysis

Does BDP enhance unpredictability?

- Comparison to base systems
- Analysis of deterrence capability
- Desk based & stats analysis

Complement baseline standards?

- How does it work alongside existing security? Additionality?
- Focus groups/ surveys, stats analysis

Development of positive security culture?

- Evidence of cross-organisational co-operation
- Survey/ focus groups

Consistent application & success rates?

- Do patterns of implementation and detection vary?
- Stats analysis

Does BDP reduce risk of terrorist attack?

- Evidence of threat identification and resolution
- Evidence of deterrence
- Statistical analysis

Is BDP consistent across contexts?

- Does Dublin trial deliver same levels of identification/detection?
- Does the programme work equally across contexts?
- Stats comparison between Dublin and other trials

# Evaluation Questions

Does BDP deliver risk based security?

- Is programme comprehensive?
- Programme design
- Desk based analysis

Does BDP enhance unpredictability?

- Comparison to base systems
- Analysis of deterrence capability
- Desk based & stats analysis

Complement baseline standards?

- How does it work alongside existing security? Additionality?
- Focus groups/ surveys, stats analysis

Development of positive security culture?

- Evidence of cross-organisational co-operation
- Survey/ focus groups

Consistent application & success rates?

- Do patterns of implementation and detection vary?
- Stats analysis

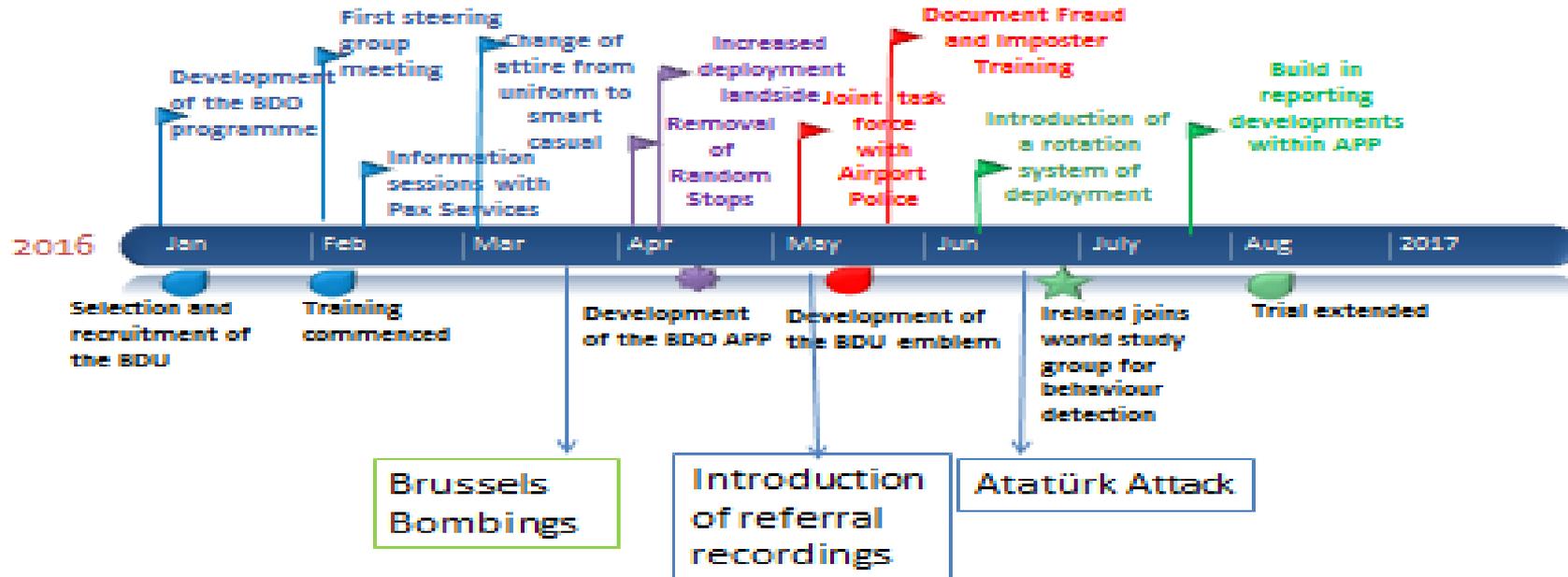
Does BDP reduce risk of terrorist attack?

- Evidence of threat identification and resolution
- Evidence of deterrence
- Statistical analysis

Is BDP consistent across contexts?

- Does Dublin trial deliver same levels of identification/detection?
- Does the programme work equally across contexts?
- Stats comparison between Dublin and other trials

# Deployment - changing landscape



# BD trial dataset

10 February 2016 –  
15 February 2017

6,378 passenger stops

- 458 randomly selected
- 5,920 based on behavioural or other indicators

13,494 behavioural  
indicators recorded

# Questions for stats analysis

- Is Behavioural Detection effective compared to random selection?
- Behavioural Indicators
  - How useful are the different behavioural indicators?
  - Are any combinations of indicators more common and/or more useful than others?
  - Does sensitivity analysis support the use of behavioural detection?
- Behavioural Detection Officers
  - Do BDOs vary in their effectiveness?
  - Does BDO performance change over time and, if so, does it improve or deteriorate?
  - Does the selection of behavioural indicators change over time and is this related to change in BDO performance?
  - Does evidence for changes and differences in BDO performance suggest need for refresher training?
  - Are any moderator variables associated with BDO performance?
- Are there improvements that can be made to the way data are collected?

# Initial Findings: outcomes

	Conversation Resolved	Escalated	Random Selection	Total
Frequencies	5,776	144	458	6,378

	AGS	Airport Police	Customs	INIS	Team Leads	Total
Frequencies	7	48	80	3	19	157

# BD vs random selection

	Clear to Travel	Escalated	Total
Conversation Resolved	5,762	14	5,776
Escalated	15	129	144
Random Selection	458	0	458
Total	6,235	143	6,378

- 143 of the 5,920 behaviour detection stops escalated
- none of the random stops escalated

(statistically significant but overall rate of escalation quite low - at 2.4%)

# Programme design and content: how accurate are the indicators?

10 behavioural indicators > statistically significant predictors of outcomes.

- best two predictors both only occasionally observed.

Most of the indicators rarely observed are non-significant

However, several indicators observed frequently are also non-significant

Total no. indicators seems to be more important predictor than which indicators

# No. indicators detected & outcomes

No. of Indicators recorded				
	1	2	3	Total
Clear to Travel	354 (6.3%)	3,023 (53.7%)	2,257 (40%)	5,634
Escalated	6 (4.9%)	34 (27.6%)	83 (67.5%)	123
Total	360	3,057	2,340	5,757

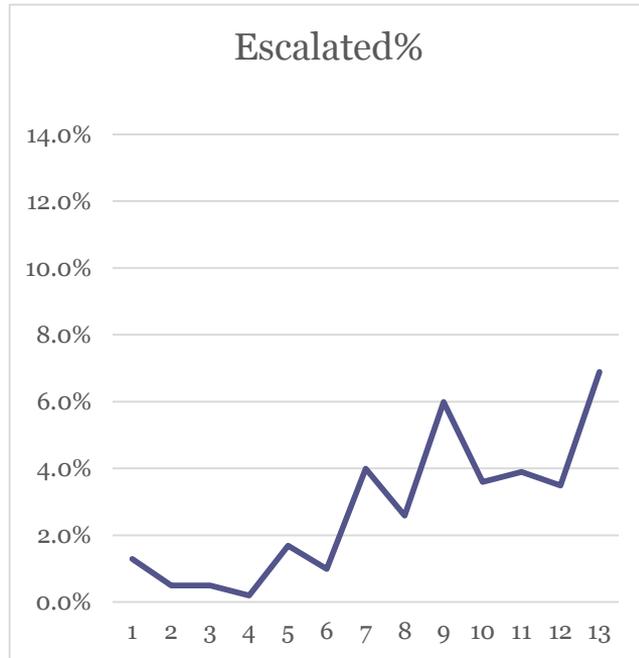
Note: work ongoing on “what happened next”

# BDO trends

Some marked differences  
between BDOs in the  
likelihood of them  
observing behavioural  
indicators

Also significant differences  
in the escalation rates  
associated with  
different BDOs

# Accuracy improves as trial beds down



- BDOs may become more selective about who they choose to stop
  - ❑ more stops made in first 2 months
  - ❑ number of stops continues to decrease
  - ❑ Lower % stopped of those whose behaviour will settle
- BDOs identify the more predictive indicators more & the less predictive indicators less

# Issues outstanding



Effectiveness of  
BD methodology  
identified

- But only to a point...
- Internally validated

How do you  
demonstrate  
effectiveness in  
identifying  
security threats?

- When these are so rare...



# Evaluation Questions

Does BDP deliver risk based security?

- Is programme comprehensive?
- Programme design
- Desk based analysis

Does BDP enhance unpredictability?

- Comparison to base systems
- Analysis of deterrence capability
- Desk based & stats analysis

Complement baseline standards?

- How does it work alongside existing security? Additionality?
- Focus groups/ surveys, stats analysis

Development of positive security culture?

- Evidence of cross-organisational co-operation
- Survey/ focus groups

Consistent application & success rates?

- Do patterns of implementation and detection vary?
- Stats analysis

Does BDP reduce risk of terrorist attack?

- Evidence of threat identification and resolution
- Evidence of deterrence
- Statistical analysis

Is BDP consistent across contexts?

- Does Dublin trial deliver same levels of identification/detection?
- Does the programme work equally across contexts?
- Stats comparison between Dublin and other trials

# Validation of BD programme design

Internal Validity vs External Validity

Most research into BD classified, no open publication, peer review

Potential to tap into research in other jurisdictions

- identifying behaviour at point of action

Evolving evidence base, need to keep validating

Comparative analysis and international collaboration key