A VIEW FROM THE INSIDE OF SYSTEM FAILURE: IS RISK MANAGEMENT THE ANTIDOTE TO SOD'S LAW?

National Curriculum Test delivery 2008 – a case study

Mick Walker

Cambridge Assessment Conference 2012

Presentation outline

- Introduction
    - Terms of reference
- Background
- 2008
    - Delivery arrangements
    - Outcomes
- 2009 onwards – lessons applied
- Implications for the future
- Conclusion

Introduction

Terms of Reference - Sod's Law (1)

- **Sod's First Law**
  When a person attempts a task, he or she will be thwarted in that task by the unconscious intervention of some other presence (animate or inanimate).

- **Sod's Second Law**
  Sooner or later, the worst possible set of circumstances is bound to occur one way or another.

- **Sod's Other Law**
  The degree of failure is in direct proportion to the effort expended and to the need for success.

Corollary –

Any system must be designed to withstand the worst possible set of circumstances.

Source - Wikipedia

Terms of Reference - Sod's Law (2)

Sod's law [uncountable]

•British English - *the natural tendency for things to go wrong whenever possible - used humorously*

Longman Dictionary of contemporary English

Mathematically, it is expressed as:

•1/8Ti x I 3/8 + O + Sb

P + M

•Task Importance (Ti)
•Inconvenience, and financial and emotional cost of task not going to plan (I)
•Optimism - the tendency to think everything will work out fine (O)
•Background Personal Stress Levels (Sb)
•Extent of Planning (P)
•Memory - especially for things that worked out well
•Using a scale of 1-5 for each factor, the minimum chance of Sod's Law striking is a score of 0.3 and the maximum is a score of 17.5.

Dr Cliff Arnall – Cardiff University

Terms of Reference - Risk Management (1)

- **Risk management** is the identification, assessment, and prioritization of risks (defined in ISO 3100 as *the effect of uncertainty on objectives*, whether positive or negative) followed by coordinated and economical application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events.

Hubbard, Douglas (2009). *The Failure of Risk Management: Why It's Broken and How to Fix It.* John Wiley & Sons. p. 46.

Terms of Reference - Risk Management (2)

Organisations have to take some risks and they have to avoid others. The big question that all organisations have to ask themselves is: just what does successful performance look like? This question might be easier to answer for a listed company than for a government department, but can usefully be asked by boards in all sectors.

The Institute of Risk Management. (2012) *Risk Appetite and Tolerance*

Risk appetite should be developed in the context of an organisation's **risk management capability**, which is a function of **risk capacity** and **risk management maturity**. Risk management remains an emerging discipline and some organisations, irrespective of size or complexity, do it much better than others. This is in part due to their risk management culture (a subset of the overall culture), partly due to their systems and processes, and partly due to the nature of their business. However, until an organisation has a clear view of both its risk capacity and its risk management maturity it cannot be clear as to what approach would work or how it should be implemented.

The Institute of Risk Management. (2012) *Risk Appetite and Tolerance*

Background

Background

- The National Assessment Agency (NAA)
- The delivery business model

  Procurement

  Contracts

  Scale and timing
- The supplier market up to 2008
- The role of National Curriculum tests

The role of National Curriculum Tests (1)

- A statutory requirement at Key Stage 2

- Introduced in the Education Reform Act (ERA) 1988

'……the arrangements for assessing pupils at or near the end of each key stage for the purpose of ascertaining what they have achieved in relation to the attainment targets for that key stage.'

ERA (1988) Ch1. *The Curriculum.* 2. (2).(a).

- Not a qualification.

The role of National Curriculum Tests (2)

Issued 1951 to 1961

21st Century

Newton's line of purposes

Available in 2012

The role of National Curriculum Tests (3)
Performance tables - Primary KS-2 changes over time

- 2012   Changes brought in following the Bew Review
  - Writing subject to teacher assessment only
  - National sample of writing test – as an estimate of national attainment
  - Overall attainment in English based on reading tests and writing teacher assessment
- 2010   White Paper – *The Importance of Teaching* (proposed future changes):
  - Removal of CVA measure
- 2010
  - Change to exclude science results for the first time as tests no longer mandatory.
  - Tests boycotted by a quarter of schools
  - Replacement of KS2 Aggregate Score year-on-year comparison with two four year time series showing year-on-year comparisons of Average Points Scores and of the proportion of pupils achieving level 4 or above in both English & maths
  - Re-introduction of publication of KS2 Teacher Assessment data
- 2009:

  Progress Measures introduced
- 2008:

  Introduction of new Key Stage 2 and Key Stage 3 combined English and maths indicators
- 2007:

  Introduction of Contextual Value Added (CVA) measure
- 2006:

  Introduction of year on year comparisons of level 5 attainment
  New pilot of KS2 Value-added measures
- 2002:

  Pilot of KS2 Value-added measures
- 1996:

  Primary schools tables introduced

2008 Delivery Arrangements

2008 Delivery Arrangements

QCA submitted to the Sutherland Inquiry that in its view, the fundamentals of test delivery remained unchanged in 2008:

"The delivery and marking of National Curriculum tests has not changed in any significant way for more than a decade. Over nine million test papers move around the country in vans…..The only development of any importance in test delivery in the past decade has been increased security: script bundles are now bar-coded and receipted, not left on doorsteps and post office counters."

QCA submission to the Sutherland Inquiry, 12 September 2008

National curriculum test processes: Top level view
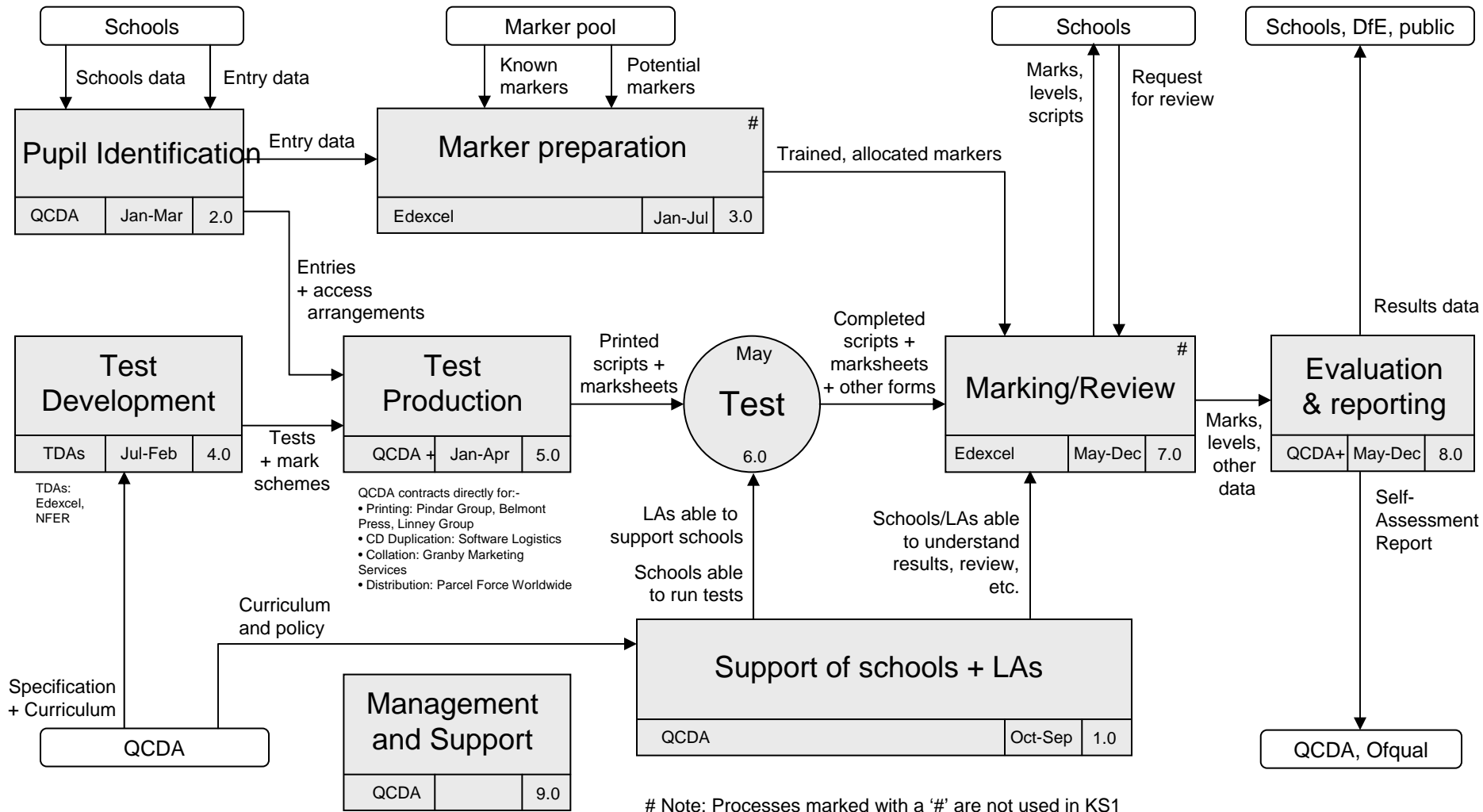
2008 - Risk registers in place….

Risk ID	Exec risk ref.	Risk description [colour code text RAG]	Risk owner	Likelihood/ Impact (Significance)	Existing controls	Future controls if necessary (inc point to escalate to Executive)	Review date
EXAMS 10	2/3	New recruitment campaign (November 2004) fails to attract new applicants in sufficient number to support summer 2005 session	Mick Walker	3 x 4 [12]	• Weekly monitoring • Focus groups • MORI poll of last year's campaign • Targeted recruitment of shortage subjects • Publicity re advantages of being an examiner • Retention strategy	• Work with The Team & Abs to improve retention strategies • Institute of Educational Assessors to raise professional status • Monitor daily	Weekly

(NAA 2005)

2008 Delivery Arrangements

ETS Europe (ETS) proposed a number of innovative elements that could be trialled for introduction between 2008-12.

These were:

- **Central distribution model (warehouse and tracking system to manage movement of scripts to and from schools and markers)**
- Onscreen marking
- **Online mark capture (submitting individual question marks online)**
- **Online training of markers**
- **Online standardisation (a quality assessment of markers' ability to apply the mark scheme fairly)**
- **Online benchmarking (as standardisation, but completed regularly during the marking process to ensure consistent quality of marking)**
- **Online attendance register**
- Development of data systems to process and present results online to schools

2008 Delivery Outcomes

2008 Delivery Outcomes
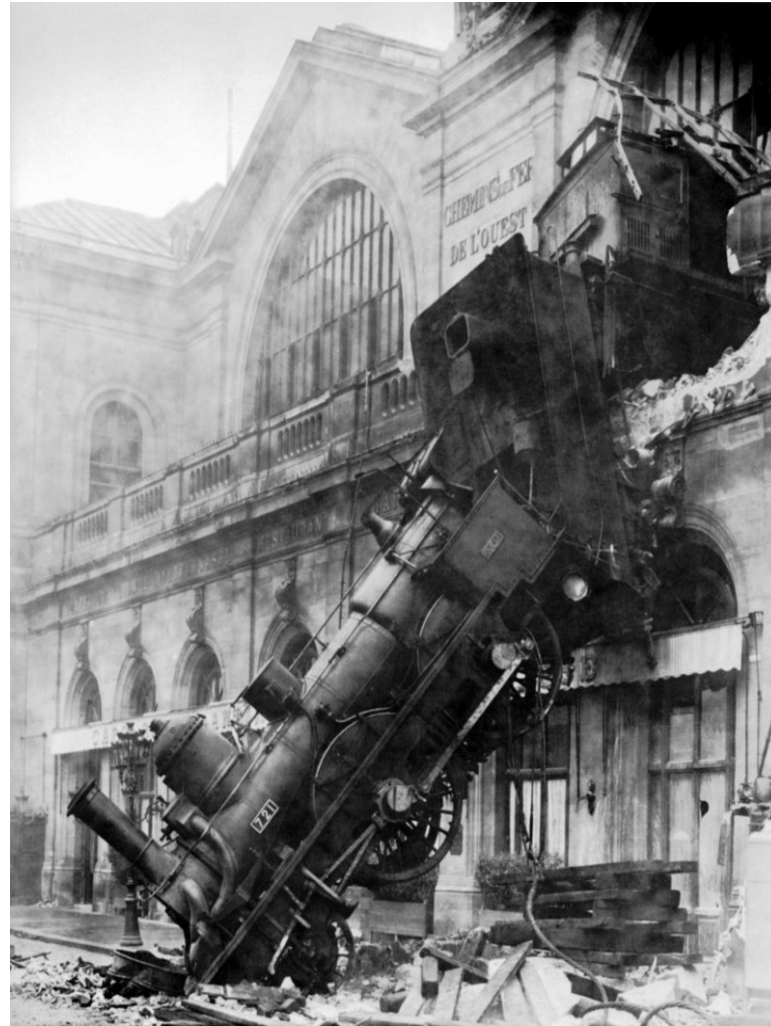
'Old Gare Montparnasse, Paris, October 23, 1895

2008 Delivery Failures

The blame game - Impact on the NAA - QCA
- Media
- Teacher Associations
- Select Committee
- Ofqual and the DCSF
- **The Sutherland Inquiry**

2008 Delivery Failures

Sutherland Inquiry (1)

Sutherland emphasised the following priorities:

1.The delivery process for National Curriculum tests should be modernised and improved, in consultation with the marking community. This should include piloting online marking, which has been used successfully for other qualifications;

2.Whatever process is used, it should be thoroughly piloted and closely project managed to ensure that schools and pupils get their results and scripts back on time; and

3.The customer service provided to markers and schools must be vastly improved to ensure that they are properly supported and able to access up-to-date information when required.

2008 Delivery Failures

Sutherland Inquiry (2)

Project and risk management: findings

1.ETS's project management was not fit-for-purpose

2.ETS failed to identify and assess risks accurately and failed to report risks to NAA transparently

3.QCA had project and risk management systems in place, but did not use these effectively

4.DCSF had good project and risk management processes, but officials may not have challenged QCA sufficiently

2008 Delivery Failures

Sutherland Inquiry (3)

Delivery of the National Curriculum tests: findings

1. The end-to-end delivery system was insufficiently tested

2. There were cumulative failures in different components and interfaces of the ETS delivery system; QCA did not make an accurate assessment of the impact of these failures

3. The 'critical path' for the project had not been identified correctly

4. The quality of the management information (MI) provided by ETS was ultimately inadequate

5. Few viable contingency options were built into the delivery system by ETS and QCA and those that were available were not put into action in a timely and appropriate way

6. ETS did not invest in its relationship with schools and markers

7. Schools and pupils have been inconvenienced by the delay in results

2008 Delivery Failures

Sutherland Inquiry (4)

Delivery of National Curriculum tests: recommendations

1.The test delivery process and timetable should be designed to allow for maximum marking time and capacity

2.Customer service must be greatly enhanced for schools and markers and a reference group should be established by the test supplier to ensure schools' and markers' views influence every stage of the process

3.The delivery process for National Curriculum tests should be modernised and improved, in consultation with the marking community. This should include piloting online marking, which has been used successfully for other qualifications

4.Full testing and piloting of the test delivery process should be integrated into the timetable, including end-to-end and user-acceptance testing

2008 Delivery Failures

Regulation: findings

1. There was insufficient clarity in the Regulator's reporting arrangements and its relationship with QCA, NAA and DCSF

2. The Regulator's risk monitoring and management processes were inadequate and the process for escalating concerns to QCA and DCSF was not properly defined

3. The focus of the Regulator's monitoring of the tests in 2008 was too narrow and insufficient attention was paid to the monitoring of delivery and systems issues

4. The Regulator was not always able to obtain accurate management information at the right time from NAA

…..the blame game can have different outcomes – G4S.

A statement from the company said the review did not hold Mr Buckles personally to blame for what went wrong and said it was in the best interests of the company and its stakeholders that he should remain in charge.

•"Whilst the [chief executive] has ultimate responsibility for the company's performance, the review did not identify significant shortcomings in his performance or serious failings attributable to him in connection with the Olympic contract."

Analysts suggested shareholders were happy for him to remain in post because of his performance over a longer period than just the last difficult six months.

•"Nick Buckles has been with the organisation for more than 25 years, [and] he has been chief executive over the last seven years," said Kean Marden, an analyst at investment bank Jefferies.

•"He has built it into the world's largest security company, the share price performance has been pretty impressive," he told BBC News.

G4S shares rose as much as 2% after it was announced that Mr Buckles would be staying on.

2009 onwards – Lessons Applied

2009 onwards – Lessons Applied

- NAA dissolved
- Stronger governance
- Clearer roles
- Refined risk management procedures
- Greater customer focus – and dialogue
- End to end testing

Risks and issues governance

High priority risks and issues are presented to the NCT Risk Committee for discussion. They are then reported to the NCT Programme Board. Risks and issues are escalated to the QCDA Executive and the QCDA Board, as appropriate. Key strategic risks are also reported to the Strategic Delivery Group.
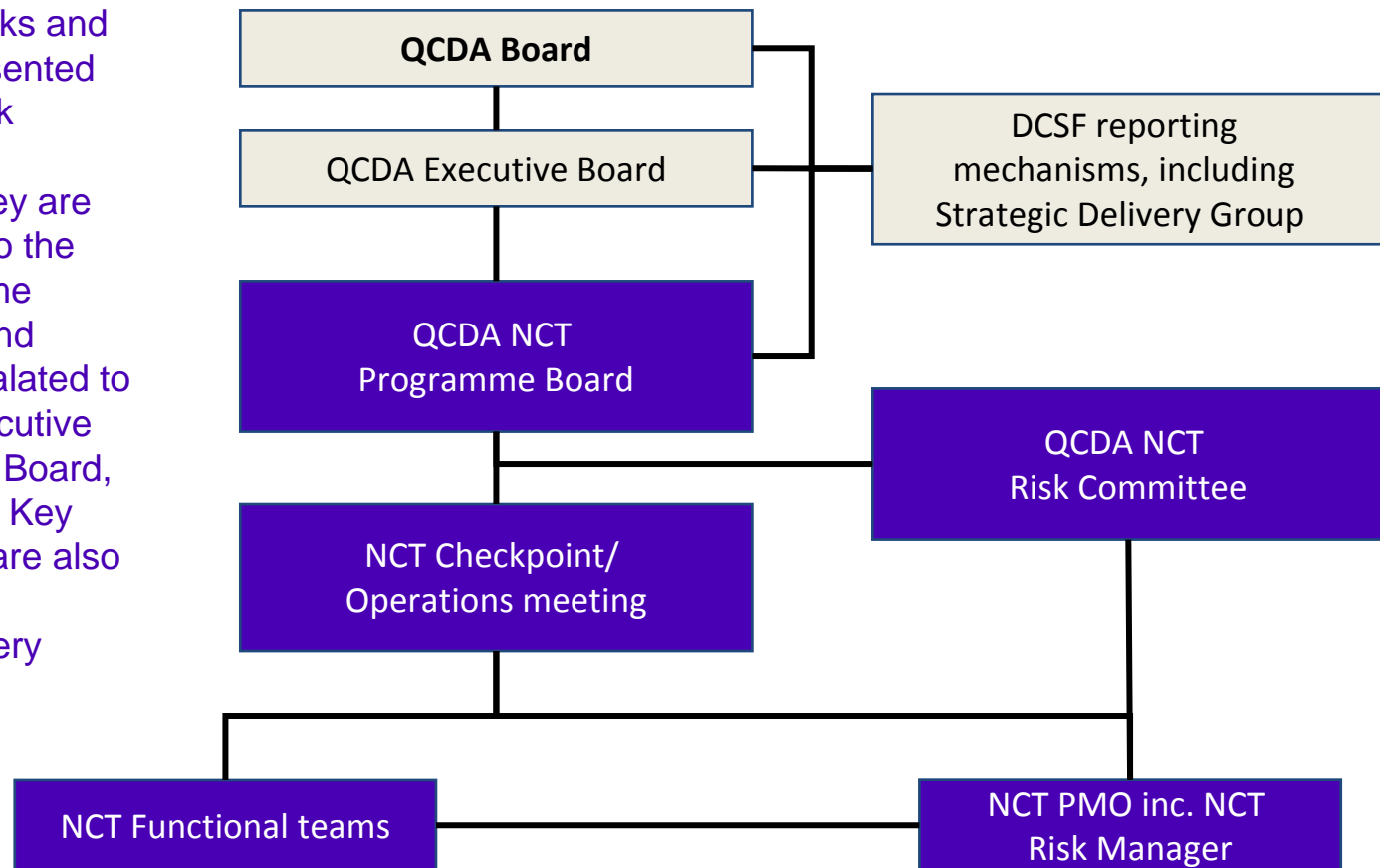
Risks and issues governance (cont.)

NCT Programme Board

Scope
Focused on the full NCT programme cycle, access to the full suite of programme risks and issues and identifies and reviews strategic matters affecting the programme. The Programme Board is the point of escalation for risks and issues from the NCT checkpoint meetings and risk committee discussions

Frequency of meetings
The Programme Board currently meets on a fortnightly basis. This becomes weekly during the higher tempo period in delivery cycle (Jan-Jul).

Attendance
Director Test and Examination Support Group (Chair), NCT Programme Director, NCT Programme Manager, QCDA Executive members, QCDA procurement and legal members (as required), DCSF Policy and Data Services Group representatives and Ofqual representatives.

NCT Risk Committee

Scope
The Risk Committee has full access to the full suite of programme risks and issues. Risks that are red-amber or red, along with issues that are amber, amber-red or red, are reported for information, discussion and action. The members of the committee also raise any risks or issues they wish to discuss and select those to be brought to the attention of the Programme Board.

Frequency of meetings
The Risk Committee currently meets fortnightly before the Programme Board meeting. This becomes weekly during the higher tempo period in delivery cycle (Jan-Jul).

Attendance
Director of Qualifications & Skills (Chair), NCT PMO Manager, QCDA Executive Members, NCT Programme Director, NCT Programme Manager, DCSF representatives, QCDA Procurement, QCDA Legal (as required), NCT Functional Leads or representatives (as required).

NCT Checkpoint / Operations Meetings

Scope
Procurement Checkpoint meetings, and delivery Operations meetings, are chaired by the programme manager for procurement programme and the key stage manager for delivery programme. Functional leads provide updates on cross cutting issues and report on key risks and issues.

Frequency of meetings
Meetings are fortnightly for the procurement programme and weekly for the delivery programme.
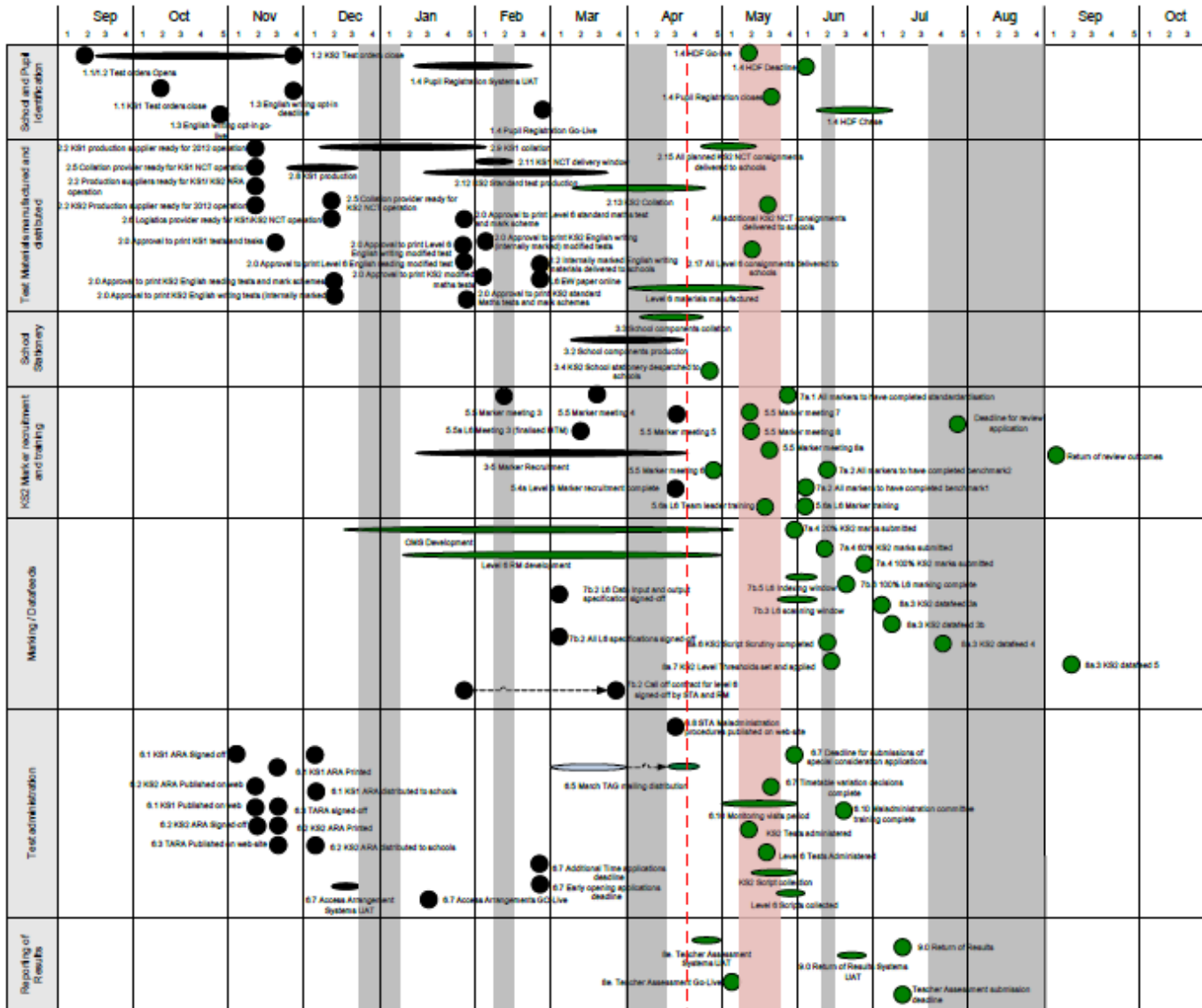
Attendance
NCT Programme Manager (procurement), Key Stage Manager (delivery), NCT PMO Manager, QCDA Head of Relationships Cluster, QCDA Delivery Director, QCDA Communications, QCDA CIOG, QCDA Procurement, QCDA Legal and DCSF Observer.

NCT KS2 English and Maths – 2012 Programme Plan on a page

2012 KS2 English and Mathematics - Risk Details							
Commercially Sensitive:		No			**Supplier:**		TSO;
Overall Status:	Closed (Approved)	**Last Modified:**	13 Jul 2012 13:51:04		**Risk ID:**	29	**Proximity:** 10 Jul 2012
Title:	Return of Results Functionality				**Risk Manager:**	IS	

Pre-mitigation	Post-mitigation	Target RAG Status
Probability: Medium, Impact: High, Severity: Amber Red	Probability: Low , Impact: High, Severity: Amber Red	Probability: Low , Impact: Medium, Severity: Amber

Risk Event:	Return of Results does not meet the final functional requirements for the 2012 cycle		
Risk Cause:	1. Policy decisions required to develop requirements are not made in good time. 2. Business processes are not scoped or agreed at the point requirements are gathered and late requirements emerge following start of development which cannot be accommodated. 3. New requirements inadequately specified. 4. Insufficient lead time to develop required changes. 5. TSO underestimate effort involved and fail to complete the work on time. 6. Insufficient systems assurance carried out on developments both by TSO and STA.	**Impact:**	Schools are unable to access their pupils results through NCA tools. There would be a large reputational impact on STA and DfE. A increase in calls to the helpline.
Predictive Risk Indicators:	1. Decisions required such as the calculation of an overall English level are delayed. 2. TSO request a lot of clarifications on the requirements specified. 3. Extended lead time in approving CRs due to uncertainty around the approvals process for IT projects. 4. TSO seek to delay the start of UAT or report overruns on their development project plan. 5. Return of results not available to schools on the 10th July 2012. Schools report significant number of issues with results on the website.	**Current Controls:**	1. TSO report development progress regularly
Update & Commentary:	08/11 - Meeting to define English writing changes for 2012 are taking place. Dec 11 - Decision regarding English overall level currently with SDG. 09/07 - System QA checks are completed - all development and testing completed. 09/02 - Draft pupil results specification produced and distributed internally. 17/04 - All requirements have been reviewed and quoted on by TSO. Contract variations are in the process of being reviewed and signed by TSO for the extension and the developments. Resources have been pre-booked by TSO. 17/05 - Project plan received from TSO showing intention to release software to UAT during w/c 12/06. Data request in place to provide data to support TSO during development. 20/06 - Development and UAT activities complete. Site released into live environment on 15/06/12 for penetration testing commencing 18/6. Data assurance activities scheduled to start 21/06. 10/07 and 12/07 - Module opened successfully being accessed by over 13000 unique schools and 22000 users in the first day. No issues reported about functionality of site.	**Linked Items:**	

Associated Actions					
Action ID	Action Title	Actionee	Action Progress Update	Due Date	Status
39	SDG English Overall Level Decision Required	Emc	13/01 - Ministerial approval on the production of an overall English level received. English levels will not be returned at a pupil level but only at school level.	16 Dec 2011	Completed

Highlighted:	No	**Escalated**	(0) Not Required		**Support Required:**	
Related Division			Related DfE Shared Service(s)		Related Workstream(s)	Related STA (Internal) Function
			Commercial / Procurement; ICT;		KS2 English and Maths; KS2 SST;	Commercial; Data and Systems;

Implications for the future

- The appetite for change in assessment & delivery
- Innovation and development
- The supplier market
- Accountability

Conclusion

- We can *minimise* risk through *realistic* mitigation and contingency planning.

- Risk management is more mature now than in 2008.

- We need an open dialogue about assessment with better education on assessment – its possibilities and limitations.

- We need a coherent approach to assessment supported by a national strategy for the design, development and delivery.

However, Sod is alive and well!