New Approaches to Validation

Alan Robinson &

Paul Glover

Defence Science and Technology Laboratory

Outline Agenda

- Why a new approach?
  - Why now?
- What overall approach ?
- What types of model ?
- How is this being implemented ?
  - Purpose
  - Transparency
  - Stages of the work
  - Common pitfalls

Alan

Paul

Discussion throughout!

Why a new approach?

Why does it matter?

- Models play a key role in most studies
  - Models come in a whole range of shapes and sizes, from human or animal experimentation through to software
- Model quality assurance (QA) ensures that:
  - The model is doing things right (verification)
  - The model is doing the right things (validation)
  - The model is appropriate to the decision being supported (fitness for purpose)
- The type and scale of QA should be proportionate to the decision being supported

When things go wrong…

- Can have major impact:
  - Economic
  - Financial
  - Legal
  - Reputational
  - Operational

- West Cost Mainline had a major impact on approaches to modelling in Government

West Coast line's two-year delay

FirstGroup vows to rebid despite franchise fiasco

The Telegraph

HOME NEWS WORLD SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFE FA

Companies | Comment | Personal Finance | Economics | Markets | Festival of Business | Your Bus

HOME » FINANCE » NEWS BY SECTOR » TRANSPORT

West Coast Main Line: scrapped bid reveals chaos at the heart of government

Scrapping the bid for the West Coast mainline has revealed chaos at the heart of government.

Even before the winning bidder was announced on Aug 15, Sir Richard had fired off a letter to Greening complaining FirstGroup's bid looked "unrealistically high" and she risked a re-run of the fiascos on the East Coast line.

First-class fiasco

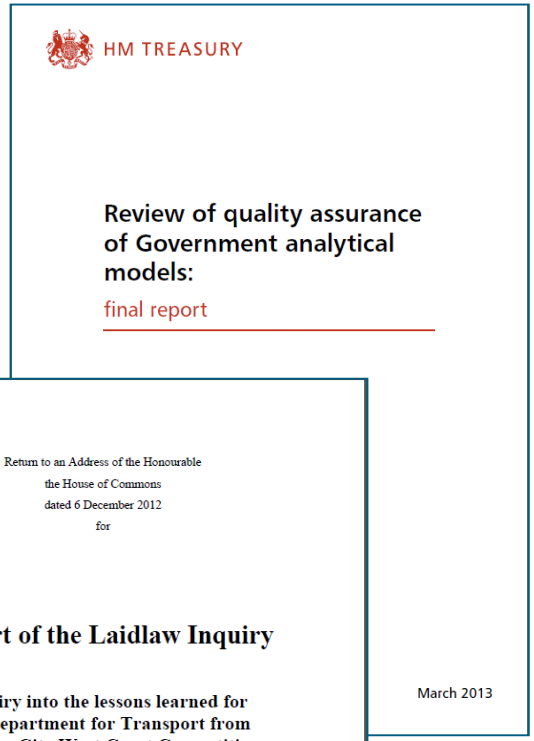
Officials blamed for failures and dodgy sums that cost Virgin the £7bn West Coast contract

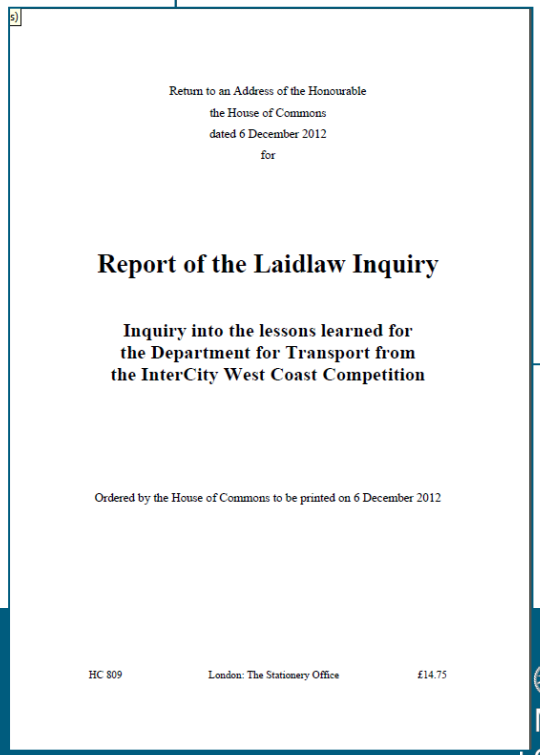
By Ray Massey

ment Secretary – should take McLoughlin, who published the

Pan-Government response

- Laidlaw Review
  - "What went wrong"
  - https://www.gov.uk/government/publications/report-of-the-laidlaw-inquiry
- MacPherson Review
  - Pan-govt implications for Business Critical models
  - https://www.gov.uk/government/publications/review-of-quality-assurance-of-government-models
  - 8 key recommendations
- Implementation underway
  - Analytical Quality Assurance (AQUA)
  - To be published later in the year



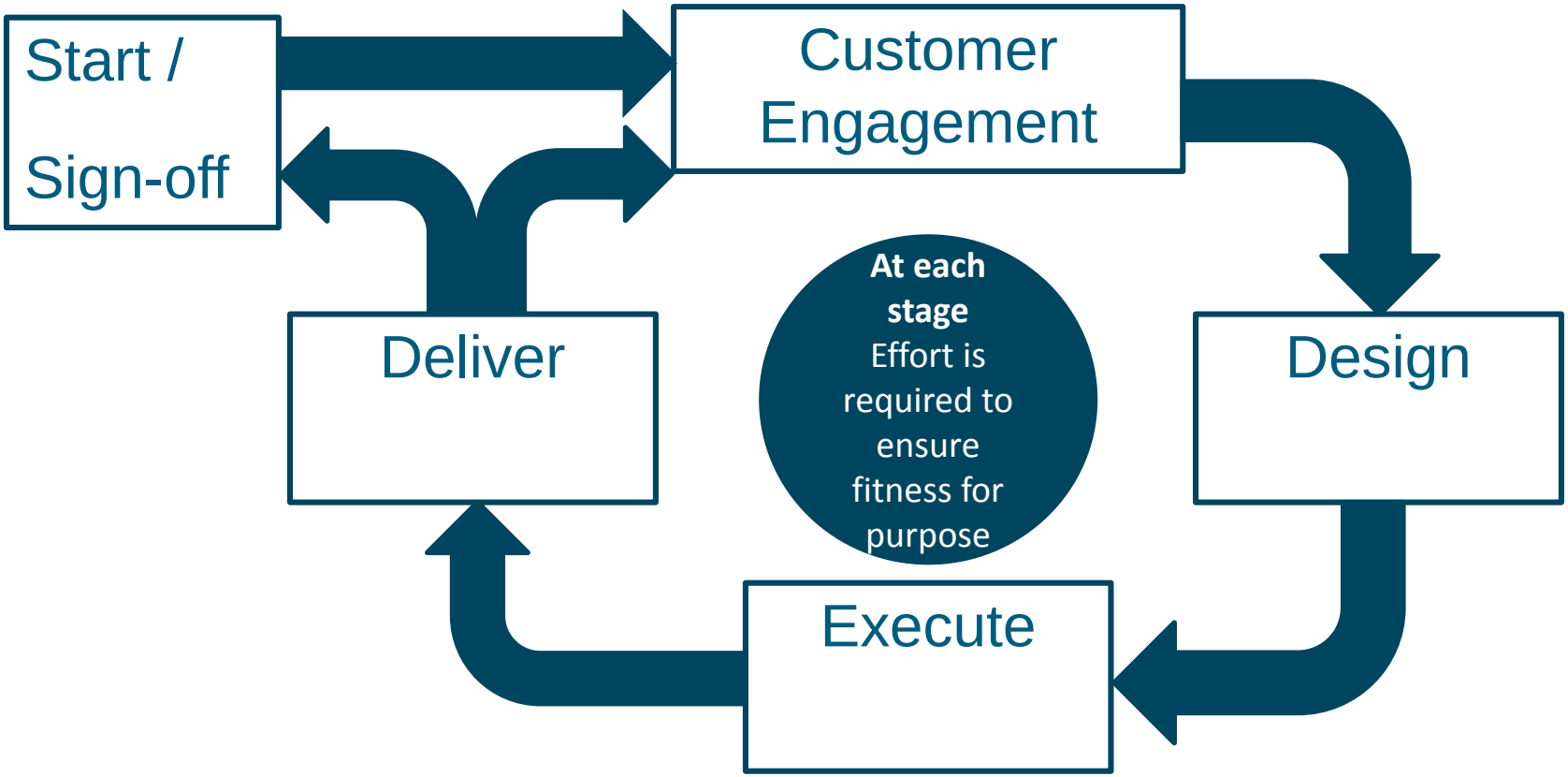


What overall approach?

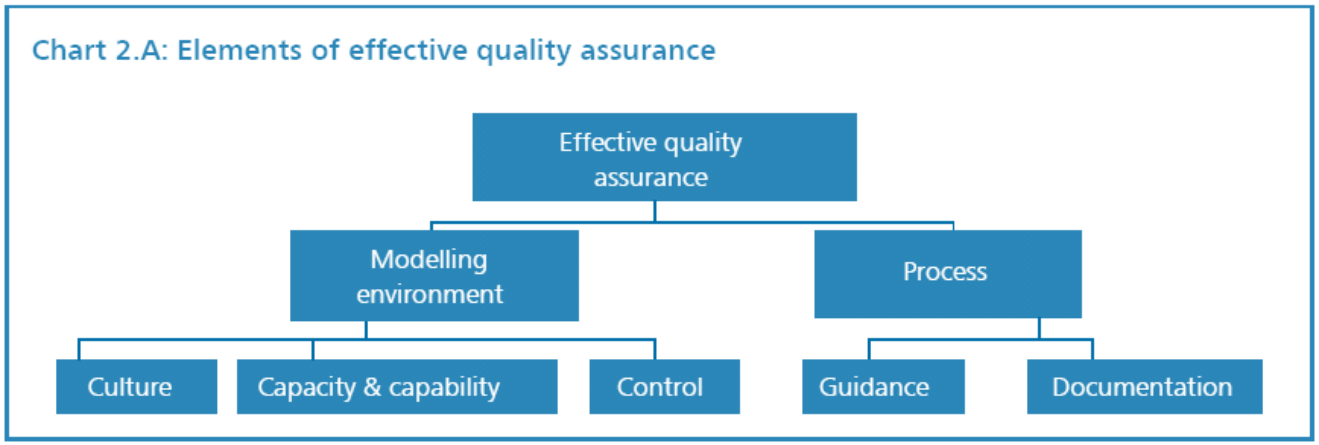
Key Points

1. Quality Assurance needs to be conducted through-life

Fitness for purpose in the analysis lifecycle

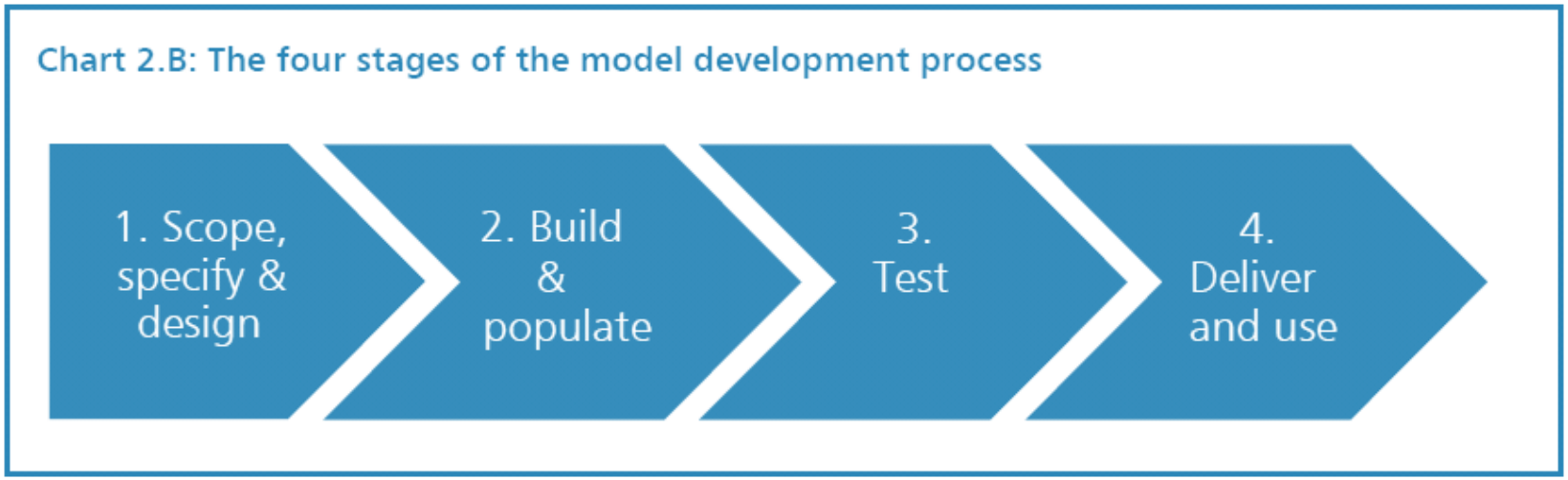


Key Elements of Quality Assurance



- *Environment: creating the conditions in which QA processes can operate effectively*
- *Process: establishing a clear process for every stage*

Models have a life-cycle



Key Points

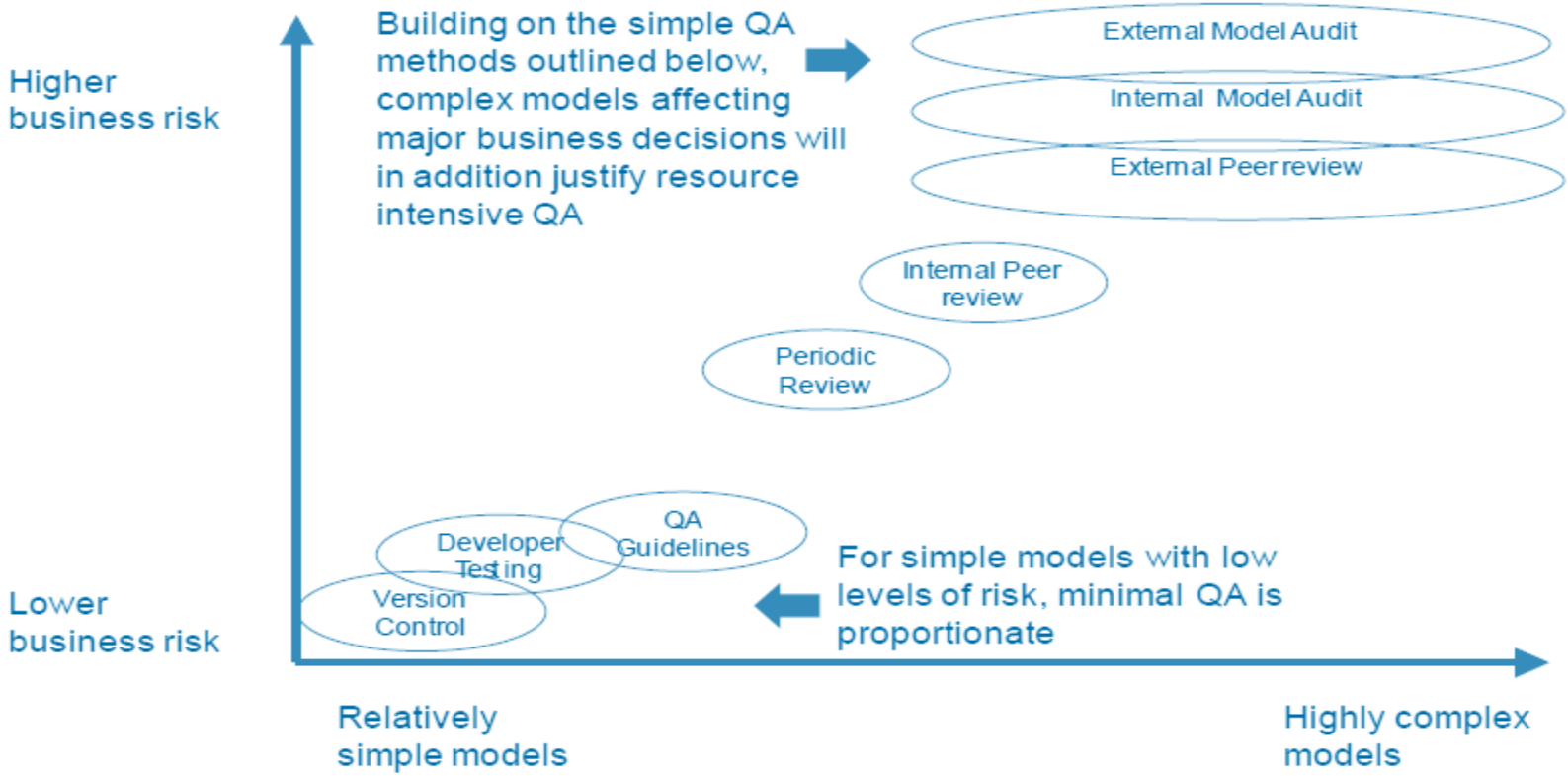
1. Quality Assurance needs to be conducted through-life
2. Quality Assurance needs to be proportionate to the decision being supported and the modelling risk being managed

Ensuring V&V is proportionate

Type and complexity of analysis	• Highly complex analysis requires more effort to assure
Novelty of the approach	• A previously untried modelling technique requires more assurance
Importance of issue	• Different issues will vary in their economic and social impact
Relevance of the analysis to the decision making process	• When a analysis forms only one component of a broad evidence base, less assurance is required than if the decision is heavily dependent on one model
Precision of the analysis outputs	• Imprecise analysis can need different QA than precise analysis, eg because of inherent limitations of the analytical technique, or lack of data on assumptions
Amount of resource available for the modelling which includes the V&V	• The value for money of any additional V&V must be balanced alongside the benefits and the risk appetite that exists

Ensuring V&V is proportionate

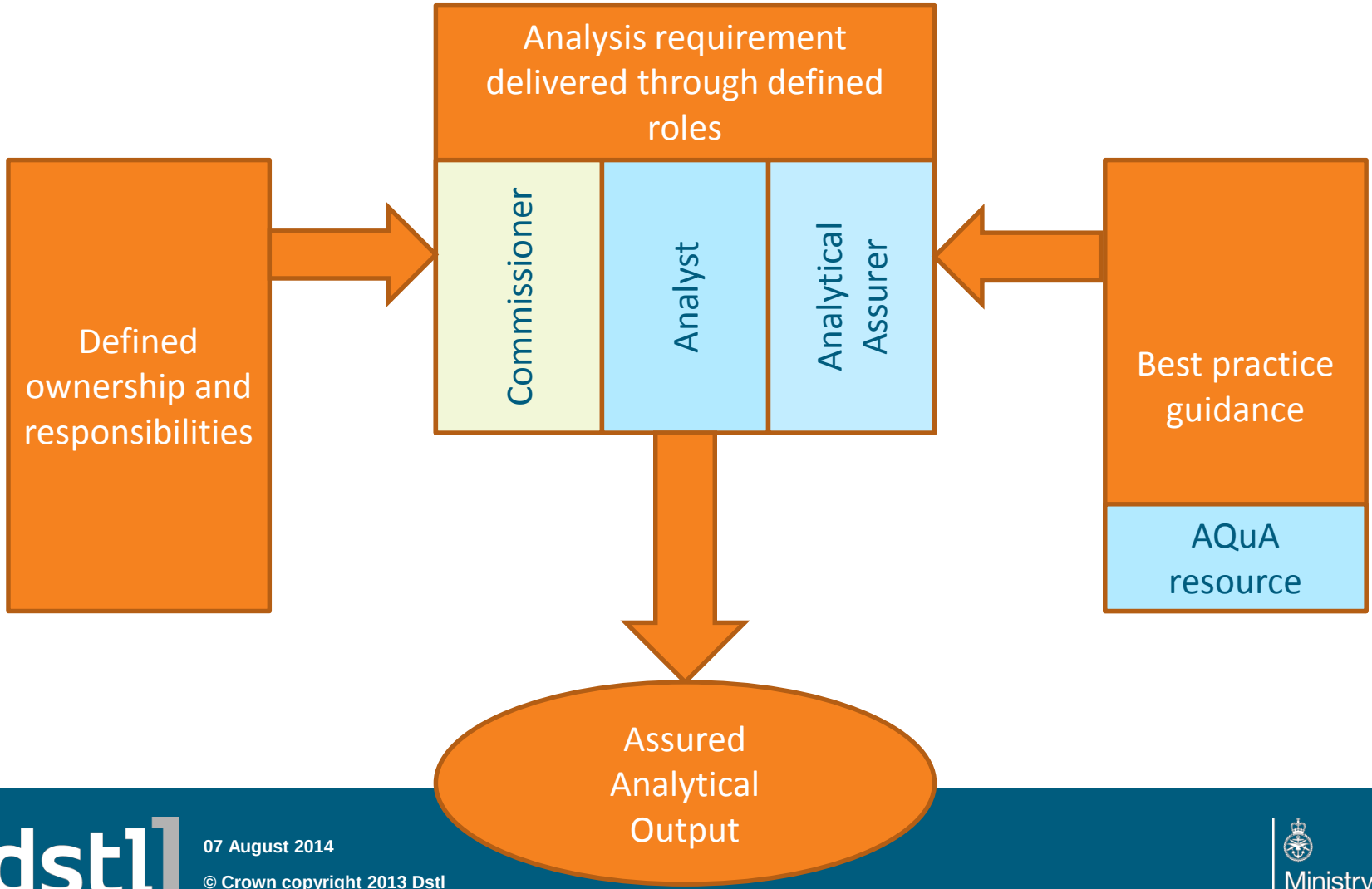
Chart 2.C: Schematic showing indicative types of QA that might be expected given different levels of risk



Key Points

1. Quality Assurance needs to be conducted through-life
2. Quality Assurance needs to be proportionate to the decision being supported and the modelling risk being managed
3. Quality Assurance requires ongoing interaction with customer and key stakeholders

Managing Study Quality



What types of model?

What types of model

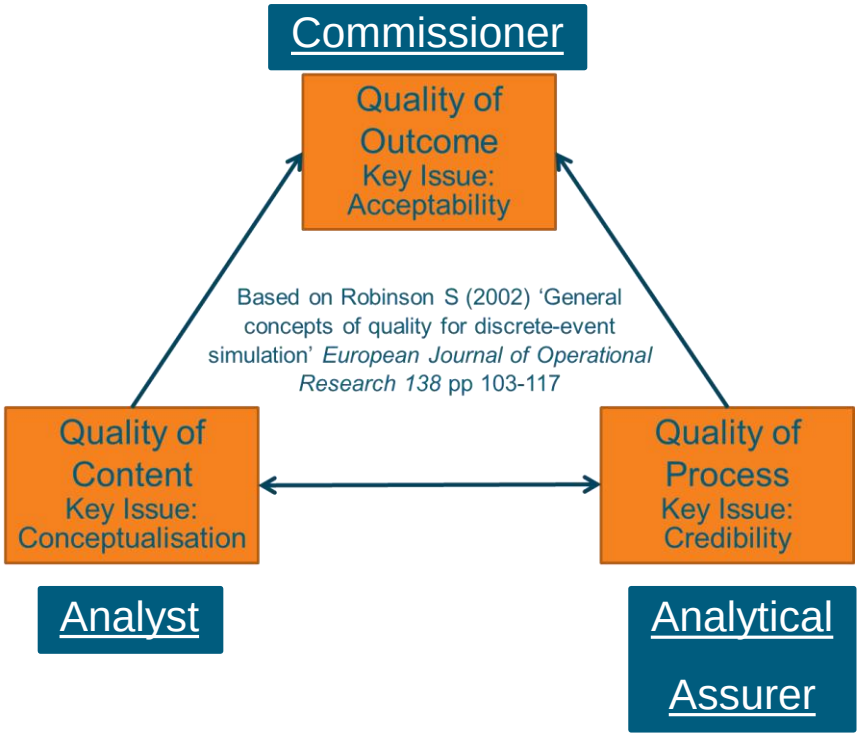
Policy Simulation	Appraisal of policy options, analysis of impact on people, finances, etc
Forecasting	Assessing the future, perhaps to provide base information for policy development or financial planning
Financial evaluation	Assessment of liability or future cost
Procurement & commercial	Evaluation of VfM or affordability and award of contracts
Planning	Planning current actions based on future forecasts
Science-based	Understanding and forecasting natural systems
Allocation	Distribution of funding across organisations responsible for service delivery

How is this being implemented ?

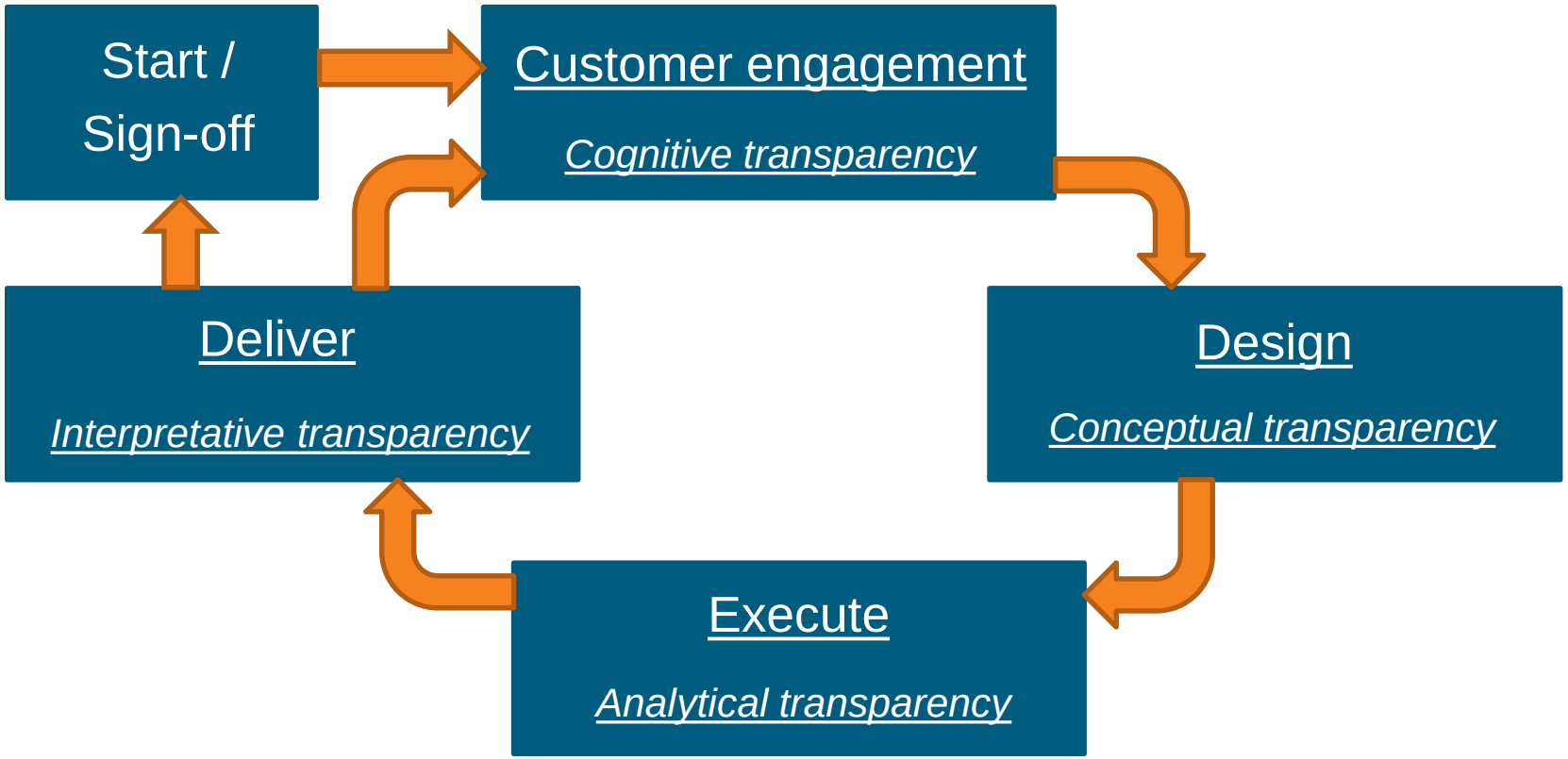
Analytical Assurance

It's all about quality

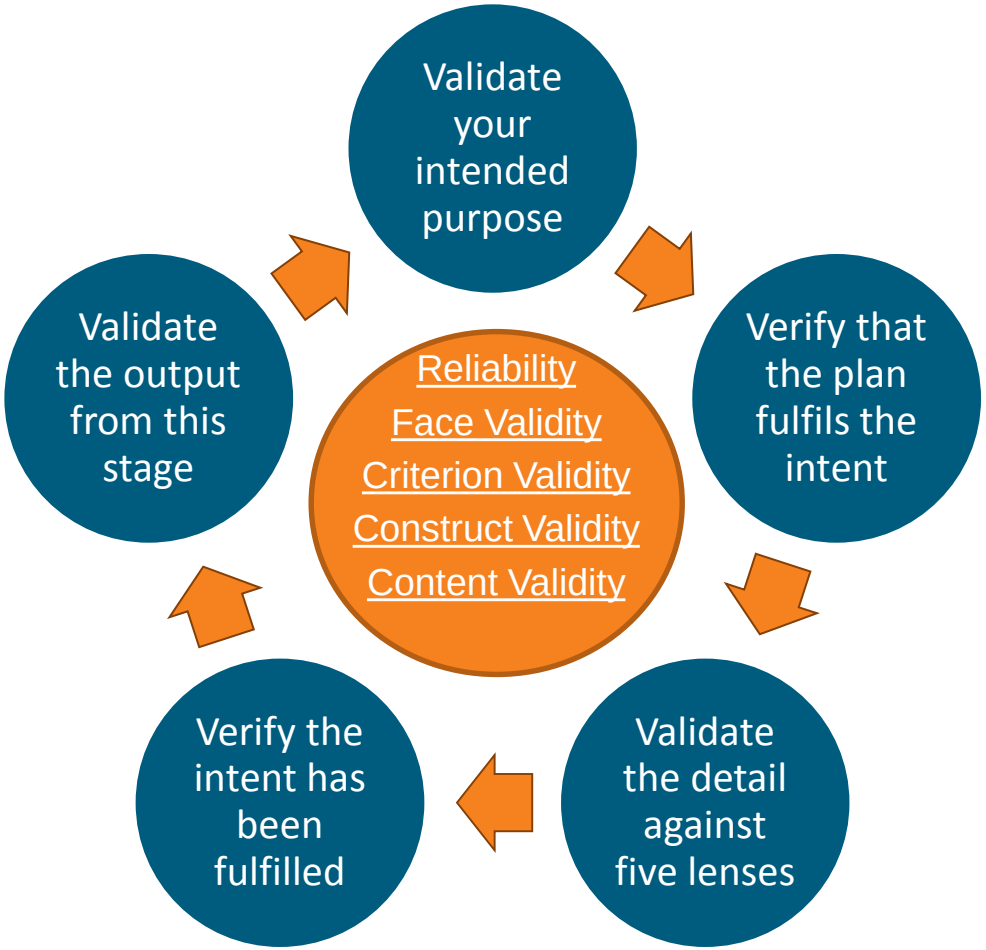
- The aim, following the MacPherson review is to provide assurance of Analytical Quality
  - Work which reliably delivers the right outcome
  - Supported by a transparent process
  - Properly founded in the evidence



Stages of the work - 1



Stages of the work - 2



- Recommended as a gated process
  - Accept and proceed
  - Accept if issues noted are resolved
  - Accept but halt
  - Reject

Common Pitfalls

- Framing the analysis – commissioner owns benefits
  - 6 key issues
    - Framing, detail, plurality, emergence, ‘facts’, ‘prediction’
- Method – analytical assurer owns approach
  - Assessing the requirements in their context
- Engagement with the work – analyst owns the detail
  - Phenomenological engagement
  - Limits of the knowable

Any Questions?

Alan Robinson
aprobinson@dstl.gov.uk
Portsdown West
Defence Science and Technology Laboratory

Paul Glover
peglover@dstl.gov.uk
Portsdown West
Defence Science and Technology Laboratory

Role Definitions

- Commissioner
  - Responsible for defining the benefits for which the work is being conducted, agreeing the analysis, understanding the product as it emerges and ensuring the work has impact
- Analytical Assurance
  - Responsible for providing evidence that the processes performed to deliver the analysis were conducted properly
- Analyst
  - Responsible for properly conducting the detail of the work

Return

Customer engagement

Activity	Commissioner	Analyst	Analytical Assurer
Ensure key aspects of the problem, scope and complexities are captured and clearly communicated.	Owner	Involved	Involved
Be available to engage with the analysts in order to appropriately shape the work.	Owner	Involved	
Clearly record the perceived purpose of the analysis and/ or modelling and the levels of quality and certainty that are required for this purpose.	Involved	Owner	Involved
Challenge and test the understanding of the problem.	Involved	Involved	Owner
Ensure appropriate resources are commissioned for the analysis.	Owner	Involved	Involved
Ensure appropriate stakeholders have been identified so that the scope and boundaries of the problem can be appropriately explored.	Owner	Involved	Involved
Explore the requirements, boundaries, and scope with all of the stakeholders ensuring a wide range of perspectives are sought.	Involved	Owner	
Challenge the requirements, boundaries and scope and assess whether sufficient views have been considered.	Involved	Involved	Owner
Ensure expectations are managed to keep stakeholders expectations aligned with what can be delivered.	Involved	Owner	Involved

Return

Design

Activity	Commissioner	Analyst	Analytical Assurer
Record and review the decision process from structuring the problem to developing the analytical plan. Consider if the process reflects due RIGOUR.	Involved	Owner	Involved
Plan appropriate resources to deliver the analysis.	Involved	Owner	Involved
Capture the specification of any necessary methods. This must be adequate to allow subsequent verification testing / validation of the analysis. It should also specify what approaches will be used to identify, quantify and communicate uncertainty.		Owner	Involved
Produce appropriate design documentation. Best practice can include a concept of analysis, user requirements, design specification, functional specification, data dictionary, and test plan.		Owner	Involved
Dry run the proposed approach to see if it delivers as intended. Then consider if the overall approach adequately addresses the complexities of the customer issue for this purpose. It is good practice to engage Subject Matter Experts in this review.	Involved	Owner	Involved
Ensure the accuracy and limitations of the chosen methods are understood – and where appropriate tested (where possible base lining their response against independent reference cases).		Owner	Involved
Ensure the basis of the work is accurate, transparent (so that the basis of the findings can be understood) and well recorded.		Owner	Involved
Ensure the approach to the analysis is well-structured for the purpose, data driven, and reflects a robust overall design.		Owner	Involved
Ensure the level of quality checking of the analysis will be appropriate for the decision being supported.	Involved	Involved	Owner
Ensure that, if required, formal ethical approval is provided.	Involved	Owner	Involved

Return

Execute

Activity	Commissioner	Analyst	Analytical Assurer
Collect and manage data. Understand data accuracy and uncertainties. Capture, manage and understand implicit assumptions made.		Owner	Involved
Engage appropriate Subject Matter Experts, at the appropriate time, when collecting data. NB: The Commissioner may be a Subject Matter Expert.	Involved	Owner	Involved
Record data and assumptions, including uncertainties and accuracy, in a master data and assumptions list to record the origin of all data used.		Owner	Involved
If applicable undertake parametric analysis to understand the consequences of missing or uncertain data and assumptions.		Owner	Involved
Ensure data formats, units, and context are properly understood and handled.		Owner	Involved
Ensure implications of any data dependencies or relationships to other analysis or methods are understood.		Owner	Involved
Ensure the level of quality checking of the analysis is appropriate for the decision being supported: All analysis requires some checks, at some level, by another competent person wherever and whenever practicable.	Involved	Involved	Owner

Return

Deliver

Activity	Commissioner	Analyst	Analytical Assurer
Communicating the analysis to te commissioner as a final part of the review process	Involved	Owner	Involved
Ensure the results are clearly and effectively communicated in the context of the problem being considered.	Owner	Involved	Involved
Ensure uncertainty, risk, limitations, and constraints are clearly communicated, along with the results, to the study commissioner.	Involved	Owner	Involved
Ensure uncertainty, risk, limitations, and constraints are clearly communicated, along with the results, to the decision makers and stakeholders.	Owner	Involved	Involved
Ensure an analytical record is provided to i) facilitate access to the analysis by broader stakeholders, ii) make the analysis exploitable for wider decisions, and iii) inform continual improvement.	Involved	Owner	Involved
Ensure a suitable audit trail is in place that clarifies the level of validation, scope, and risks associated with the analysis. Best practice includes the production of validation log books.		Involved	Owner
Undertake reflective learning to capture successes and difficulties and ensure these lessons are available to improve future analysis.	Involved	Owner	Involved

Return

Cognitive transparency

- Purpose:
  - Develop a shared understanding of:
    - Requirements, constraints and the ‘art of the possible’
- Established via:
  - Analytical estimate
    - The task – the benefits the analysis is seeking to support
    - Cost, time, quality constraints
    - Analytical tractability within these bounds

Return

Conceptual transparency

- Purpose:
  - Proposed means to satisfy requirements
- Established via:
  - Concept of analysis
    - Intended data sourcing, measures, use of methods, models and techniques
  - Validation logbook / logsheet for each analytical means
    - Purpose and known competence; history of usage; user experience; evidence of V&V; known limitations

Return

Analytical transparency

- Purpose:
  - To ensure that the work is conducted with due RIGOUR
- Established via:
  - Technical reporting
    - Method, measures, results, any significant limitations, preliminary interpretation setting results in the context of previous reporting, and technical insights from the work
  - Master data and assumptions list

Return

Interpretative transparency

- Purpose:
  - Clarity of interpretation from study findings to customer need
- Established via:
  - Customer reporting with key caveats
    - Presentation; 2-sides for seniors; 5-sides contextual considerations; 10-sides key findings and mechanisms
    - Interpretative roles
      - Analyst: ensuring a fair interpretation of the process
      - Analytical assurer: results in wider context
      - Commissioner: customer context ensuring impact

Return

Validity criteria

- Reliability:
  - Alignment between the work and the intent of that work
- Face validity:
  - Alignment between the work and stakeholder perceptions
- Criterion validity:
  - The extent of engagement with the intended phenomena
- Construct validity:
  - The adequacy of representation of key processes
- Content validity:
  - The interpretative weight that the work can bear

Return

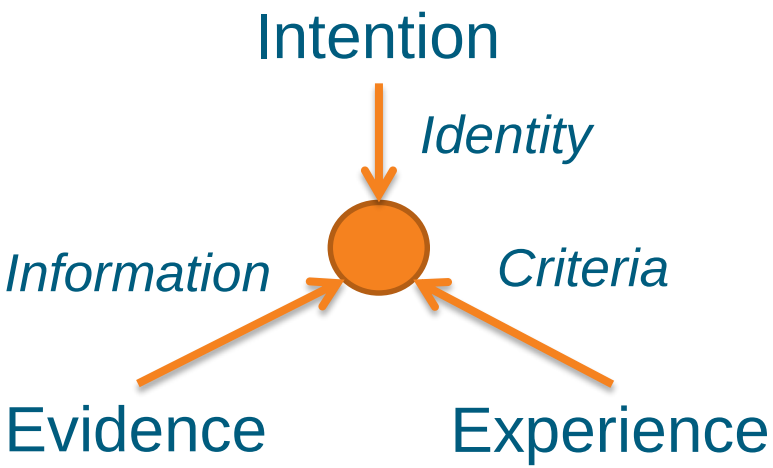
RIGOUR

- Repeatable – noting the fundamental limits
- Independent – engaging with the range of views
- Grounded in Reality – analysis is a journey
- Objective – engage and appropriately challenge
- Uncertainty Managed – identified and assessed
- Robust – communicating residual uncertainty

Return to analytical transparency

Return to phenomenological engagement

Fundamental limits



Produces a conjunction of engagement with a phenomena

- References
  - Fuenmayor R (1991) 'Truth and Openness: An Epistemology for Interpretive Systemology' *Systems Practice*, Vol. 4, No. 5 pp 473 – 490 <see pages 473-474 concerning *intentionality*>
  - Foucault M (1969) '*L'archéologie du savoir'*, Paris: Gallimard ('*The Archaeology of Knowledge'*, translated by Allan Sheridan, New York: Harper and Row, 1972)
  - Derrida J (1992) 'Force of Law' tr. Quaintance M in *Deconstruction and the Possibility of Justice* Eds. Cornell D et al New York: Routledge, 1992 pp 3-67 <see pages 26 – 28 concerning "the urgency that disrupts the horizon of knowledge">

Return

Framing the analysis

- The commissioner owns the benefits intended
  - The following aspects of framing should be led by the analytical assurer role
    - Framing: commissioner often too close to the problem
    - Detail: may bring assurance not necessarily benefit
    - Plurality: can help ‘triangulate’
    - Emergence: real leverage is often in dynamic complexity
    - ‘Facts’: ensuring appropriate inputs are used and appropriate interpretation is made of the work
    - ‘Prediction’: results indicative with uncertainty bounds

Return

Phenomenological engagement

- Prior beliefs can cause the commissioner of the analysis and the customer to start too close to the problem, while the analyst often starts too far away
  - The gap to the problem can be closed through
    - The framing of the analytical question
    - Identification of stakeholder group
    - Verification and validation processes
    - Transparency of process
    - RIGOUR
    - Reflexive journal

Return

Limits of the knowable

- The Ravens paradox
  - Cause: offering assertions of a pre-held view as evidence
  - Mitigation: assessment of criterion validity
- The Grue paradox
  - Cause: future asserted to be different from the past
  - Mitigation: testing the construct validity of this claim
- Underdetermination of theory by evidence
  - Cause: evidence does not allow us to distinguish between a range of divergent hypotheses
  - Mitigation: clarity concerning the uncertainty

Return

Framing

- Foundational narrative
  - Compare findings, including outliers, against expectations
  - Identify gap and propose explanations
  - Progressively and rigorously sift this understanding
- Boundary issues
  - Test breadth - more broadly framing a conceptual model
  - Test depth – through comparison of key processes with more detailed work
  - Test granularity – through comparison of key findings with more detailed work

Return