International Journal of Managing Projects in Business, Volume 14, Issue 4, May 2021, pp. 856-897
DOI: 10.1108/IJMPB-10-2019-0242

An empirical study of assurance in the UK government major projects portfolio: from data to recommendations, to action or inaction

Abstract

Purpose: Effective and robust governance of major projects and programmes in the public sector is crucial to the accountability of the state and the transparency of state spending. The theoretical discourse on governance, in the context of projects and programmes is not fully mature although is now sufficiently well developed to warrant an increased scholarly focus on practice. In this paper, we contribute to the empirical literature through a study of assurance routines in the UK Government Major Projects Portfolio (GMPP). **Design/methodology/approach**: A framework analysis approach to the evaluation of a subset of GMPP database generates original insights into i) the framing of assurance review recommendations, ii) the treatment of assurance review data and iii) the subsequent tracking of the implementation of actions arising from the assurance review process. **Findings**: Our analysis reveals that the 'delivery confidence' of the major projects and programmes included in this study improves during the time that they are assured on the GMPP. This would suggest that 'enhanced' governance routines are desirable in programmes and projects that exhibit high degrees of complexity and scale. **Originality/value**: Our research findings contribute to the wider conversations in this journal and elsewhere on project governance routines and governance-as-practice in the context of government and public services.

Keywords: project governance, project assurance, review recommendations, delivery confidence assessment, government

1. Introduction

Brunet (2018) suggests that project governance '*has gained recognition as an important object of inquiry*', but suggests that '*what is actually done by different actors having to manage those projects has been studied much less*'. Brunet's conceptualisation of '*governance as practice*' draws on a theoretical basis advanced in the *International Journal of Project Management's* 2014 special edition and includes transaction cost economics (Williamson 1998), agency theory (Eisnhardt 1989) and stakeholder theory (Donaldson & Preston 1995). In this paper, we structure our understanding of governance practices, and assurance in particular, through the literatures on project governance (see Müller 2009), project decision-making (see Cantarelli & Flyvbjerg 2015, Williams & Samset 2010) and organisational learning (see Miterev et. al. 2017, Söderlund et al. 2014, Gardiner et. al. 2018). The concept of "governance" has generated academic debate (see McGrath and Whitty (2015, 2019) and Musawir et al (2020)), we therefore accept that the definitions used in this paper may not necessarily reflect other conceptualistions of governance described in the literature.

The study described in this paper is situated in the major public projects arena, and our interest is stimulated by the authors observations on governments who have increasingly sought to use legislative instruments to frame their governance (or oversight) practices. The United States (US) Program Management Improvement and Accountability Act (2015) is one example while in the UK, a 'three-lines-of-defence' approach (HM Treasury, 2014) operates – this includes the role of a Senior Responsible Owner (SRO) accountable to the UK Parliament for the successful delivery of the project and its intended benefits. Recent inquiries by the UK Parliament Public Accounts Committee in 2016 and Public Administration & Constitutional Affairs Committee in 2019 further emphasise the important role that governance assumes in enabling SRO's to navigate projects through the difficult challenges that complexity brings; this is often amplified by the asynchronous nature of policy and project delivery (see King and Crew (2013) and Kirkham (2019)).

In this paper, we build on existing 'governance-as-practice' literature through an empirical study of the UK government's database of assurance review recommendations for all major projects and programmes overseen by the Government Major Project Portfolio (GMPP). Assurance reviews are undertaken not to penalise project team actions, but to provide meaningful and practicable reflections on project progress. As well as to communicate to project stakeholders any insights that can ultimately improve the success of a given project (Bourne & Parr, 2019) rather than to ensure adherence to any legislative requirement. Although, assurance reviews may be mandated as part of the approvals or business case process, and are designed to 'provide support and constructive challenge to senior responsible owners' (HM Government, 2019) Reviews are undertaken by a pool of independent professionals drawn from the civil service, industry and commerce.

Specifically, we investigate documented recommendations arising from assurance reviews and equate these with observed performance changes in the database. A content analysis of review reports taken from a subset of the GMPP combined with face-to-face semi-structured interviews with project participants to provide additional insights and context to the data forms the basis of this study.

The purpose of investigating the assurance recommendations is not to provide a 'one size fits all' solution but rather to exploit the dense GMPP database to update the contemporary understanding of how project assurance activities are enacted in practice. As such the essential research question is: *how are assurance review recommendations formulated and why are they not always implemented?* – but by looking at this, we hope to develop a deep understanding of the practical dynamics surrounding assurance and review processes within contemporary government projects and to provide actionable guidance for practitioners on how the assurance review process could be improved with a view to delivering better project outcomes.

In order to achieve this, our analysis is framed by a series of questions (see Table 1) that are

developmental in nature and based on an incremental familiarity and understanding of the assurance review data set. The questions are supplemented by a researcher-facilitated workshop, undertaken with external assurance reviewers and designed to equip the authors with additional contextual understanding of the data and to prompt questions on methodological issues underlying the assurance review process. In the following section we characterise the extant literature relevant to this study and therefore establish the basis of the research questions.

2. Literature review

2.1. Governance

Governance, according to Samset and Volden (2016), derives from the Latin *gubernar* ('to steer') and pertains to the administration of a nation or state; Fentrop (2003) asserts that the origins of modern corporate governance practices can be traced back to the legal and allied professions of the 17th century Dutch Republic. Subsequently, economists became interested in the decision-making processes of organizations and observed that governance theories might be helpful to organisations in shaping policies and procedures (Gilson, 1996). More recently, Gompers (2003) and Basu et al., (2007) observe that improvements in a company's governance processes may lead to increases in shareholder return and that independent management and greater shareholder rights can lead to increases in organisational performance (see Millstein & MacAvoy 1998 and Hirschey et. al. 2009). Conversely, weak governance mechanisms potentially lead to undesirable organisational outcomes including poor business performance (John & Senbet 1998) and 'black-swans' (Taleb 2007). These studies have stimulated researchers to extend governance theories to other sectors, organisations and academic disciplines such as healthcare, justice, education and the legislature (Too & Weaver, 2014). In this context, we understand governance as the means by which, organizational goals are achieved through policies, processes, and structures in the realm of projects (Müller, 2016). The recognition of projects/programmes as instruments of organisational strategy delivery has led to a

growing acknowledgement of the requirement for good project governance (Biesenthal & Wilden, 2014) – the benefits being a framework for decision-making that is logical, robust and aligned with other corporate activities incluing risk management and compliance. Whilst the literature on project governance is generally positioned as a subset of the corporate governance literature (see, e.g. Müller, 2009), the focus of project 'front-end' governance research is generally quite sparse (see Williams et. al 2019a).

2.2 Governance and governmentality

"Governmentality" provides a useful lens to view governance practices. The concept of Governmentality, this is, the behaviour or actions designed to influence, direct, or affect the conduct of people was first introduced by Michel Foucault in his '*Security, Territory, Population*' lectures. Li (2007) describes Foucault's theory as the '*conduct of conduct*' whilst Müller et. al. (2016) develops this notion, in the context of projects, to reflect human ('soft') aspects of project governance in contrast to 'hard' methodologies that underpin governance routines, commenting that '*while governance provides a particular project management methodology in an organization, it is governmentality that regulates how the use of this methodology is enforced*'. Müller considers governmentality to be the "*mentalities, rationalities, and ways of interaction, chosen by the governance roles to implement, maintain, and change the governance structure*"

Biesenthal and Wilden (2014) suggest that the development of studies in the project field is inextricably linked to organisational and corporate governance research. Turner and Simister (2000) argue that the theories of corporate governance can be rationally transferred to the practice of project and programme delivery; Miller & Hobbs (2005) and Crawford & Helm (2009) argue that project governance involving multiple companies is, in effect, a contract organisation. Sanderson (2012) presents an analysis of risks and uncertainties in an organization through governance theories; Müller et al. (2013) focus on governance methods in a project team, and their relationship

with ethics and trust. Weill and Ross (2004) discuss how to conduct a governance procedure for a particular department. The relationship between project governance and project success is also well developed in the literature: Müller and Martinsuo (2015) discuss the relationship between project owners and their suppliers; their conclusion that project governance promotes effective communication talks to the wider literature on the importance of collaboration, trust and ethics in project organisations. Wang and Chen (2006) analyse the impact of project governance mechanisms in the IT industry. Yazici (2009) shows that the cultural background of an organization will also affect the success of the project. Where a project's participants hold common interests and values, the likelihood of project success will increase.

In contrast, Clases et al. (2003) emphasise the consequences of 'over-control' in situations where the environment facing the organisation changes, and unreasonable response mechanisms emerge that lead to agency tendencies (see 2.3). Project governance research also generally seeks to emphasise the importance of good governance structures and governance processes (Peterson et al., 2002). Whilst particular researchers believe accountability to be one of the main driving forces behind of project governance, overtly controlling methods in an organization may increase resistance in agency tendencies towards accountability, resulting in effects at significant variance to those initially desired (Knodel, 2004). Complex governance systems may lead to decision inertia (Keyes-Pearce, 2002) and a tendency towards overty risk averse behaviours.

Whilst the literature describes the importance of sound governance mechanisms and their possible impact on projects/programmes (Müller et al. 2016; Aliza et al. 2011; Müller & Lecoeuvre 2014), the application of these systems in the organisational setting may not necessarily lead to performance improvements (Peterson et al., 2002). The enacted governance model may exhibit conflicts in practice so the co-ordination of the resolution of conflicts is a key 'dynamic capability'.

In 'conventional' democratic systems, good governance of major public programmes and projects is a key component within a broader framework of accountability and transparency of public

spending; consequently, project-studies scholars have increasingly turned their attention to theoretical and practical aspects of governance, considering how decisions arising from governance and assurance routines may influence short-run and long-run benefits (and disbenefits) and the tension between policy formulation and implementation through projects/programmes and operational delivery.

Crawford & Helm (2009) argue that effective governance frameworks could unlock potential improvements in performance and better attain desired outcomes while offering "traceability, transparency, and accountability". The implication is that learning, and the creation of new insights into performance, is a product of governance practices.

There are clearly defined governance procedures (Klakegg et al., 2009), influenced by the emergence of New Public Management since the 1980s. Indeed, these are particulalry well articulated in the UK (e.g., HMT & Cabinet Office, 2011). In this paper, we consider a governance framework as "*an organised structure established as authoritative within the institution, comprising processes and rules established to ensure projects meet their purpose*" (Klakegg et al., 2009, p. 60), which Klakegg et al base on previous literature such as Miller and Hobbs (2005).

2.3 Governance and agency problems in the context of major projects and programmes

While a number of theoretical lenses such as stakeholder theory and customer relationship management theory (Zwikael et al., 2015) have been usefully applied to the study of major projects and programmes, the units of analysis within this research naturally lean towards the explanatory power of agency theory. Renz (2007) describes agency theory as essentially concerned with the contractual relationship between the provider of projects resources and the user of those resources (Hannafey et al., 2013); the owner of project resources is usually identified as the 'principal' whilst the manager responsible for using and controlling these resources is identified as the 'agent' (Turner

et al., 2003). Therefore, when examining assurance reviews and their recommendations, there is a naturally defined agent (manager) and a naturally defined principal (owner). Agency theory provides a necessary theoretical lens to help us understand and interpret these key interactions in the context of the projectified environment. Developing this agency theoretic view, Zumbansen (2007) suggests that the separation of ownership and control is a common organisational form in modern organisations and in projects specifically because project investors often lack the specific relevant expertise needed for project delivery **and** transition to operations.

The corollary is that delivery teams may require more information than the principal, the project sponsor, and this information asymmetry may adversely affect the principal's effective oversight of whether the agent serves the principal's interests appropriately (Chang, 2014). Specifically, project owners and project managers represent different interests and in this case, it is possible that a moral hazard or conflict of interest may occur. Mahaney and Lederer, (2011) claimed that if the agency dilemma is neglected or not effectively dealt with through good project governance, undesirable risk and uncertainty may increase and thus delivery confidence may decrease.

Whilst agency theory is closely related to the concept of accountability (see Jensen, 2000), Trkman (2010) emphasise contracts and contractual responsibility as instruments to resolving agency problems; emphasising the advantage of different principal-agent relationships in project environments. Zwikael et al., (2015), in a more turbulent project environment, suggests that project sponsors should strengthen the control of project governance to reduce the negative impact of risk and uncertainty.

2.4 Learning arising from governance practices in projects and programmes

Grabher (2004) proposes the notion of a 'project ecology' as a means by which to conceptualise 'the essential processes of creating and sedimenting knowledge at the interface between projects and the agents and networks in and through which projects operate'. Gardiner et. al (2018) propose

the 'strategy-project system' as an alternative conceptualisation, that emphasises the crucial role of projects in organisational learning and strategy delivery; this work promulgates the 'dynamic capability' of learning ability and agility (seel also Lichtenthaler & Lichtenthaler, 2009), arising from the interaction of tacit accumulation of experience (Goffin et al., 2010), knowledge articulation, and knowledge codification. Hartmann and Dorée (2015) consider learning to be a contextually embedded social activity, using the theory of "social learning," so learning is not "copy & paste" but a translation and reinvention activity to enact knowledge in a new context.

Organisational learning is also concerned with how projects deal (or intend to deal) with unexpected events (Makridakis et. al 2019). Christianson et al. (2009) demonstrate how learning from rare events changed organisational routines at four sites of learning, suggesting project failure might be seen as a consequence of the failure of an organisation to learn.

Projects are often perceived as unique and learning from failure rarely happens due to organisational boundaries (Baumard & Starbuck, 2005). Learning from success also falls victim to survivor bias, whereby pitfalls are often overlooked as the sample examined did not experience these pitfalls. However, project learning through project reviews is one form of learning where organisations have learned from "samples of one". This depth reflection provides opportunities to examine outcomes that were potentially possible in that cause of the project but were not eventually realised, thus avoiding the problem of survivor bias. Williams (2008) shows it is unclear whether current practices achieve their aims for learning, identifying factors which inhibit learning, including lack of resources and top-management support, but also cultural (e.g., blame-culture), the temporary organisation nature, and project complexity. The literature on "lessons-learned" is pervasive yet this remains a major challenge to enact in large, complex organisations; Anbari et al. (2008) stresses the importance of post-project reviews/after action reviews in helping future project teams plan for success, yet recognises that the PMBOK "defines lessons learned narrowly … pertaining mainly to the closeout phase" (p. 13).

Gardiner et. al. (2018) emphasise many of the learning challenges that project/programmes experience through the lens of the 'strategy-project system' such as cultural, physical and political influences.

Later in this paper we explore the role of project reviews in triggering improvement actions, and how learning from implementation of recommendations may (or may not) occur. The change that reviews effect can be viewed through the lens of learning for the project itself (stage/phase/gate reviews) or other projects (ex-post reviews).

Ajmal & Koskinen (2008) underline the importance of preparing the organisation to foster a culture that facilitates and encourages knowledge creation, sharing, and utilisation. One knowledge-broker could be the Project Management Office (PMO) (Williams, 2008), but Pemsel & Wiewiora (2013) find that PMO functions are not always fully aligned with project managers' knowledge-sharing behaviour.

Aerts et al. (2017) suggest that public organisations (and their actors) learn from and share working experience (in the context of large-scale infrastructure projects). Of note was the transfer of (public sector) project teams from one project to another to allow project learning to be disseminated and knowledge transferred from the individual to the permanent organisation. Nevertheless these projects, perhaps particularly so for public-private partnership projects, were considered weak regarding organisational learning. We have considered project governance and how to learn from project to project (El-Tigani et. al, 2019). But often implementing that learning does not seem to actually happen – commonly known as the "knowing-doing gap", we consider this literature in Section 2.5.

2.5 Implementing recommendations and the 'knowing-doing gap'

Pfeffer and Sutton (2000)'s "knowing-doing gap" is well known, concentrating on "let[ting] talk

substitute for action." They consider over-reliance on past custom and practice, and problems of fear, mistrust and internal competition and ways to minimise these (in context of project studies, see Styhre et al., 2004; Williams, 2008). They note the importance of "closing the loop"; and point to measures and measurement process as major barriers to turning knowledge into action, concluding with eight "guidelines for action."

This idea has been taken up as part of the discourse within management (e.g., Zheltoukhova, 2014; Baldwin et al., 2011), particularly why people and organisations do not learn from events and successes/failures. Levinthal and March (1993) describe "failure myopia", "Any learning process tends to [seek to] eliminate failures … learning produces confidence and confidence produces favourable anticipations…" (p. 110), but balanced by Madsen and Desai (2010), who claim we learn better from failures than from successes. However, some scholars observed little learning from failures (Edmondson, 2011; Baumard & Starbuck, 2005). Downham and Lingham (2009), looking at national enquiries, say although investigations are often mandatorily conducted following serious incidents, "nowhere is there any focus on the process by which recommendations arising from such investigations should be followed through to implementation" (p. 58).

3 The coalescence of these considerations

The theoretical literature on governance, organisational learning weakness, inherent principal-agent problems and the knowledge-doing gap suggests that empirical studies into the practice of assurance regimes may offer an original opportunity to interrogate the dynamics of how government projects learn from past experience and how these agency problems are manifesting. Thus, in this study we focus on the UK Governments Major Projects Portfolio (GMPP) and, inter-alia, seek to understand the scale and scope of the gap between knowing and doing.

The GMPP represents the "most complex and strategically significant projects and programmes" in the UK (IPA 2017b, p. 1); it is overseen by officials in the UK Infrastructure and Project Authority

(IPA) (hereinafter referred to as 'the authority'). The work of the authority includes the annual oversight of > 200 independent assurance reviews using the Office of Government Commerce (OGC) Gateway™ process methodology. Gateway reviews (see Fig. 1) are episodic and occur at critical decision points in the project life cycle, known as Gateways 0 to 5: start-up (or strategic assessment); business justification; delivery strategy; investment decision; readiness for service; and operation review / benefits evaluation (IPA, 2017a). Gateway Review 0 scrutinises at the program level and is repeatable; Gateway Reviews 1-5 are conducted at the project level (IPA, 2017a). Reviews are usually commissioned by Senior Responsible Owners (SRO's), as part of a wider approach to reflection on project delivery.

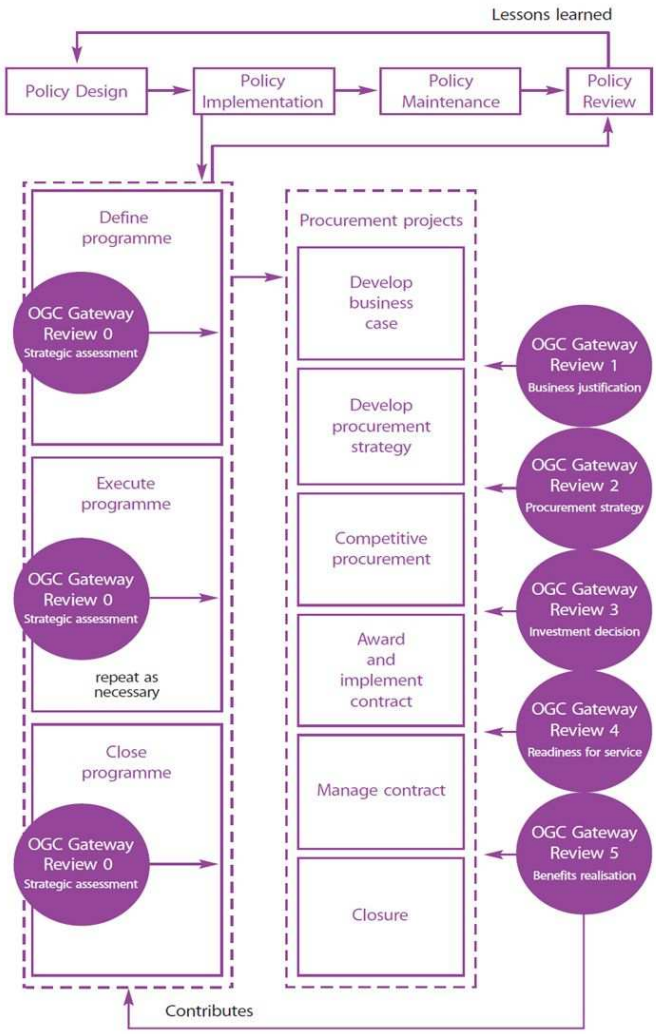


Fig 1: OGC Gateway Review Process

In addition to Gateway reviews, UK major projects and programmes are subject to other types of assurance such as Project Validation Reviews (PVR) and (maybe) Project Assessment Reviews (PAR). A PVR (formerly Starting Gate Review - SGR) is conducted during the early stage of a project, often before the SRO is appointed, to explore project initiation (IPA, 2017a). A PAR is usually carried out instead of one of Gates 0 to 5 if bespoke terms of reference (ToR's) are required and can be repeatable throughout the project lifecycle (IPA, 2017a). For further information, see IPA (2017a, pp. 5-10) and HM Treasury (2016, p. 24).

The benefits of the authority's reviews include insights generated by the independent perspective and the reflective nature of the conversations with the Senior Responsible Owner (SRO) and project/programme team. Reviewers form a temporary team for 3-4 days. At the review end, a brief report of evidence-based findings and recommendations is produced for the SRO, awarding projects with a non-negotiable delivery confidence assessment (DCA) using a five-point rating: Green ("Successful delivery of the project … appears highly likely"), Amber-Green, Amber, Amber-Red, or Red ("Successful delivery of the project appears to be unachievable") (IPA, 2012, p. 11). The report is discussed with the SRO and forwarded to the IPA; with proposed remedial actions agreed before the review team departs.

The UK National Audit Office (NAO) (NAO 2016) identifies some variability in responses to assurance review recommendations and suggests that the authority was unable to establish any causality between recommendations and improvements in project performance. The OGC (using a method approved and audited by the NAO) claimed the process had saved the taxpayer over £2.5 billion by 2007 as a result of more than 1,500 Gateway reviews (HMT, 2007); the validity of this is questioned in Xu et al. (2011).

The "project owner" (the principal) delegates their project accountability to a "project sponsor" who is responsible for the realisation of project outcomes and monitoring the behaviour of the "project manager" (the agent). The "project sponsors" of UK major projects/programmes – the SROs - use assurance reviews facilitated by the authority to address the issues caused by the separation of project ownership and control (Ahola et al., 2014, p. 1330). The UK's 34-year-old '*Osmotherly Rules'* were revised in 2014 to mandate accountability of the SRO to Parliament and "be expected to explain the decisions and actions they have taken during the period they are responsible for delivery of their project" (Riddell, 2014). The corollary is that the SRO takes ownership of assurance review findings "and ensures that appropriate action is taken to address recommendations in the advised time-scale" (IPA, 2016b, p. 2).

Project review is a term with various definitions in practice and in the literature. Here, the term will be used to refer to assurance activities that scrutinise projects in order to assess what happened and why and to decide appropriate actions to enhance the results of following activities. Reviews might look at resources (people, money, and time), performance (outputs achieved) and/or readiness (for the near future). Project reviews can take place at any time during a project life cycle. Various types of project review have been mentioned in, e.g., APM BOK (APM, 2012), PMBOK (PMI, 2013), and PRINCE2 (OGC, 2009). Oakes (2008) categorises project review by "frequency," "degree of formality," "degree of independence," "attributes being reviewed," and "review style" (pp. 34-39). Different review types typically come together in harmony over a project's lifecycle and provide a system for lessons learned and sharing, thus potentially improving project knowledge and performance (Kotnour & Vergopia, 2005). Different governments have different ways of categorising project reviews (Klakegg et al., 2010), with UK processes described below.

Much of the literature on project review considers stage-gate approaches, conducted at critical points during the project lifecycle. Key decisions are based on decision makers' comprehension of the project; however, the quality of decisions is regularly questioned. Bowman et al. (2015)

summarise failures of UK public projects as either projects undeliverable from the start or projects insufficiently governed; reviews aim to halt the former and correct the latter. Many governments have developed formal stage gate review approaches (Nalewaik & Mills, 2014).

The outcome of a gate review is textual information in the format of recommendations. Xu et al. (2012) show efforts made by Governments to develop methodologies to quantify review outcomes but without much success. They suggest that a gateway review could generate the most significant impact on the project outcomes during the business justification phase and least during the benefits realisation phase. Although they solely analysed Australian projects, this introduced some ideas for quantifying the value of gateway reviews.

Aside from SROs, other key actors are the authority's Strategic Delivery Advisors (SDAs).The SDA acts as the principal point of contact between the authority and a Government Department (Appendix 1), usually at the interface with the Departmental Portfolio Office (DPO) (or equivalent) and the GMPP projects/programme. The SDA initiates planning for a review, determines the skills and experience required in the review team, agrees ToR with key stakeholders, and recruits and briefs the review team. The review team is empowered with responsibility for the review, but with support from the SDA available throughout. The final report prepared for the SRO will form part of the evidence base for day-to-day engagement with the SDA going forward to provide continuity. (see Annex 1 and 2 for visualisations of the assurance regime and the wider context of Treasury approvals processes for major government projects).

4 Research Methodology

The aim of this study is to increase our knowledge and understanding of assurance routines in UK government major projects and programmes; specifically we seek to evidence how action (or inaction) may arise from the recommendations that arise. We define a framework, similar to that described in Volden (2019), of seven 'framework research questions' (FQs) with associated sub-

questions, addressing the findings of the literature review. The questions are situated in a broader framework of analysis that seeks to enact a methodological approach capable of i) sourcing a disparate data set comprising quantitative and qualitative data, ii) managing sensitive data in accordance with strict and exacting requirements and iii) contextualising the data to understand the nuances that may affect the researchers' understanding and interpretation. We propose a modified version of 'Framework Analysis' (see Gale et al 2013; Srivastava & Thomson, 2009), with the intention of generating a deeper and nuanced set of insights into the data available to the researchers. Framework analysis is agnostic of the epistemological and theoretical lens adopted by the researchers, and can be used with a range of qualitative methods (Gale et al 2013 in Hackett 2018). In this paper, we enact a six-step approach (see Figure 2):

(a) **Theoretical structuring, governance processes and data familiarisation**. An extensive literature review provided researchers with a theoretical basis to enact a semi-structured interview instrument, for use in establishing the current desideratum of the assurance review processes and data capture. The researchers interviewed an 'assurance review coordinator' and thirteen external (to the government) assurance reviewers at a workshop ("Workshop attendees"), we were able to establish an impression of the issues facing assurance reviewers, and derive premises for the first interview protocol and later analysis.

(b) **Pilot analysis**. Drawing on data from seven review reports of a GMPP project to develop and test the planned data-inspection, analysis, and interview protocol. The project was chosen based on the project director (PD)'s willingness in providing access to project materials.

(c) **First-cut quantitative analysis of the authority's assurance database (AD)**: An analysis was carried out in 2018 to understand the data, and identify key constructs. Some transparency data on the performance of the GMPP is made available annually and consolidated into the government's annual report on major projects. In parallel, the authority also actively monitors and update an internal assurance database; however, this is not publicly available as a data-base; to do so would "undermine the whole integrity of the assurance system" and "foster a

greater sense of risk aversion" (Meggs, 2018).

(d) **Assurance review data qualitative analysis framework:** to generate deeper insights into the data set, the researchers identified seven central themes, and a twenty-five further sub-questions emerged from (a)-(c) as shown in Table 1.

(e) **Detailed qualitative analysis.** A qualitative content analysis (Hsieh & Shannon, 2005) of assurance review reports from five GMPP projects (per Table 2) was undertaken, using the sub-questions in Table 1 as a coding scheme. Three of four types of GMPP projects (see below) were included; more transformation projects were investigated as they made up nearly half of the database (see below). Selection criteria included: the project had recently left GMPP, had existed on GMPP for at least two years, had DCA ratings not exclusively green/amber-green while on GMPP, and was not subject to data access restrictions (so military projects were not studied). The names of the projects/programmes selected are confidential.

(f) Deep-dive analysis: Following the detailed analysis, two GMPP projects (Project 1 and 2, with improving and worsening DCA rating respectively) were investigated using semi-structured interviews to address the research questions listed in Table 1. We recruited eight participants based on their willingness to be interviewed: for Project 1 (only), we were able to interview the project's IPA Strategic Delivery Advisor (SDA) and one of its SROs; for Project 2 (only), we were able to talk to the project's IPA SDA and two of its Programme Directors, of whom one was the project's SRO for a brief period of time; and we were also able to gain further information from our IPA assurance coordinator, another IPA SDA, and one departmental portfolio officer (who gave information regarding the relationship between the projects, assurance reviews, the IPA, and the departments). Many reviewers were "outsourced," and it was, therefore, difficult to secure their participation. The interviews lasted between 30 and 60 minutes; notes were agreed with the interviewees and analysed using the same approach described in (e). An overview of the interview details can be found in table 3.

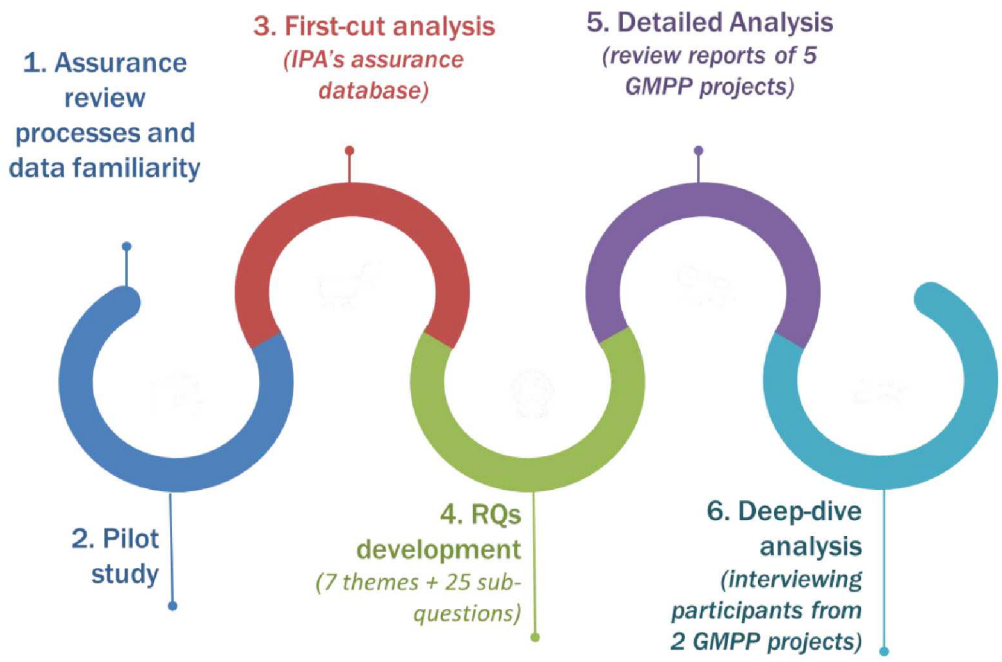


Fig 2: Research Methodology

5 Analysis and results

In this section we present an analysis of the assurance review data captured from the GMPP, this is supplemented by an analysis of the findings generated from the assurance review reports and interviews undertaken with the review teams. The researchers were provided with access to a database maintained by the Infrastructure and Projects Authority (and subject to a non-discolsure agreement); the structure of which reflects the guidance set out in '*A guide to choosing which IPA Assurance Reviews to include in the integrated Assurance and Approvals Plan for your Major Project*' (IPA, 2017a). The guidance is designed to assist project teams in developing Integrated Assurance and Approvals Plans (IAAPs) for the projects that receive assurance from the authority. Procedural explanations are set out in more detail in the '*Guide to planning Integrated Assurance and Approvals*' (IPA, 2017c).

5.1 Descriptive analysis of the IPA assurance review data

240 major projects and programmes were captured in the GMPP in the period 2011-2016. Our analysis of the portfolio identifed 747 assurance reviews that subsequently generated 5,500 recommendations. Project Assessment Reviews (PARs) and Gate 0 reviews, being repeatable, account for the most frequent review types, with 33.4% and 25.9% of all recommendations made respectively. Gate 3 (8.9%), Gate 4 (7.7%), and Stage Gate Reviews (SGR's) or Project Validation Reviews (PVR's) (7.1% in total) made up the remainder of recommendations of the five most frequently enacted review types (see Fig 2). Each GMPP project is assigned to one of four classifications: military capability (MC), infrastructure and construction (IC), government transformation/service delivery (transformation), or information & communications technology (ICT). Transformation projects account for c. 50% of all recommendations while c. 30% are ICT, as could be expected given the proportion of these projects on GMPP and their relative complexity

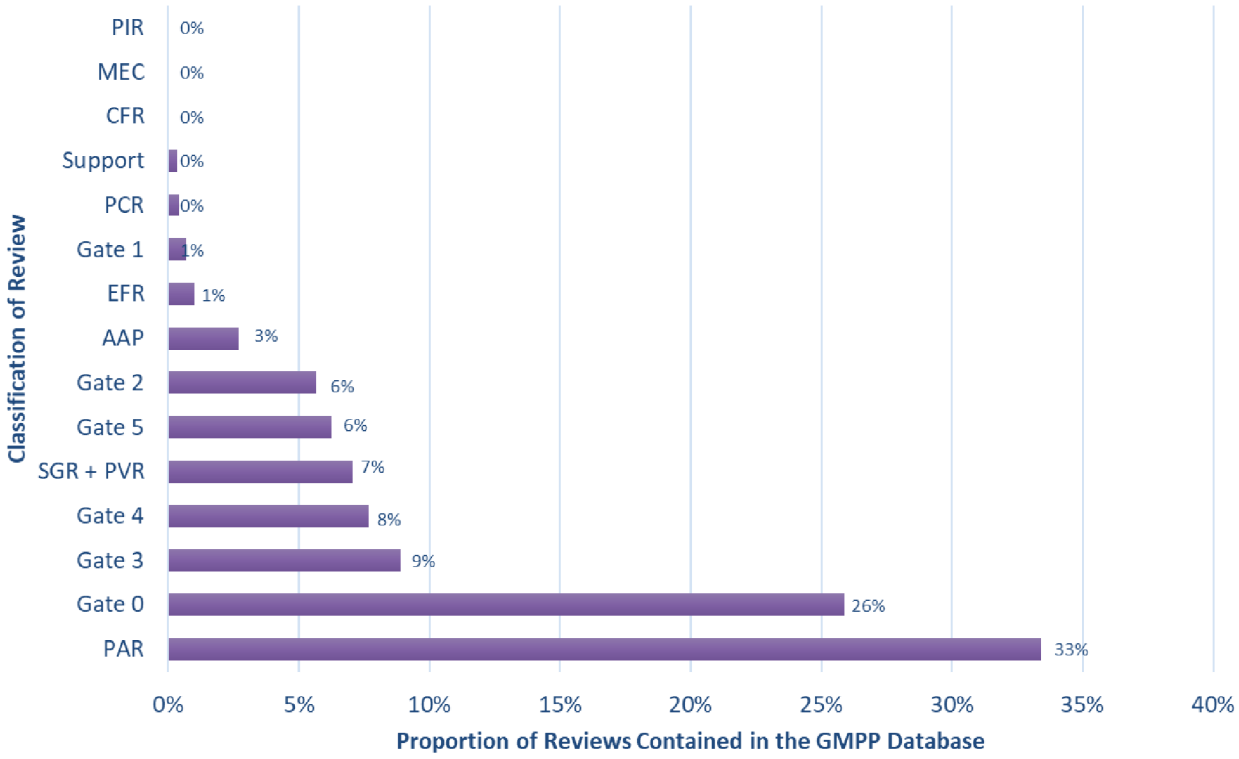


Fig 2: Proportion of assurance reviews by classification

5.1.1 Change in Delivery Confidence Assessment (DCA) of projects over the time spent on GMPP

On average, Delivery Confidence Assessment (DCA) ratings improved over time; a DCA rating of amber/amber-red at first review tended to steadily improve during the project's life on GMPP. Typically, eight reviews occurred prior to a green DCA. This is likley to reflect the uncertainty that characterises the initial stages of a project (Samset & Volden, 2016), in that its is more likely that projects will be assigned amber/amber-red (or even red) on entry to the GMPP. Overall, the prognosis for successful delivery improves during the period that a project remains on the GMPP.

5.1.2 Recommendations' degree of urgency

The recommendations are divided into three different levels subject to their urgency: critical (do now), essential (do by), or recommended (should benefit from following this recommendation). 43% of recommendations were marked "essential" (medium urgency), with the proportions of "critical" (most urgent) and "recommended" (least urgent) roughly equal. One interviewee suggested that reviewers might become involved in projects too late, so "critical" issues have mostly happened. That raised the question of whether RDU varied with the change of the reviews' DCA ratings (whether a "Red" review had a greater proportion of "critical" recommendations than a "Green" review). Analysis indicated that the proportion of "essential" recommendations tended to fluctuate 'around 40% irrespective of DCA ratings; however, the proportion of "critical" recommendations was noticeably different by DCA ratings, being 46%, 37%, 31%, 22%, and 4% for reviews rated "red" down to "green", respectively. This perhaps explains why projects tended to take recommendations more seriously when the review was rated red or amber/red (Fig 3).

Trend analysis shows that RDU tended to decrease gradually as projects progress from review to review whilst on GMPP. On average, projects' first review on GMPP had RDU around "essential," RDU then declined slowly and reached midway between "essential" and "recommended" after

approximately eight reviews. Average RDU was more likely to fall moderately as a consequence of being included within the GMPP, reducing the issues identified as needing immediate action. The rate of decline was different for different project categories, and IC projects (alone) (which on average started with fewer urgent recommendations) saw an increase in RDU over time. Although MC projects began their first review with more "critical" recommendations, their decline toward less urgent recommendations occurred faster; MC project selection and review processes are different, perhaps explaining why RDU decreased more during the project lifecycle.

5.1.3 Distribution of recommendations by their primary and sub-categories

Fig 3: Proportions of recommendations degree of urgency by delivery confidence assessment classification

13 primary categories are used to classify recommendations. Program/Project Management – PPM (19% of all recommendations); Governance (14%); Resource and Skills Management (13%); Context, Aim and Scope (11%); and Risk, Issues, and Dependency Management (9.3%) form the most frequent five primary categories. Only 3% of recommendations were classified as "knowledge management," with just 9% of these "critical." Each recommendation category is further sub-classified to enable deeper insights to be generated. "Project planning", a sub-category of PPM, was

a consistently critical concern for reviewers.

A typically high proportion of recommendations in certain classifications led to some speculation amongst workshop attendees that there was a sense that commonly used categories may cognitively associate with reviewers in general because they feel more confident in identifying issues related to particular primary categories or sub-categories.

5.1.4 Continuity of issues addressed

We examined changes in the pattern of recommendations during the duration of the project whilst on the GMPP. We assigned a value of one where a particular recommendation was in the same category in consecutive reviews and a value of zero where a particular recommendation did not have the same category in consecutive reviews. The average (i.e., correlation between the main issues raised in one review and the next) was 0.46, so some common concerns recur; this might appear high, however each primary category is divided into sub-categories – thus having the same primary category does not necessarily imply recommendations address precisely the same issues. In some instances the recommendations move into a sub category may indicate the recommendation has been broadly addressed previously but remains relevant from a changed perspective.

5.1.5 Continuity of Senior Responsible Owners (SRO's)

Senior Responsble Owners (SRO's) rarely deliver a project from 'cradle to grave'; in our analysis some 5% of GMPP projects had a single SRO responsible for the project over a 4-year period; the NAO notes that some "56% of projects have experienced at least 2 changes in SRO" (NAO, 2016 p. 36). Our analysis reveals a 50% chance of continuity of SRO 'in post' between one review and the next. Pettigrew (2008) suggests that continuity of leadership is an "absolutely critical factor in the performance of an organisation over time." However, our research also suggests that changes in SROs might not necessarily be problematic. Revisions to the "*Osmotherly Rules*" require an

enhanced level of seniority of SRO appointees. The limited pool of sufficiently senior SROs may increase the likelihood of individuals leading multiple projects. Other reasons for SRO turnover included "resetting" a project, changes to ministerial leadership, and changes in required civil service competencies. It is nevertheless desirable to monitor the effect of changes in project leadership (see Parker and Skitmore 2003).

5.1.6 Qualitative analysis of assurance review recommendations using textual analysis

We analysed the "narrative text" of recommendations to identify common themes using a simple NVivo frequency query. The term "plan" is quoted frequently, consistent with the findings of the assurance database analysis, showing the sub-category "project planning" as the most common category of issues identified.

5.2 Findings from the analysis of assurance reviews and interviews with reviewers.

In this section we present the results from the content analysis of assurance review reports gleaned from the five selected UK major projects and programmes (activity e) and from the 'deep-dive' analysis into two of those projects, (activity f). The findings are presented thematically, and in line with the set of framework research questions identified in Table 1.

5.2.1 A tailored approach to defining the scope of the review is preferred

Whilst the scope of project validation review (PVR) and gate reviews was tailored, standard terms of reference are used. However, the definition of the review scope - and the role of SROs in scope design - could not be evidenced. The majority of interviewees indicated that review scope was agreed by the strategic delivery adviser (SDA) and the SRO/project team before the review, and that the SRO could request that particular idiosyncratic issues be brought into scope. Assessment and planning meetings usually enabled review scope to be defined to suit the requirements of the SRO,

however this was not always the case. In Project 2, we observed (unusually) that the predefined scope of most of the reviews of left little in terms of flexibility, hence the SRO had little chance to shape the review; we also recorded evidence that the PVR was set up 'at significant pace' and that recommendations were made with insufficient testing of policies, leading to challenges later. Some workshop attendees felt assurance reports should be disseminated beyond the SRO, possibly within the authority and hence suggested improved interaction between the wider authority and the review team.

5.2.2 The pre-defined scope did not restrict the exploration of issues during the review

Review reports show that most recommendations focused on addressing the pre-agreed review scope. However, the research indicated that the pre-defined review scope seemed not to restrict the issues that could be addressed: rather than recommendations always addressing known issues, the team could amend the scope during the reviews. Most interviewees said review scope could be amended to accommodate emergent issues and it was impossible to know all issues before the review start. However, the review scope of Project 2, which was imposed on that project's reviews, did not always address what were felt to be its most critical issues. Interviewees indicated the usefulness of having skilled people on the review team, who could ask the types of questions that ensured the review focused on the significant issues.

5.2.3 Reviews were generally pre-planned, not all were commissioned by the SRO

Most reviews were pre-planned, with some reviews slightly delayed or advanced from baseline milestones. The only exception was one project brought to an early closure due to changes in departmental objectives, so two previous assurance reviews were not pre-scheduled but triggered by the closure proposal; it was unclear how the project team viewed the recommendations in those two reviews since we were not provided with an implementation update. We also found that not all reviews were commissioned by SROs; it is perhaps worthwhile to consider whether different types

of review commissioners led to a variation in the level of implementation of recommendations.

SDAs stressed that review outcomes needed to be beneficial to both the IPA and the SRO, so it is crucial that the SRO is involved in pre-defining the review scope. That was typically done during the assessment and planning meetings where the scope could be influenced (although as in FQ1.1 and FQ1.2, the SRO (for one unusual project) did not always have an opportunity to thus participate, leading to adverse effects on the receipt of recommendations – (Ref: FQ5.1).

5.2.4 The effect of SRO turnover on the implementation of recommendations was not clear.

All five projects exhibited high SRO turnover, but the effect varied. There was no evidence that implementation of recommendations or the DCA decreased due to SRO change. Participants suggested that reducing SRO turnover was a priority, but that the change in SRO was not necessarily an indication of a problem (see 5.1.5). SRO turnover appeared to produce negative effects (e.g. the difficulty of the SRO adapting to the new environment, and understanding the continuity of that environment, or a potential shift in priorities.) but also positive effects, including new SROs presenting a fresh perspective on the project, or the decision of the 'right person at the right time' to ensure project success. SRO turnover might not always reflect reality, e.g. one SRO interviewed had to step down to PD and was replaced by the Executive Director due to the new Osmotherly rules (see above), so the nominal change in SRO made little difference to the project.

5.2.5 Continuity of reviewers is perceived to be important, but specific expertise in the project less so

The sample of reviews illustrates that c.50% of reviews featured at least one review team member (RTM) from the previous review(s). Our analysis suggests that the benefits of the continuity of reviewers are inconclusive. All interviewees suggested that there should not be continuity of the whole review team, but that it may be beneficial to retain at least one reviewer. A mixture of reviewers was suggested, comprised of some with experience of the project and others with

expertise from other domains. The authority aimed to maintain the continuity of at least one reviewer; but this was dependent upon availability.

Whether reviewers looked only within their expertise could be investigated by comparing the expertises of each RTM with issues addressed in their reviews, but this information was not available. The SRO of Project 1 (a successful project) thought the reviewers were excellent in that they looked at all the project-related issues, no matter what kinds of expertise they had. Some Workshop attendees raised a question about the degree of training received by external reviewers; although the assurance process for Gates 0-5 was well understood, additional training and explanation on the other types of assurance (e.g. Critical Friend Review) could be usefully given to reviewers. Review team leaders had the opportunity to give negative feedback on unproductive RTMs during peer assessments, but it was suggested that they often avoided doing so in practice..

5.2.6 Project reports and interviews suggested that most recommendations made were based on both reviewing project documents and interviewing relevant people.

Generally, review teams were well informed before carrying out a review; project documentation usually provided the team with "groundwork" pointing to key issues to be covered. Interviewees stated that one-to-one interviews were vital; correlation crosschecks between documents and speaking with key project stakeholders helped reveal outstanding issues needing further investigation. Reviewers often formed views after reading documents, but many issues only came out from interviews, which might change initial perceptions.

Some Workshop attendees felt it was easy to get overwhelmed by the immense amount of information given, and it was hard to judge which documentation to read in a short timescale. This view was reinforced by an SRO's comment in one review report about the review team having "insufficient time to read and absorb all the evidence presented."

"Many early warning signs are of a less measurable nature and thus depend on more 'gut feeling'

approaches" (Williams et al., 2012, p. 46). Although we could not find any evidence about reviewers using "soft" indicators whilst they were undertaking project review interviews, some Workshop attendees noted that gut feeling and experience helped them direct interview questions, and suggested a range of "soft" indicators, which they might use during an interview to help detect the extent to which an interview was misleading (e.g., recalcitrance, inconsistency of message/answer, and parrot fashion answers). We found these "bad signs" not prominent in either review reports or interviews with key informants. We also did not find any commentary by the reviewers in their review reports on the conduct of the reviews. For the one unusual project, it was believed that the reviews only tried to handle emerging delivery issues instead of concentrating on addressing the main issues of policy assumptions.

5.2.7 Only a minority of interviewees expressed concerns over the insufficiency of time allocated to reviews

Typically, the authority's assurance review was restricted to a 2-4 day period. Review reports did not indicate reviewers complaining about time issues, except for one SRO who suggested that the IPA should consider lessons from one review as the short time available for the Team had led to significant rework of the report for factual accuracy; however, this was felt not to be typical. One interviewee suggested that it might be beneficial for complex projects to have a review over four days in order to include a larger number of interviews and leave enough time to write the report.

5.2.8 Recommendations focused mostly on delivery rather than benefits, with over 70% of recommendations addressing delivery issues

Literature on "project success" divides this concept into "delivery success" and "benefits success." PVR and Project Closure Reviews had a slightly higher percentage of recommendations focusing on benefits, perhaps implying a skewed emphasis on benefits towards project initiation (to get projects started) and closure (to get projects signed off). This finding agrees with a more general

finding in the PMI Project Benefits study (Williams et al., 2019b). Interviewees commented that recommendations were much more linked to what was directly needed to get to the next stage of the project cycle, rather than project benefits. Chipulu et al. (2018) show that when asked about project success, people tend to talk about effectiveness (benefits) criteria, but when talking about project failure, they talk about efficiency (delivery) criteria. However, we found no signs of the balance between recommendations focused on project delivery and those focused on benefits varying with DCA (e.g., reviewers concentrating on delivery if DCA was red, and benefits when DCA is green).

5.2.9 Review reports confirmed that most recommendations were aimed at the SRO and the project/program level, but some were aimed higher.

While recommendations were generally directed toward the SROs or the project teams, others were "messages" for higher levels such as Permanent Secretaries. For example, Project 1's SRO believed that although most review recommendations for his project were aimed at the SROs, a minority were targeted above the project to convince the Minister or the Department to take actions.

Workshop attendees from phase (a) indicated that giving too many or combined recommendations might inhibit recommendations being implemented. We found that only 8% of recommendations made for the five selected projects were combined ones and that most combined recommendations tended to originate from the same review teams. For the selected projects, the number of recommendations made in each review report ranged widely, from three to 26; only 11% of the review reports had over ten recommendations, whilst over half had 7-10 recommendations. Thus given the continuity between recommendation type and review team, it seems the intended recipient of a recommendation may vary in accordance with who conducted the review rather than for a systematic rationale.

5.2.10 Review reports show that few SROs gave comments on the reviews; amongst those who did, most were positive

It was unclear whether these comments had real meaning or were only a formal complimentary gesture. In only two cases that were observed did SRO give negative comments, including (i) the review team did not focus on addressing the pre-agreed primary issues and tended to explore other areas, and (ii) recommendations were ambiguous or made based on hearsay rather than material evidence. Here, the lack of buy-in from SROs was evident. It was interesting that these SROs disagreed with the review findings while, theoretically, they were the customers and probably commissioners of the review; this raises questions about the general role of the SRO in the assurance process. Some Workshop attendees felt there was pressure to include in the report recommendations that the SRO specifically wanted to see, e.g., a need for additional resources. They also noted that some SROs could become defensive when presented with negative findings or a worsening DCA.

A review recommendation cannot compel action; rather it directs the SRO towards action that will support improved delivery. It is therefore at the SROs discretions as to weather to enact these recomendations Project 1's SRO thought the effect of this depended on for whom the reviews were intended. The interviewee, as an SRO, was satisfied with the current policy in which a review's DCA rating was not negotiable, whereas recommendations might be discussed before finalisation of the review report. S/he said recommendations should be made for the SRO and should be seen as a means to support the SRO; if recommendations were considered made for the Department, there would be a danger of recommendations being manipulated and might not get buy-in from the SRO and project team. It was evident from review reports and interviews that if the SROs were not convinced of a recommendation, s/he would discuss it with the review team to clarify the situation rather than implement it. Although an assurance review is not an audit and should only be seen as complementary to other internal assurance activities, there was still a strong opinion amongst most interviewees that the IPA assurance process should provide a valid picture of the project for the SRO as part of the control mechanism. It would be worth looking further whether and for how long recommendations remained appropriate or valid after a review. As projects progress,

recommendations' objectives might no longer be appropriate.

5.2.11 There was no evidence to suggest that a 'blame-culture' featured in the review process

"Blame-culture" (Brown and Jones, 1998) is commonly observed and perceived in project-oriented settings, Williams (2008) observes this specifically in a UK Government setting. However, matters related to "blame culture" were not noticeable in either review reports or interviews. In Project 2 there were discussions between parties on how the project was undertkane in a 'theoretically effcive manner,, but the project did not achieve its desired outcomes; however, the interviewees did not consider those discussions as a "blame game," instead as an attempt to seek the lessons learned for continuous improvement.

5.2.12 The reasons why review recommendations were not implemented is unclear

Workshop attendees expressed a variety of opinions and observation on possible causes of suggested recommendations having not been implemented. Whilst we observed a few recommendations that were unpersuasive (Ref: FQ5.1-5.2), we could not find any evidence of other causes. Project 1's SRO said s/he always encouraged candour and frank discussion throughout the review. The two PDs of Project 2 pointed out an occasion when the review was not considered as helpful; in that particular case, the interviewees agreed that the reviewers did not explore the real problems facing the project (although raised to them by the project team).

If a project receives a "red"/"amber-red" DCA following a review, an 'Assurance of Action Plan' (AAP) review ought to be undertaken to "to check whether actions taken since the previous review have significantly improved the likelihood of successful delivery" (IPA, 2016a, p. 3). However, the review reports studied show that not all such reviews were followed by an AAP. One SDA indicated a lack of resource to carry out reviews was one reason.

Interviewees indicated that, as far as they knew, recommendations from current and previous

reviews were tracked by the Programme Management Office (PMO) and/or DPO (or equivalent), where applicable (see Figure 1). We suggest that the DPO could be the catalyst for enabling organisational learning from projects (see Gardiner et al., 2018). It was indicated that DPOs had existed since 2000; by 2010 Departments with a DPO were still the minority, but all were aiming to have one. DPOs provide assurance and report to their Departmental Boards on their portfolio and are linked to the authority, with reporting on GMPP schemes and overall transformation progress.

5.2.13 Our analysis emphasises a high correlation between the primary categories of recommendations from one review and those in the following review

The desire to respond to 'over-bureaucratic' assurance control in the public sector causes staff to 'tick boxes', "whether this compliance is for real or merely legitimating activity" (Arnaboldi & Steccolini, 2015, p. 6). Although most interviewees suggested that assurance review recommendations were implemented earnestly, review reports show some recommendations, already marked as "completed" or "done" recurring in subsequent reviews. The data analysis also emphasises a high correlation between the primary categories of recommendations from one review and those in the following review. It would be worth looking further into whether the repetition of recommendations is a sign that some recommendations were just done as a box-ticking exercise to progress to its next stage (see the literature on the "knowing-doing" gap), or were due to changes in project circumstance that made the involved issues arise again. One interviewee said sometimes repeated recommendations were made to reinforce work that had already been done.

Most workshop attendees noted they had observed some assurance review recommendations not being implemented for reasons detailed in FQ5.4. Two of Project 2 review reports showed that some recommendations were not acted on as the SRO remained unconvinced of their appropriateness. One interviewee believed that non-implemented recommendations were uncommon, but when they did occur the main reason was a lack of senior leadership or 'buy-in' (i.e. the SRO did not have sufficient authority in the organisation to get such recommendations

implemented or believed that the recommendation was inappropriate).

5.2.14 Interviewees acknowledged that there was almost no follow-up on recommendations by the reviewers and that there was little sanction for not implementing recommendations

A review team was only responsible for checking actions that resulted from recommendations made in a preceding review. All other recommendations (i.e. outstanding recommendations and recommendations made for the final review) were monitored by the PMO and the DPO or its equivalent. Figure 1 illustrates, in principle, how recommendations are tracked and reported. NAO (2010) suggested incorporating continuous assurance, having assurance reviewers working alongside project teams in its system, at least for most high-risk projects; however, NAO (2012) shows that the IPA did not consider it had the resources to conduct continuous assurance and thus did not implement this suggestion.

5.2.15 The delivery confidence rating might not automatically improve despite the implementation of all recommendations.

Changes in the project/programme environment may lead to declining currency of review recommendations, or recommendations in themselves not having the desired impact. NAO (2016, p. 7) states that the authority "has not yet established a link between assurance review recommendations and project performance and 'departments' responses to assurance recommendations are varied". The authority now requests SROs to complete a "Review Feedback Form" to indicate their view of the review (IPA, 2016b). However, this practice is optional, and as reviewers are entitled to see this feedback, it is not clear whether SROs will necessarily voice criticism.

Our analysis shows very few SROs gave feedback directly in review reports. Where they did, most feedback concerned the helpfulness and usefulness of the reviews, either by offering advice or reinforcing work already being done. Nearly all interviewees agreed that the reviews provided the

SRO with insight and meaningful recommendations. Project 1's SRO was positive about the outcomes of assurance reviews and felt Project 1 would have been less successful if the reviews had not been conducted and recommendations not been made. In contrast, the two PDs of Project 2 thought its assurance review recommendations only concentrated on project delivery and were irrelevant to the main issues facing the project (underlying policy) and the realisation of intended benefits; that project's situation did not improve despite the project team implementing most of the recommendations; they thus did not think the reviews were of value for the SROs and felt the assurance system restricted the reviews' value (i.e., limiting the way people think from imposing a pre-determined structure - Hodgson's (2002) Foucaldian analysis). Finally, interviewees thought the assurance processes was not updated to reflect what was taught in the governmental courses for senior project leaders; a general view amongst interviewees seemed to be that the process was about how to progress to the next stage instead of trying to understand the projects' real operating environment.

Project 1's SRO noted differences in opinions on whether it was the DCA or the implementation of recommendations that made the difference to a project's further progress. For Project 1, when the DCA was Amber-Green, the senior management probably did not take recommendations as seriously as when the DCA was Amber-Red; if the DCA rating was Red/Amber-Red, recommendations would have a higher likelihood of being executed to move the project forward. It is therefore both the ratings and the recommendations that make a difference.

6 Discussion and conclusions

The findings from this study, and the learning that has occurred, may be viewed through the lens of Ostrom's (1990) mainstream "rational thought" perspectives. These outline eight principles of managing "common" goods:

- scope and boundaries of the common (in this context, who can conduct what projects - e.g.,

digital transformation programmes should include the UK Government's Digital Service);

- appropriation and provision (the legitimacy arising from authoritative Treasury guidelines);
- collective-choice processes for appropriation of the common (a "rational" process of project appraisal applied by the Treasury);
- effective monitoring and control of resource consumption;
- a scale of sanctions for violators of resource use rules (this would include Departmental Spending Limits (DSL's), Parliamentary Select Committee's);
- mechanisms of conflict resolution (judicial review, public enquiries);
- self-determination recognised by a higher authority (perhaps not currently);
- multiple layers of nested governance.

The gateway process is one of a number of means of covering all eight of these issues, and the principles show that the appraisal and follow-up of recommendations from such reviews form an important governance structure to prevent departments or indeed individual projects simply doing what they want. An agency theory lens provides a theoretical basis to explain how the government manages its interrelations in the context of major projects and programmes and why the assurance process is necessary. There are a number of important concluding remarks that can be drawn from this study.

The delivery confidence of the projects included in this study generally improves over the period of time spent on the GMPP, but it is not possible to generalise this conclusion to the wider population of projects and programmes. We also observed that 50% of assurance review recommendations were assigned a medium level of urgency; but it is not clear whether reviews are executed at the most strate. The proportion of recommendations at the highest level of urgency decreased as a project's DCA rating improved. RDU usually decreased over time (except with IC projects), perhaps because project issues needing immediate action reduced while projects were within the GMPP system (receiving the authority's support). Military projects began with many

more most-urgent recommendations. **Continuity of recommendations between reviews:** The correlation between recommendations from one review and those in the next (regarding primary categories of issues) was almost 0.5, which seems quite high. **Commissioning a review:** Not all reviews were commissioned by the SROs, e.g., a PAR can be commissioned either by the SRO, the authority or the chair of HM Treasury's Major Projects Review Group. **The review scope:** Each review's scope was pre-agreed (assessing standard ToR as a minimum). Scope could generally be defined to suit the SRO. Most recommendations addressed specific concerns agreed within the scope, but the scope could be amended to accommodate emergent critical issues. Most reviews were pre-timed, with some reviews slightly delayed or advanced from baseline.

Effect of SRO turnover: There was only a 50% chance of continuity of SROs in post on projects. However, the impact of SRO turnover was unclear. There was no evidence that implementation or DCA was related to a change in SRO. SRO turnover was not necessarily an indication that something was wrong- it results in both positive and negative effects. Discontinuity in project leadership, its causes and impacts need to be better understood.

Review teams and members: Half of the reviews had no review team member from the previous review. It appeared difficult to maintain continuity of RTMs/leaders due to availability. It seemed there should not be continuity of the whole review team, but beneficial to have continuity of at least one RTM and to have a combination of reviewers who had knowledge of the project and those who could bring expertise related to the project stage/type of review. Reviewers looked at all issues and were not limited to their areas of expertise. However, external reviewers appeared sometimes to have insufficient training. Review team leaders might be reluctant to give negative feedback on RTMs to maintain good relationships, despite having the opportunity to do so.

The formulation of assurance review recommendations: Most respondents felt that in general there was enough time to carry out the review, with a few exceptions on very complex projects/programs; it might be beneficial for these to have reviews over four days or more. Most

recommendations were made after both reviewing documentation and interviews, but sometimes reviewers needed to review a large amount of information in a short timescale.

Issues addressed by assurance review recommendations: Recommendations focused more on delivery (70%) than benefits. PVRs (or SGRs) and Project Closure Reviews focussed on benefits slightly more often, implying more emphasis on benefits during project initiation and closure. Recommendations were typically more linked to what was directly needed to get to the next stage of the project than to addressing project benefits. Issues relating to the achievement of benefits/strategic fit should be continuously assessed in all types of reviews. Our analysis shows no indication that reviewers paid more attention to benefits when DCA was green and to delivery when DCA was red.

Number of recommendations per review. This ranged from 3-26; 11% had more than ten; the number should perhaps be limited to more critical issues. Combinations of issues were covered within 8% of recommendations - each issue should be addressed in a separate recommendation.

Behavioural dynamics of reviews: Very few SROs gave comments on reviews. Of those who did, most were positive. It was unclear whether comments had real meaning. SROs are now asked to complete a feedback form after each review. This is still new and optional, so the impact of this on assurance is unclear. Reviewers are entitled to see feedback, so SROs may be inhibited in commenting.

In some cases SROs gave negative comments on reviews, even though they were the review customers and probably commissioners; here lack of SRO buy-in with recommendations was evident. Recommendations should be seen as intended to help SROs. Reviewers should ensure recommendations are convincing, unambiguous, well written and understandable. Perhaps review DCA rating should be non-negotiable but recommendations debatable.

We did not detect signs of a “blame game”, but rather a debate of how or why such projects could

not achieve their desired outcomes despite best endeavours. Perhaps this debate should be encouraged in a more transparent (to the public), as a means by which the sharing of lessons learned can be improved. Project reviews are one specific way to socialise learning from previous projects.

Implementation of recommendations: The analysis suggests that the DCA rating influences the propensity to implement review recommendations (or not): amber-green DCA's are, for example, unlikely to attract little enthusiasm to take forward, in contrast to 'red-amber' DCA's. We note that AAP reviews should be undertaken for 'high-risk' projects receiving a poor DCA, but this did not necessarily happen. We also observed inconsistencies in tracking (follow-up) of implementation of recommendations. DPOs (where they exist) could play leading role in tracking implementation, which not all currently do; and there could be a common mandate regarding the formation of a DPO and their role.

Contribution to the literature

This paper presents a first ever indepdent research investigation into assurance review data contained in the UK Government's Major Project Portfolio. The results generated from this study provide insights into the value of enhanced governance for complex projects and programmes and furthermore, the broader issues associated with the accountability and responsibility of those changed with their leadership. The research also provides additional contributions to the literature on benefits realisation management in the sense that evidence arising from the study illustrates that recommendations (in assurance reviews) tend to focus typically on (project) delivery rather than long-run benefits, this complements the wider literature, which suggests that project success (and the realisation of benefits) may be seen as secondary to project performance in the context of governance routines.

Further research

The data analysis suggests that it may be appropriate to assess if different project selection and

review processes in defence influenced the observed reduction in 'recommended degree of urgency', and conversely, why the level for infrastructure & construction projects tended to worsen (slightly) over their project cycles. This could be achieved through establishing a control group of projects to be reviewed earlier than currently recommended to compare the level of urgency of resulting review recommendations.

Uncovering the reasons for 'repeated' recommendations (see 5.2.13) may offer improved insights into the challenges of organisational learning; a "deep dive" into a sample of recommendations that are repeated in subsequent reviews may offer some useful answers. We also suggest looking further into whether recommendations are taken more seriously if reviews are called ad-hoc rather than pre-timed.

Clarifying whether different types of review commissioners lead to a different level of recommendation implementation and studying antecedents and consequences of SRO turnover, specifically the apparent correlation between SRO tenure & changes, and stability/ improvement in DCA ratings. Finally, we suggest a detailed examination of the correlation between the focus of recommendations (on benefits or on delivery) and the review's DCA ratings.

Limitations of the study

The main weakness of this study was the paucity of projects studied and participants interviewed due to restricted data access. Therefore, the findings are not generalizable beyond the subset of projects that were examined. The data-base covered all types of project, but the two cases only covered two of the four GMPP categories (infrastructure and transformation). There is also a methodological weakness in the research, in that the cases are regarded as post-case-studies, while trying to contribute to the governance-as-practice literature, which would be better studied by contemporaneous detailed studies. However, the study, however, to the best knowledge of the authors, is the first to provide unique insights into the assurance of UK major projects and

programmes through a systematic analysis of assurance review recommendations. The results of this study suggest a number of further avenues for research including benchmarking improvements in confidence ratings for GMPP projects with other public sector projects delivered by government departments and 'arms-length' bodies.

In conclusion, there are a number of practical recommendations that arise from this study, The findings suggest several courses of action that may lead to potential improvements in the approach to the design and implementation of the assurance review processes described in this research, as outlined in Table 4.

Acknowledgements

This research is funded by the Project Management Institute (PMI) and executed under the governance of the Economic and Social Research Council (ESRC) Project X. The authors wish to recognise the ongoing support of Mr Stephen Townsend, PMI Director of Network Programmes.

Table 1. Assurance review data analysis framework

Analysis framework questions (FQ's)	Analysis framework questions from (a) Pilot study (b) First cut analysis of governassurance review data (c) Interviewing Workshop Attendees (d) Interviewing project delivery professionals
1. The review scope	
1.1. How is the assurance review scope defined?	(a)
1.2. How does the definition of scope affect the reviews' usefulness?	(a)
1.3. Are the reviews pre-planned? Would recommendations be taken more seriously if the reviews were called ad-hoc?	(b)
1.4. What is the significance of the SROs defining the scope?	(b)
1.5. What is the effect of SRO turnover?	(b)
2. Review Teams	
2.1. Is having consistency (or non-consistency) of review team leaders/members good or bad regarding the effect on Assurance review recommendations?	(b)
2.2. Is there evidence of reviewers only looking within their expertise?	(a)
3. How are recommendations made?	
3.1. Are recommendation derived mainly from interview or documents? How good is the quality of documents provided to the reviewers?	(b)
3.2. What is the effect of the SRO's oversight in planning the reviews?	(a)
3.3. Whether there were any bad signs (such as recalcitrance, the inconsistency of message/answer, parrot fashion answers, etc.) during the reviews as pointed out by Workshop attendees?	(c)
3.4. Whether enough time was given to the reviewers to do the reviews?	(c)
4. What do recommendations cover?	
4.1. Are recommendations concentrating on project benefits or operational delivery? Does the type of review make a difference here?	(b)

4.2. At what level of authority (i.e., project, SRO, departmental, ministerial, etc.) are recommendations aimed?	(b)
4.3. How often were there too many or combined recommendations?	(c)

5. Receipt of recommendations

5.1. What influences comments by the SROs? Can we detect "lack of buy-in" from the SROs, and why?	(a) and (c)
5.2. What is the effect of the implementation of review recommendations not being mandated? Should recommendations be mandated anyway? What are the implications of the SRO being responsible for implementation?	(b)
5.3. Is there any evidence of the influence of a "blame culture" when the project goes wrong? What is the effect of the various parties' incentives on the review process?	(b)
5.4. Any evidence of issues raised by workshop attendees? - Reviewers as "outsiders" who could not understand the project well. - Reviews as organisational bureaucracy. - Unconvincing recommendations. - Bandwidth, capability, or understanding issues. - Badly written recommendations. - Subsequent/separate 'commercial' decisions (by the Department and/or externally) rendered the recommendations 'not fit for purpose'.	(c)

6. Implementation of recommendations

6.1. When do Assurance Action Plans (AAPs) happen and why?	(b)
6.2. If there are not AAPs, how do projects record the progress of implementing recommendations in their systems?	(b)
6.3. What evidence is there that recommendations were taken to heart or just done as a "tick box" exercise?	(b)
6.4. Why were some recommendations not implemented?	(c)
6.5. Do we see evidence of a lack of follow-up on recommendations?	(c)

7. Effect of recommendations

7.1. Where recommendations are followed, why does delivery confidence not always get better? Do the reviews do any good?	(b)

7.2. Did the review's delivery confidence rating or the implementation of recommendations make the difference? If there was only the delivery confidence's colour but no recommendations, would that have any effects on the subsequent project performance?	(d)

Table 2. Selected projects for detailed analysis

	Review 1	Review 2	Review 3	Review 4	Review 5	Review 6	Review 7	Review 8
Project 1	Amber	Red	Amber/ Green	Amber	Amber/ Red	Amber/ Green	Amber/ Green	Green
Project 2	No DCA	Amber	Amber	N/A	Amber/ Red	Amber/ Red	Amber	Amber/ Red
Project 3	No DCA	Amber/ Green	Amber/ Green	Amber	Amber/ Green	Green		
Project 4	No DCA	Amber	Amber	Amber	Amber/ Red	Amber/ Red	Amber/ Red	Amber
Project 5	Red	Amber/ Red	Amber/ Red	Amber				

Table 4. Recommendations arising from the research.

Theme	Recommendation	
1. Roles and responsibilities	1.1	Ensure the input of the Senior Responsible Owner in the definition of assurance review's terms of reference.
	1.2	Consider responsibility for following up implementation
	1.3	Consider mandating a 'departmental portfolio officer' role, which can ensure that AAP reviews are conducted for all high-risk projects receiving an "Amber-Red"/"Red" DCA rating
2. Assurance review teams	2.1	Explore the potential for appropriate but not excessive continuity of review team membership
	2.2	Ensure that external reviewer training remains accessible and available
	2.3	Ensure that peer review remains confidential.
3. Assurance review guidelines	3.1	Develop guidelines for the project team and the relevant department to aid in optimal selection of materials for the review team
	3.2	Develop guidelines to ensure that benefits receive as much attention as delivery, in the light of evidence that desired benefits are likely to be delivered only if they are managed throughout the project lifecycle.
	3.3	Guidelines should suggest concentrating recommendations on a manageable number of the most critical issues and discourage the coalescence of different issues within a single recommendation.
	3.4	Guidelines should distinguish those parts of the review that are

		open to discussion from those that are non-negotiable once the review team has completed its review.
4. Continuous improvement	4.1	Continually review and improve the "lessons learned" process, encouraging the widest possible involvement. Learning from the assurance reviews should be fed back into education and training
5. Further research	5.1	Undertake a reflection on the efficacy of the "Review Feedback Forms"
	5.2	Compare the effect of the process for defence projects compared to other GMPP categories
	5.3	Look into the reasons for 'repeated' recommendations
	5.4	Consider the effect of SRO turnover.

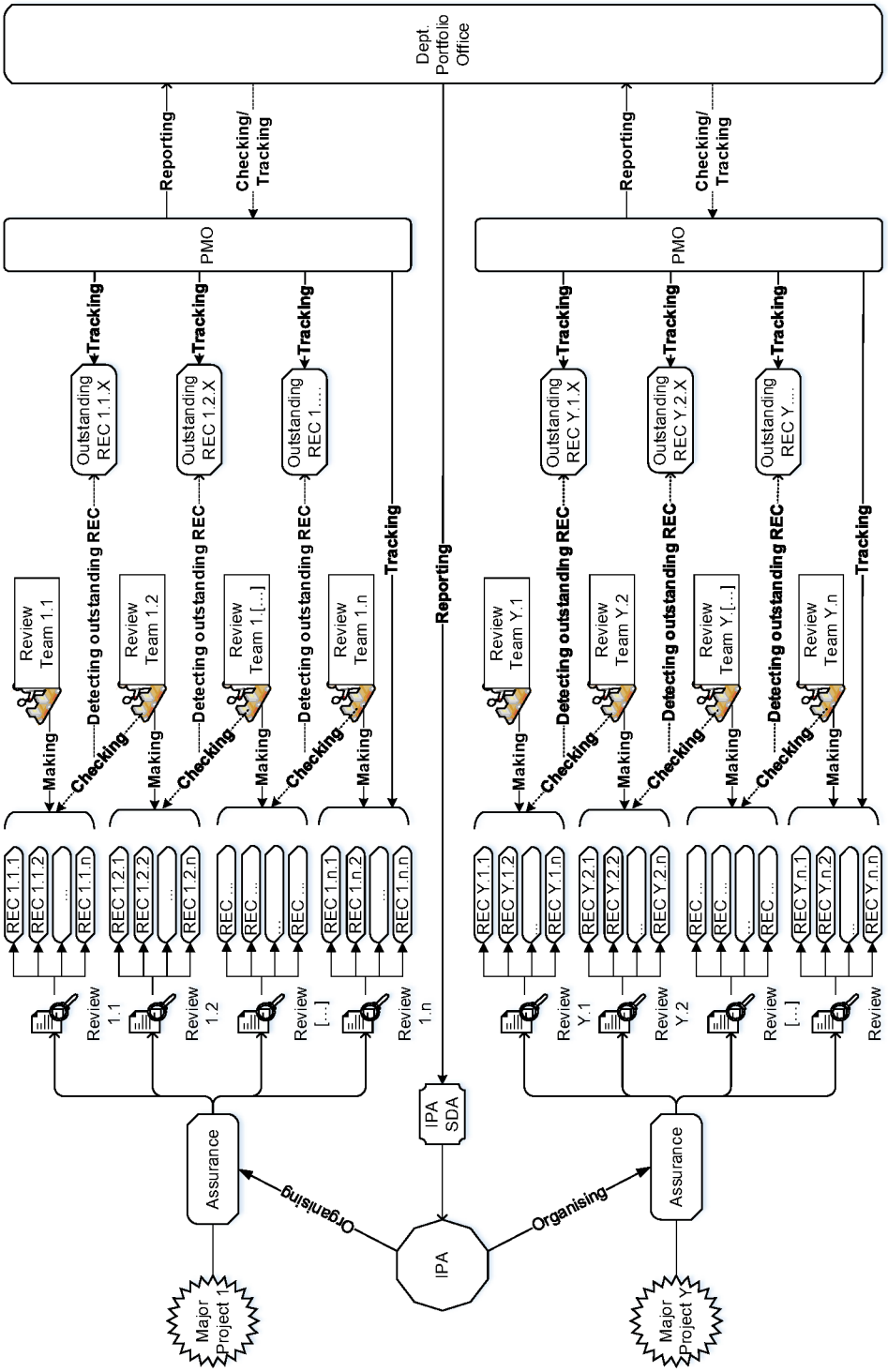


Figure 1 - Relationship between DPO and IPA

References

Aerts, G., Dooms, M., & Haezendonck, E. (2017). Knowledge transfers and project-based learning in large scale infrastructure development projects: an exploratory and comparative ex-post analysis. *International Journal of Project Management*, 35(3), 224-240.

Ahola, T., Ruuska, I., Artto, K., & Kujala, J. (2014). What is project governance and what are its origins? *International Journal of Project Management, 32*(8), 1321-1332.

Ajmal, M.M., & Koskinen, K. (2008). Knowledge transfer in project‐based organisations: an organisational culture perspective. *Project Management Journal, 39*(1), 7-15.

Albert, M., Balve, P., & Spang, K. (2017). Evaluation of project success: a structured literature review. *International Journal of Managing Projects in Business, 10*(4), 796-821.

Anbari, F.T., Carayannis, E., & Voetsch, R. (2008). Post-project reviews as a key project management competence. *Technovation, 28*(10), 633-643.

APM. (2012). *APM Body of Knowledge* (5 ed.). Buckinghamshire: Association for Project Management.

Arnaboldi, M., Lapsley, I., & Steccolini, I. (2015). Performance management in the public sector: The ultimate challenge. *Financial Accountability & Management*, *31*(1), 1-22.

Baldwin, T.T., Pierce, J.R., Joines, R.C., & Farouk, S. (2011). The elusiveness of applied management knowledge: A critical challenge for management educators. *Academy of Management Learning & Education, 10*(4), 583-605.

Baumard, P., & Starbuck, W.H. (2005). Learning from failures: Why it may not happen. *Long range planning, 38*(3), 281-298.

Biesenthal, C., & Wilden, R. (2014). Multi-level project governance: Trends and opportunities. *International Journal of Project Management, 32*(8), 1291-1308.

Bowman, A., Ertürk, I., Folkman, P., Froud, J., Haslam, C., Johal, S., … Tsitsianis, N. (2015). *What a waste: Outsourcing and how it goes wrong*. Manchester, UK: Manchester University Press.

Brown, A.D., & Jones, M. (1998). Doomed to failure: Narratives of inevitability and conspiracy in a failed IS project. *Organisation studies*, *19*(1), 73-88.

Brunet, M. (2018). Governance-as-practice for major public infrastructure projects: A case of multilevel project governing. *International Journal of Project Management*. 37(2), 283-297

Cantarelli, Chantal & Flyvbjerg, Bent. (2015). Decision Making and Major Transport Infrastructure Projects: The Role of Project Ownership. *Handbook on Transport and Development. Edward Elgar, Cheltenham*, 380-393..

Cerpa, N. & Verner, J.M. (2009). Why did your project fail? *Communications of the ACM, 52*(12), 130-134.

Chang, Chen-Yu. (2014). Principal-Agent Model of Risk Allocation in Construction Contracts and Its Critique. Journal of Construction Engineering and Management. 140. 04013032.10.1061/(ASCE)CO.1943-7862.0000779.

Chipulu, M., Ojiako, U., Marshall, A., Williams, T., Bititci, U., ... & Stemati, T. (2018). A dimensional analysis of stakeholder saliences in the assessment of project outcomes. *Production Planning & Control*, 1.

Christianson, M.K., Farkas, M.T., Sutcliffe, K.M. & Weick, K.E. (2009). Learning through rare events: Significant interruptions at the Baltimore & Ohio Railroad Museum. *Organisation Science, 20*(5), 846-860.

Crawford, L.H., & Helm, J. (2009). Government and governance: The value of project management in the

public sector. *Project Management Journal, 40*(1), 73-87.

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. Academy of management Review, 20(1), 65-91.

Douglas, M. (1986). *How institutions think*. Syracuse University Press.

Downham, G. and Lingham, R. (2009). Learning lessons: using inquiries for change. *Journal of mental health law*, 18, 57-69.

Edmondson, A.C. (2011). Strategies for learning from failure. Harvard business review, 89(4), 48-55.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of management review*, 14(1), 57-74.

Fawcett, P., & Marsh, D. (2012). Policy Transfer and Policy Success: The Case of the Gateway Review Process (2001-10). *Government & Opposition, 47*(2), 162-185

Gardiner, P., Söderlund, J., Eltigani, A., Williams, T., Kirkham, R., Ou, L. & Calabrese, C. (2018). *Evolutionary Learning in Strategy-Project Systems*. Newtown Square, PA: Project Management Institute.

Goffin, K., Koners, U., Baxter, D., & Van der Hoven, C. (2010). Managing lessons learned and tacit knowledge in new product development. *Research-Technology Management, 53*(4), 39-51.

Gale NK, Heath G, Cameron E, Rashid S, Redwood S. (2013) Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol. 2013 Sep 18;13:117.

Grabher, G. (2004). Temporary Architectures of Learning: Knowledge Governance in Project Ecologies. Organization Studies, 25(9), 1491–1514. https://doi.org/10.1177/0170840604047996

Hackett, A., & Strickland, K. (2018). Using the framework approach to analyse qualitative data: a worked example. Nurse researcher, 26(3).

Haq, SU, Gu, D., Liang, C., & Abdullah, I. (2019). Project governance mechanisms and the performance of software development projects: Moderating role of requirements risk. *International Journal of Project Management* , 37 (4), 533-548

Hartmann, A., & Dorée, A. (2015). Learning between projects: More than sending messages in bottles. *International Journal of Project Management, 33*(2), 341-351.

HM Treasury (2007). *Transforming Government Procurement*. Retrieved from https://fairsnape.files.wordpress.com/2007/03/government_procurement_pu147.pdf

HM Treasury, & Cabinet Office (2011). *Major Project approval and assurance guidance*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/179763/major_projects_approvals_assurance_guidance.PDF.pdf

HM Treasury (2014) Assurance Frameworks https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/270485/assurance_frameworks_191212.pdf

HM Treasury. (2016). *Treasury approvals process for programmes and projects*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/567908/Treasury_approvals_process_guidance_final.pdf

Hodgson, D.E. (2002). Disciplining the professional: the case of project management. *Journal of Management Studies,* 39(6), 803-821

Hsieh, H.F., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288

Ika, L.A. (2009). Project success as a topic in project management journals. *Project Management Journal, 40*(4), 6-19.

IPA. (2012). *Undertaking Project Assurance Reviews - A briefing note for a Review Team Leader (RTL).* Retrieved from https://www.gov.uk/government/publications/rtl-briefing-note-relevant-documentation

IPA. (2016a). *Guidance for departments and review teams - Assurance of Action Plan (AAP).* Retrieved from https://www.gov.uk/government/publications/assurance-of-action-plan-guidance-and-templates

IPA. (2016b). *Undertaking Project Assurance Reviews: A briefing note for a Senior Responsible Owner (SRO).* Retrieved from https://www.gov.uk/government/publications/sro-briefing-note-relevant-documentation

IPA. (2017a). *A guide to choosing which IPA Assurance Reviews to include in the integrated Assurance and Approvals Plan for your Major Project (v2)*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/638202/guide_to_choosing_an_infrastructure_and_projects_authority_assurance_review.odt

IPA. (2017b). *Annual Report on Major Projects 2016-17*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/629622/IPA_Annual_Report_2017.pdf

IPA. (2017c). *Integrated Assurance Strategies and Integrated Assurance and Approval Plans: A guide to implementing integrated assurance (v2)*. Retreived from https://www.gov.uk/government/publications/implementing-integrated-assurance-for-major-projects

Jensen, Michael C. (2000). "Value Maximization and the Corporate Objective Function," in M. Beer and N. Nohria, Breaking the Code of Change (Boston: Harvard Business School Press), pp. 37-58. Reprinted (2002) as "Value Maximization, Stakeholder Theory, and the Corporate Objective Function," Business Ethics Quarterly 12(2): 235-256

Kirkham, R.J. (2019) Written evidence to the Public Administration and Constitutional Affairs Committee Enquiry into the government's management of major projects. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/public-administration-and-constitutional- affairs-committee/the-governments-management-of-major-projects/written/92642.html

King, A. and Crewe (2014) The Blunders of Our Governments, Oneworld Publications

Mehandjiev, N., Allmendinger, R. and Kirkham, R.J. (2017), Risks of legacy ICT infrastructure in government transformation programmes, Cabinet Office

Klakegg, O.J., Williams, T., Magnussen, O.M., & Glasspool, H. (2009). *Governance frameworks for public project development - estimation*. Newton Square, PA: Project Management Instittute.

Klakegg, O.J., Williams, T., Walker, D., Andersen, B., & Magnussen, O. (2010). *Early warning signs in complex projects*. Newtown Square, PA: Project Management Institute

Kotnour, T., & Vergopia, C. (2005). Learning-based project reviews: Observations and lessons learned from the Kennedy Space Center. *Engineering Management Journal, 17*(4), 30-38.

Levinthal, D.A., & March, J.G. (1993). The myopia of learning. *Strategic management journal*, *14*(S2), 95-112.

Lichtenthaler, U., & Lichtenthaler, E. (2009). A capability ‐ based framework for open innovation: Complementing absorptive capacity. *Journal of Management Studies, 46*(8), 1315-1338.

Madsen, P.M., & Desai, V. (2010). Failing to learn? The effects of failure and success on organisational learning in the global orbital launch vehicle industry. *Academy of Management Journal*, *53*(3), 451-476.

Mahaney, R. C., & Lederer, A. L. (2011). An agency theory explanation of project success. Journal of Computer Information Systems, 51(4), 102-113.

Makridakis, S., Williams, T., Kirkham, R. & Papadaki, M. (2019). Forecasting, uncertainty and risk

management. International Journal of Forecasting. Volume 35, Issue 2, April–June 2019, Pages 641-643 10.1016/j.ijforecast.2018.10.002.

McGrath, SK & Whitty, SJ (2015) Redefining governance: From confusion to certainty and clarity', International Journal of Managing Projects in Business, 8 (4), pp. 755-87

McGrath, SK & Whitty, SJ (2019) Governance terminology confusion in management and project management reference documents, Journal of Modern Project Management, vol. 7(2), no. 20, pp. 146-71.

Meggs, T. (2018). *Delivering Government Major Projects in a Modern Age [Speech]* . Retrieved from https://www.gov.uk/government/speeches/delivering-government-major-projects-in-a-modern-age

Miller, R., & Hobbs, B. (2005). Governance regimes for large complex projects. Project Management Journal, 36(3), 42-50.

Müller, R. (2009). *Project Governance*. Farnham, UK: Gower.

Müller, R., & Lecoeuvre, L. (2014). Operationalizing governance categories of projects. International Journal of Project Management, 32(8), 1346-1357.

Müller, R. and Martinsuo, M. (2015), The impact of relational norms on information technology project success and its moderation through project governance, International Journal of Managing Projects in Business, 8 (1), 154-176.

Müller, R., Turner, J. R., Andersen, E. S., Shao, J., & Kvalnes, Ø. (2013). Ethics, Trust and Governance in Temporary Organizations: Building trust through the governance paradigm. In Proceedings of the 2013 IRNOP Conference (pp. 17-18).

Müller, R., Shao, J., & Pemsel, S. (2016). Organizational enablers for project governance. Newton Square, PA: Project Management Institute.

Musawir, AU, Abd-Karim, SB & Mohd-Danuri, MS (2020) Project governance and its role in enabling organizational strategy implementation: A systematic literature review, International Journal of Project Management, 38(1), 1-16

Nalewaik, A., & Mills, A. (2014). The Path to Assurance: An Analysis of Project Performance Methodologies. *Procedia-Social and Behavioral Sciences, 119*, 105-114.

National Audit Office. (2010). *Assurance for high risk projects*. Retrieved from https://www.nao.org.uk/report/assurance-for-high-risk-projects

National Audit Office. (2012). *NAO Assurance for major projects*. Retrieved from https://www.nao.org.uk/report/assurance-for-major-projects/

National Audit Office. (2016). *Delivering major projects in government: a briefing for the Committee of Public Accounts*. Retrieved from https://www.nao.org.uk/press-release/delivering-major-projects-in-government-a-briefing-for-the-committee-of-public-accounts/

Oakes, G. (2008). *Project Reviews, Assurance and Governance*. Aldershot, UK: Gower.

Office of Government Commerce. (2009). *Managing successful projects with PRINCE2* (2009 ed.). London: TSO.

Ostrom, E. (1990). G*overning the commons: The evolution of institutions for collective action.* Cambridge, UK: Cambridge University Press.

Parker, S. and Skitmore, M. Project management turnover: causes and effects on project performance. *International Journal of Project Management*23, no. 3, (2005): 205-214.

Pemsel, S., & Wiewiora, A. (2013). Project management office a knowledge broker in project-based organisations. *International Journal of Project Management, 31*(1), 31-42.

Pettigrew, A. (2008, July). *Innovative forms of organizing – international perspectives*. Plenary presentation at the PMI reserch Conference, Warsaw, Poland.

Pfeffer, J., & Sutton, R.I. (2000). *The knowing-doing gap: How smart companies turn knowledge into action*. Boston, MA: Harvard Business School Press.

PMI (2013). *A guide to the project management body of knowledge (5 ed.)*. Newtown Square, PA: Project Management Institute.

Riddell, P. (2014, October 20). The updated guidance on the Osmotherly Rules [Web log post]. Retrieved from https://www.instituteforgovernment.org.uk/blog/updated-guidance-osmotherly-rules

Samset, K., & Volden, G. (2016). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management, 34*(2), 297-313.

Srivastava, A. & Thomson, S. B. (2009). Framework Analysis: A Qualitative Methodology for Applied Policy Research. JOAAG, Vol. 4. No. 2

Styhre, A., Josephson, P.E. & Knauseder, I. (2004). Learning capabilities in organisational networks: case studies of six construction projects. *Construction Management & Economics, 22*(9), 957-966.

Taleb, N. N. (2007). The black swan: the impact of the highly improbable. New York, Random House.

Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In C. Teddlie & A. Tashakkori (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: SAGE

Too, E.G., & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. *International Journal of Project Management, 32*(8), 1382-1394.

Trkman, P. (2010). The critical success factors of business process management. International Journal of Information Management, 30(2), 125-134.

Volden G H (2019) Assessing public projects' value for money: An empirical study of the usefulness of cost–benefit analyses in decision-making, International Journal of Project Management Volume 37, Issue 4, May 2019, Pages 549-564

Williams, T. (2008). How do organisations learn lessons from projects—And do they? *IEEE Transactions Engineering Management, 55*(2), 248-266.

Williams T, Vo H, Bourne M, Bourne P, Cooke-Davies T, Kirkham R, Masterton G, Quattrone P and Valette J (2019b) A Cross-national Comparison of Project Benefits Management Practices – The Effectiveness of Benefits Management Frameworks in Application To appear in *Production Planning & Control*

Williams, T., Klakegg, O., Walker, D. H., Andersen, B., & Morten Magnussen, O. (2012). Identifying and acting on early warning signs in complex projects. *Project Management Journal*, *43*(2), 37-53.

Williams, T., Vo, H., Samset, K., & Edkins, A. (2019a) The front-end of projects: a systematic literature review and structuring. Production Planning and Control 30, 14, 1137-1169

Williamson, O. E. (1998). Transaction cost economics: how it works; where it is headed. De economist, 146(1), 23-58.

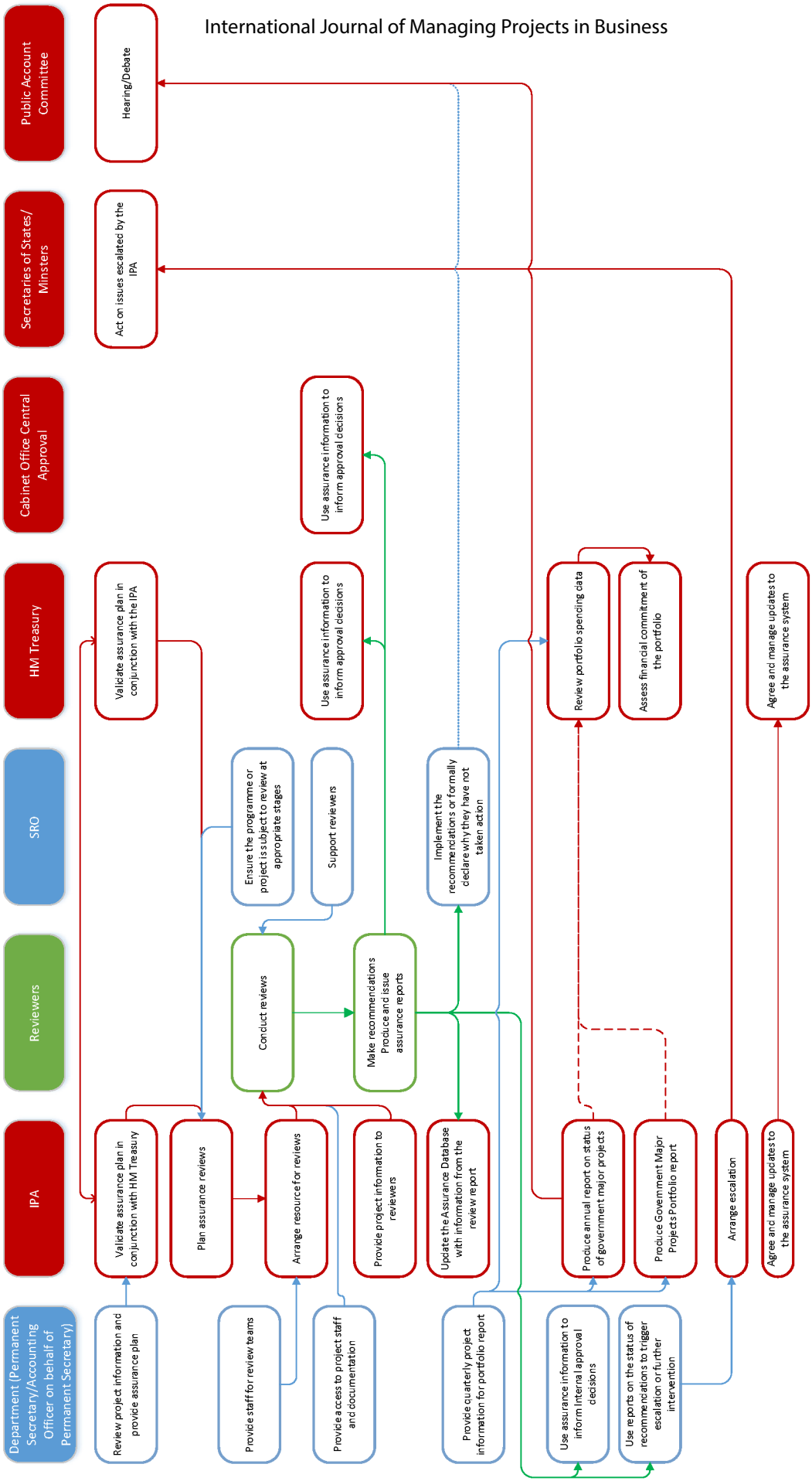
Xu, M., Duffield, C., & Ma, J. (2011). Performance of mid-project reviews (MPRs): quantification based on fuzzy recognition. *Built Environment Project and Asset Management,* 1(2), 137-155.

Xu, M., Duffield, C., Pelham, N., & Bradley, M. (2012). When Do Mid-Project Reviews (MPRs) Deliver the Greatest Project Benefits? Comparative Analysis of Australian Review Outcomes. *Journal of Management in Engineering, 29*(2), 140-149.

Zheltoukhova, K. (2014). Leadership in organisational practice: closing the knowing-doing gap. *Strategic HR Review, 13*(2), 69-74.

Zwikael, O., & Meredith, J.R. (2018). Who's who in the project zoo? The ten core project roles. *International Journal of Operations & Production Management*, *38*(2), 474-492.

Appendix 1: The responsibility of the organisations which are parts of the assurance process for GMPP projects


Public Account Committee
Secretaries of States/ Ministers
Cabinet Office Central Approval
HM Treasury
SRO
Reviewers
IPA
Department (Permanent Secretary/Accounting Officer on behalf of Permanent Secretary)
Hearing/Debate
Act on issues escalated by the IPA
Use assurance information to inform approval decisions
Use assurance information to inform approval decisions
Validate assurance plan in conjunction with the IPA
Review portfolio spending data
Assess financial commitment of the portfolio
Agree and manage updates to the assurance system
Ensure the programme or project is subject to review at appropriate stages
Support reviewers
Implement the recommendations or formally declare why they have not taken action
Conduct reviews
Make recommendations Produce and issue assurance reports
Validate assurance plan in conjunction with HM Treasury
Plan assurance reviews
Arrange resource for reviews
Provide project information to reviewers
Update the Assurance Database with information from the review report
Produce annual report on status of government major projects
Produce Government Major Projects Portfolio report
Arrange escalation
Agree and manage updates to the assurance system
Review project information and provide assurance plan
Provide staff for review teams
Provide access to project staff and documentation
Provide quarterly project information for portfolio report
Use assurance information to inform internal approval decisions
Use reports on the status of recommendations to trigger escalation or further intervention

Appendix 2: Treasury's assurance review process for a major project/programme