



Australian Government

Department of Foreign Affairs and Trade

# 2015 Quality Review of Aid Program Performance Reports

Office of Development Effectiveness

June 2016

# Contents

Executive summary	4
Findings	4
Conclusions and recommendations	6
DFAT management response to recommendations	7
<b>1 Introduction</b>	<b>8</b>
1.1 Aid Program Performance Reports	8
1.2 Coverage of 2015 Quality Review of APPRs	8
1.3 Changes to country and regional program performance management	9
1.4 Methodology	10
<b>2 Quality review</b>	<b>12</b>
2.1 Overall quality	12
2.2 Context and strategic orientation	13
2.3 Progress assessment	15
2.3.1 Gender Assessment	17
2.4 Management responses	18
<b>3 Key observations for learning</b>	<b>21</b>
3.1 Performance benchmarks	21
3.2 Mutual obligations	22
3.3 Performance assessment frameworks	24
3.4 Compliance with guidance	25
<b>Appendix 1: Review Terms of Reference (excerpt)</b>	<b>27</b>
Background	27
Objectives	27
Audience	28
Approach	28
<b>Appendix 2: Method</b>	<b>29</b>
<b>Appendix 3: All ratings assessment criteria</b>	<b>32</b>
Context and strategic orientation	32
Progress assessment	33
Management responses	34
<b>Appendix 4: Comparison between 2013–14 and 2014–15</b>	<b>35</b>

## Office of Development Effectiveness

The Office of Development Effectiveness (ODE) is an independent branch within the Australian Government Department of Foreign Affairs and Trade (DFAT). ODE monitors the Australian aid program's performance, evaluates its impact, and contributes to international evidence and debate about aid and development effectiveness. ODE's work is overseen by the Independent Evaluation Committee (IEC), an advisory body that provides independent expert advice on ODE's evaluation strategy, work plan, analysis and reports.

[www.ode.dfat.gov.au](http://www.ode.dfat.gov.au)

© Commonwealth of Australia 2016 ISBN 978 0 9944202 5 1

Published by the Department of Foreign Affairs and Trade, Canberra, 2016.

Disclaimer: The views contained in this report do not necessarily represent those of the Australian Government.

With the exception of the Commonwealth Coat of Arms and where otherwise noted all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia (<http://creativecommons.org/licenses/by/3.0/au/>) licence. The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence (<http://creativecommons.org/licenses/by/3.0/au/legalcode>). The document must be attributed as *2015 Quality Review of Aid Program Performance Reports*.

## Executive summary

Since 2008, country and regional programs have reported progress against their objectives in annual Aid Program Performance Reports (APPRs). APPRs are published, and serve three main purposes: strengthening program management and effectiveness; providing accountability; and facilitating policy dialogue with development partners.

The Office of Development Effectiveness (ODE), under the guidance of the Independent Evaluation Committee (IEC), provides independent quality assurance of departmental performance reporting. As part of this role, ODE has carried out quality reviews of APPRs every year since they were introduced. These examine the quality of APPRs and highlight strengths and opportunities for improvement.

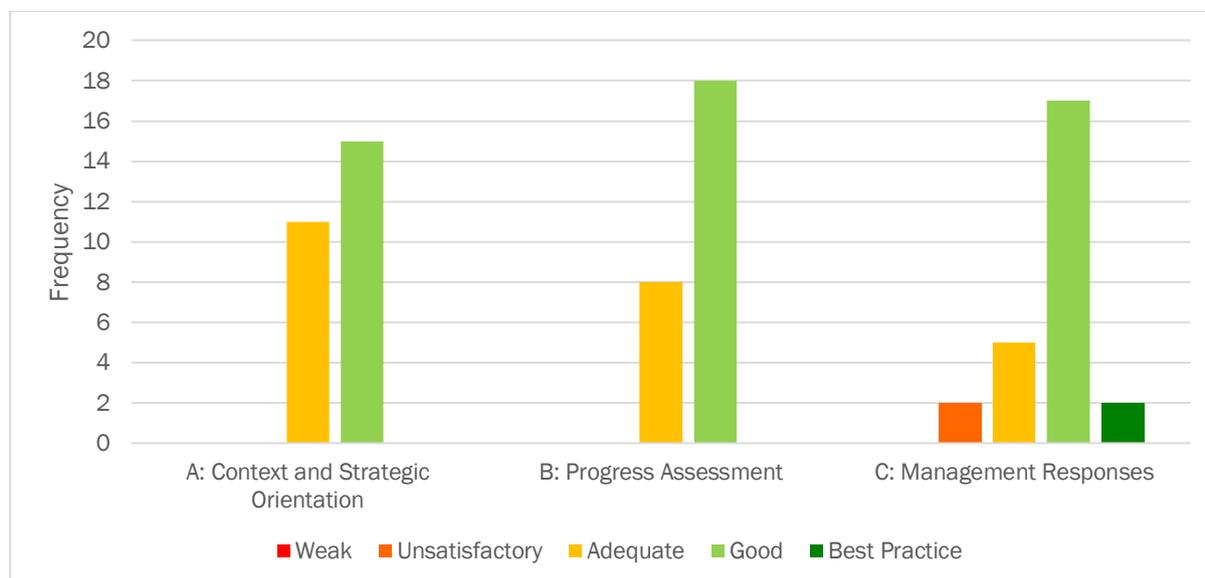
The objective of the 2015 Quality Review of APPRs is to assess and report on the reasonableness of the performance narratives and the use of evidence in the 2014–15 APPRs. The 2015 ODE Quality Review of APPRs assessed the quality of all 26\* of the 2014–15 APPRs. Key findings, conclusions and recommendations follow.

## Findings

### APPR quality is largely adequate or good—consistent with 2014

Against the three quality criteria of context and strategic orientation, progress assessment and management responses, with few exceptions, the 2014–15 APPRs are adequate or good quality (Graph 1).

Graph 1 Overall APPR quality ratings by section



\* 27 APPRs were prepared in 2015, however, while conducting the review ODE was only aware of the 26 APPRs that were published on the DFAT website. As at 04 April 2016 the 2015 APPR for the Mongolia program remains unpublished.

Without exception, the reports were at least adequate in providing a clear and balanced description of program progress within the country or regional context and aid policy framework. This demonstrates the 'embeddedness' of the aid policy across programs.

More than three quarters of APPRs (20) were good quality in terms of how well program objectives were framed and measurable in terms of outcomes.

On the whole, the description of Australian aid's contribution to partner countries was good. However, only a few countries managed to provide a clear description of the donor landscape, including the role of Australian aid within the national and/or regional context. Now that the aid policy is embedded into country and regional programs and as programs move to implement their Aid Investment Plans (AIPs), there is every reason to see clear alignment between AIP objectives and performance benchmarks reflected in APPR reporting.

Overall, the quality of progress assessments in APPRs was good this year. Out of 26 APPR reports, 18 reports were rated good. The remaining eight APPRs were adequate and met minimum standards.

The quality of the management responses in this year's APPRs (that is the actions which DFAT will take to maintain or improve performance in the next reporting period) was generally high and better than the 2013–14 reports. However, the results in this area showed the most variance, with an equal number of reports (2) rated as inadequate and best practice quality.

### **Ratings in APPRs are reasonably well justified overall, with important exceptions**

The APPRs use a green-amber-red colour code to indicate progress towards objectives. Green indicates progress is as expected and it is likely that the objective will be achieved. Amber indicates progress is somewhat less than expected at this stage of implementation and restorative action will be necessary if the objective is to be achieved. Red indicates progress is significantly less than expected at this stage of implementation and the objective is not likely to be met given available resources and priorities. Mostly the traffic light ratings against each objective were well justified: the reports provided a frank, open and balanced discussion and used evidence well to validate their findings. While there was a spread of quality in justifying traffic light ratings (from best practice to inadequate), a high number of reports (6) demonstrated best practice, while 3 reports were inadequate (Papua New Guinea [PNG], Tonga and Nauru).

### **Policy dialogue reporting is improving**

There has also been a general improvement and consolidation of the quality of reporting on policy dialogue. Policy dialogue refers to engagement between the Australian aid program and program partners with the policies of partner governments or other stakeholders.<sup>1</sup> More than half of the APPRs (69 per cent) were rated good or best practice in 2014–15, compared to less than half (42 per cent) in 2013–14. Better reports (for example Bangladesh and Vietnam) provided a detailed understanding of the role of advocacy and policy dialogue, how it shaped progress against intended outcomes and Australia's role in key achievements.

### **Gender equality reporting was strong in several reports**

When assessing gender, 13 APPRs provided a good discussion of progress against some of their outcomes. Timor-Leste, Pakistan, Bangladesh and Afghanistan provided excellent narratives on progress that demonstrated an integrated understanding of how change happens and identified actions to take next year. However there is varied performance in this area.

### **Performance benchmarks met minimum standards but are not yet well integrated**

All programs met reporting requirements for performance benchmarks with half (13) of the programs setting good quality performance benchmarks and just over half (14) providing a considered discussion of progress on last year's performance benchmarks. The performance benchmarks are not yet at a level to confidently assess the performance of each program in meeting the strategic priorities set out in its AIP. Currently the performance benchmarks are not well integrated with

program objectives. Drawing performance benchmarks from the performance assessment framework (PAF) is one way to help increase confidence that they are reflective of the country program and also integrate them more clearly with progress against objectives. A PAF is a management tool for tracking progress towards the objectives of a program. PAFs have been completed by some programs to establish a basis for results management and to provide program managers with better data about performance.

### Performance Assessment Frameworks appear to be helpful for better quality reports

Reports that drew upon a PAF in the APPR were good or better quality. Although having a PAF does not appear to be a determinant of the quality of APPRs (six APPRs were high quality and did not draw on a PAF), all programs with a PAF clearly framed objectives at the outset. PAFs also helped in setting good quality and well-integrated performance benchmarks, discussed above.

### Mutual obligations continue to be of mixed quality

Programs were slightly more effective in identifying clear mutual obligations than they were in assessing progress against them, but overall the quality was mixed in both dimensions. Better practice was evident when programs specified mutual obligations, often sector-specific, beyond simply shared development objectives with the partner country. Weaker practice in mutual obligations was characterised by reporting alignment with partner governments but failing to articulate mutual obligations between the two governments.

### Report page length appears to have a bearing on quality of reports

All reports complied, in large part, with guidance as set out in the APPR template and DFAT's good practice note *Preparing Aid Program Performance Reports*. However, more than half of reports did not comply with specifications for page limits and exceeded the maximum page limit for APPRs as determined by the size of the program in terms of total DFAT official developmental assistance (ODA). Reports' page length does appear to have some bearing on quality. While there was some variability in the quality of those reports that stayed within the requirements for page limits, those reports that went under the page limit were in the lower quality tier and those reports that exceeded their maximum page length were, with one exception, in the upper quality tier. In view of these findings, DFAT should consider increasing the maximum page limit for APPRs.

## Conclusions and recommendations

In 2014–15 APPR quality is largely adequate or good and consistent with the previous year. APPRs exhibit increased embeddedness of the aid policy, clear explanations of policy dialogue engagement and good quality objectives. The APPR narratives also performed well with the quality of their progress assessments.

While ratings in APPRs are mostly well justified, three reports were inadequate in justifying their traffic light ratings. There is also room for improvement in the quality of the management responses, mutual obligations, the integration of the performance benchmarks and reporting performance in addressing gender inequality.

Given the APPRs are of central importance to program management and decision-making, there is a question for DFAT as to whether it should accept the good standard of reporting it has achieved in the last two years or whether it should try to lift the bar.

ODE considers relatively small changes such as requiring performance assessment frameworks and easing the page length restrictions would help improve the quality of APPRs, without significant resource implications.

The 2013–14 APPR Quality Review recommended that *'All country and regional programs that produce APPRs should have, and be supported to develop, high-quality performance assessment frameworks as part of their Aid Investment Plans. Program objectives, performance benchmarks and indicators should all be drawn from these frameworks.'* This recommendation was noted by DFAT, but

PAFs were not required of programs below \$50 million value. On the basis of the 2015 Quality Review, ODE continues to advocate for PAFs for all programs as a means to promote consistently higher quality APPRs.

### 2014–15 Recommendation

1. Improve the integration of Aid Investment Plan objectives, performance assessment frameworks and performance benchmarks (PAFs). Performance against AIP objectives should be measurable through PAFs, from which performance benchmarks should be drawn. All country and regional programs that produce APPRs should prepare and use a high-quality PAF.
2. Change guidance to increase the maximum page limit for APPR reports for all program sizes. Page length should not be a limiting factor for providing a strong, well-evidenced performance narrative.

## DFAT management response to recommendations

### Recommendation 1: Agree in part

ACD agrees with improving the integration of AIP objectives, PAFs and performance benchmarks as far as possible. ACD agrees that performance against AIP objectives can be measurable through PAFs, but notes that other information will also be relevant e.g. Pacific Partnerships for Development. ACD does not agree that all country and regional programs that produce APPRs should prepare a PAF, as PAFs are not mandatory for programs below \$50m.

#### Action:

The 2016 APPR guidance note has been updated to include:

- › the requirement to include PAFs (where available) as an attachment to the APPR report
- › the importance of drawing on PAF milestones and indicators, as well as performance benchmarks, when assessing performance against objectives (June 2016)

The Aid Investment Plan section of the Aid Programming Guide will be updated to include the need for performance benchmarks to be part of PAFs. (August 2016)

### Recommendation 2: Disagree

The ability to draft concise, well-argued reports is a key skill for departmental staff. The current page length limits for APPR reports are sufficient for the drafting of strong, evidence-based program performance reports.

# 1 Introduction

## 1.1 Aid Program Performance Reports

DFAT produces Aid Program Performance Reports (APPRs) for each of the major countries and regions where Australian aid is delivered. They are a central part of DFAT's system for assessing program performance. Annually APPRs assess the performance of regional and country programs against their objectives allowing the Australian aid program to monitor its contribution to change at country and regional levels.

APPRs have three main purposes:

- › **To strengthen program management and improve effectiveness**, serving as management tools to inform program adjustments based on performance.
- › **To ensure accountability and transparency** to the Australian Government and the public, demonstrating how Australian funds are contributing to international development.
- › **To strengthen relationships with partners, especially partner governments and multilateral agencies**, reflecting the extent that the program is working towards mutual commitments and shared development outcomes.

APPRs are peer reviewed, which provides an opportunity for programs to have their assessment of program performance considered by sector specialists, program and quality managers and advisers, and whole-of-government partners. They are also published, contributing to the transparency and accountability of Australia's aid program.

APPRs draw together reporting from Aid Quality Checks, which report at the investment (project) level and against higher order objectives set out in country and regional Aid Investment Plans (AIPs). These include aid investments delivered by DFAT as well as other Australian government departments and agencies. Information from the APPRs provides a basis for DFAT senior management and the Australian Government to understand how the aid program is performing, which in turn informs decision-making and resource allocation.

### Role of the Office of Development Effectiveness in reviewing APPRs

The Office of Development Effectiveness (ODE) provides independent oversight of the effectiveness of the Australian aid program. As part of its work to test and quality-assure DFAT's aid performance assessment systems, since 2008, ODE has carried out quality reviews of APPRs to report on the strengths of the reporting and highlight opportunities for improvement. ODE's reviews are overseen and endorsed by the Independent Evaluation Committee. A response to review recommendations is sought from the DFAT Executive.

## 1.2 Coverage of 2015 Quality Review of APPRs

This report reviews the 2014–15 APPRs against the quality criteria developed in line with APPR guidelines. Comparison with findings from 2013–14's Quality Review is woven through this report. This year's review discusses the quality trends observed in the 2014–15 set of APPRs in terms of:

- › description of overall program progress within the context of the aid policy framework
- › justification of progress ratings against each program objective
- › linkage of management responses to report narratives and identified risks to future progress
- › additional observations for learning about performance benchmarks, mutual obligations, performance assessment frameworks and compliance with guidance.

### 1.3 Changes to country and regional program performance management

Three recent changes in DFAT's program performance management system are highlighted as they had implications for the quality review. AIPs were published on the DFAT website in November 2015 to reflect the June 2014 aid policy *Australian aid: promoting prosperity, reducing poverty, enhancing stability* and the accompanying performance management policy *Making Performance Count*. The AIPs establish objectives for Australian aid to each country or region and outline the strategy by which the aid investments for that country or region will contribute towards realising those objectives. The APPRs report progress annually against achieving those objectives. In addition, they report against performance benchmarks and mutual obligations, also introduced with the June 2014 policy. A short description of each now follows.

#### Aid Investment Plans

The purpose of AIPs is to set the strategic framework for the provision of Australian official developmental assistance (ODA) administered by DFAT and other Australian Government departments and agencies to a country or region. Although they were finalised at the same time as the APPRs were prepared (APPRs were also due in November 2015), they informed the quality review in helping to understand the setting of program objectives within the broader strategic orientation of the country or regional program.

#### Performance benchmarks

One year on from the requirement that country and regional programs establish performance benchmarks, as set out in *Making Performance Count*, 2014–15 is the first year in which programs were expected to report against progress towards these pilot benchmarks. In 2013–14, ODE assessed the quality of the first year of performance benchmarks. For the present review, performance benchmarks are assessed in terms of:

- › their suitability for assessing program performance
- › how well programs have reported on progress towards these performance benchmarks, including highlighting key achievements and challenges and noting any changes that have been made to the pilot performance benchmarks.

#### Mutual obligations

This was the second year in which programs were asked to identify and report on the progress of mutual obligations as set out in Target 6 of *Making Performance Count*.<sup>\*</sup> The 2014–15 Quality Review assessed the extent that programs:

- › identified key mutual obligations for the partner government and Australia

---

<sup>\*</sup> *Making Performance Count*, Target 6: 'progress against mutual obligations agreed between Australia and its key partner governments and organisations will form part of program performance assessments... [Moving] away from traditional donor recipient relationships to more mature partnerships. Greater mutual accountability is critical in making this transition. While Australia will be held to account for the performance of its investments, partner governments and organisations also need to contribute to shared development goals...The mutual obligations of partner governments and organisations will generally reflect their own reform commitments of particular importance to their development and Australia's aid investments.'

- › reported on progress made by the partner government and Australia against the mutual obligations identified in last year’s APPR. Particularly with a view towards progress that demonstrates a mutual commitment to obligations and that Australia is delivering on its part of the obligations.

APPRs were guided by DFAT’s good practice note and the APPR template.

## 1.4 Methodology

The objective of the 2015 Quality Review of APPRs is to assess and report on the reasonableness of the performance narratives and the use of evidence in the 2014–15 APPRs, and on the reasonableness of the performance narratives against the country and regional program-level performance benchmarks proposed in the 2013–14 APPRs.

To inform the review, evidence has been gathered from all of the 26 2014–15 APPRs (Table 1), corresponding AIPs, available performance assessment frameworks (PAFs), the most recent APPR guideline, template and previous ODE APPR quality reviews and working papers.

Table 1 2014–15 APPRs assessed in this quality review

APPR type	Program
Country APPRs	Afghanistan, Bangladesh, Cambodia, Fiji, Indonesia, Kiribati, Laos, Myanmar, Nauru, Nepal, Pakistan, Palestinian Territories, Papua New Guinea, Philippines, Samoa, Solomon Islands, Sri Lanka, Timor-Leste, Tonga, Vanuatu, Vietnam
Regional APPRs	East Asia, Pacific, South Asia, Sub-Saharan Africa
Other APPRs	Australian NGO Cooperation Program (ANCP)

The quality review was structured by three main assessment criteria, articulated as ‘cornerstone questions’, with 15 subquestions.

1. **Context and strategic orientation:** Does the APPR provide a clear and balanced description of overall program progress within the country/ regional context and aid policy framework?
2. **Progress assessment:** Does the APPR narrative explain and justify progress ratings against each program objective?
3. **Management responses:** Are the management responses clearly linked to the ratings narratives and the identified risks to future progress?

A fourth quality criterion around the contribution of partners was dropped from this year’s quality review because partner performance is now measured through alternative systems introduced by *Making Performance Count*. International, private sector and civil society partner performance is assessed annually by programs, rating performance against criteria in a separately managed partner performance process.

Each APPR was reviewed and rated (with comments) against these criteria according to the same rating scale as last year (Figure 1), following an extensive moderation process with the review team members. They were reviewed against last year’s assessments to maximise comparability. APPRs that complied with current guidance received a rating of at least 3 (adequate) rather than 2 (unsatisfactory). Examples of good practice, as well as common weaknesses in all areas of APPR reporting, were identified. A full description of the methodology is in Appendix 2.

Figure 1 APPR quality rating scale

Ratings	5	4	3	2	1
Explanation of ratings	Example of best practice	Good	Adequate / complies with guidelines	Unsatisfactory / below expectations	Unsatisfactory/ weak

Note: Ratings 3 and above indicate adequate quality; a rating of 3 is given where there has been compliance with the guidelines; ratings of 4 indicate that additional information and interpretation has been given to enhance the quality area; a rating of 5 reflects high quality in the given area. Ratings of 2 or below indicate that the quality is unsatisfactory.

### Structure of report

The body of the quality review is dedicated to the quality assessment in Chapter 2. The narrative is broken down into three subsections as per the main assessment criteria (cited above). Each of these subsections is structured in the same way: it opens by listing the cornerstone and subquestions, then provides a description of what is covered by that assessment criteria, presents the high-level findings and finally summarises how better reports tackled the assessment. Examples from various APPRs are peppered throughout the chapter in boxes to illustrate key findings. Appendix 3 contains the ratings of each subquestion and Appendix 4 sets out critical comparisons with last year's APPRs.

Chapter 3 sets out additional observations to learn from. This year the report looks at performance benchmarks, mutual obligations, performance assessment frameworks and compliance in a little more detail.

## 2 Quality review

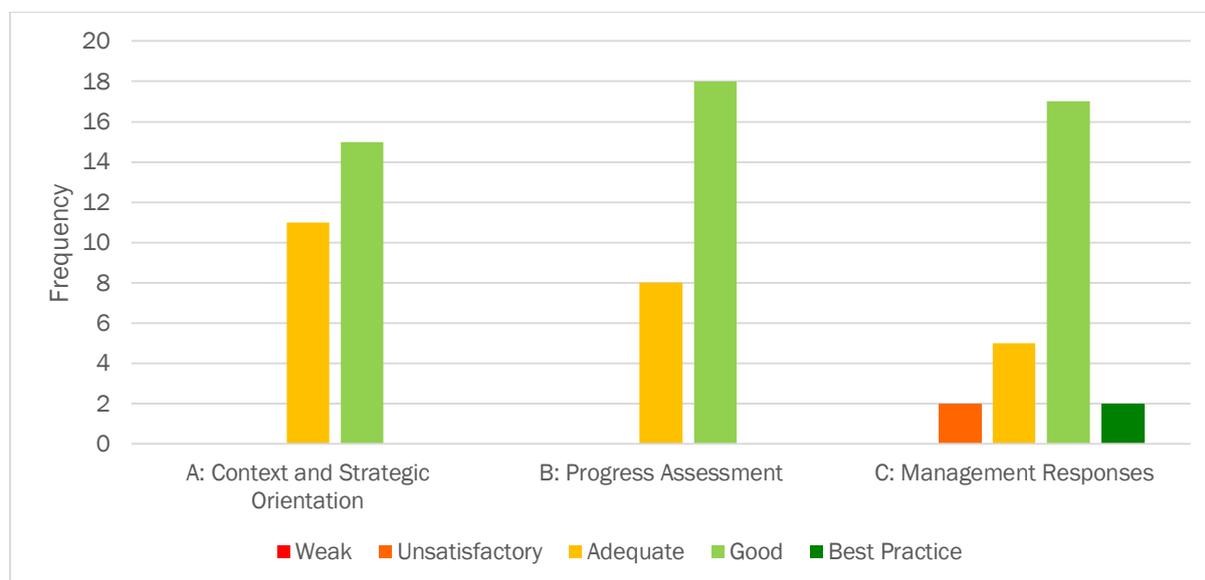
### 2.1 Overall quality

The quality of the 2014–15 APPRs is assessed against the three major quality criteria: country and strategic context, progress assessment and management responses. The following sections of this report set out the key findings by criterion.

The 2014–15 APPRs are consistent with those produced in 2013–14 in terms of the quality of the reports. There are fewer best practice ratings, but more ‘good’ ratings on some criteria. Almost all APPRs were rated as either good or adequate in all sections (Graph 2). Without exception, the reports produced an adequate or good description of program progress within the country or regional context and aid policy framework. In 2014–15, a strong majority (69 per cent) of APPRs provided good quality progress assessments. The continued quality of the management responses in the APPRs from previous years suggests that they are being used effectively to manage and mitigate risk. Almost all of the APPRs provided well-thought-out management responses that addressed relevant issues and risks to future progress.

Graph 2 shows the overall quality ratings for the 26 2014–15 APPRs against the three quality criteria. Each criterion is discussed in turn, beginning with context and strategic orientation.

Graph 2 Overall APPR quality ratings by section



## 2.2 Context and strategic orientation

<b>Subquestion 1a</b>	<b>Does the APPR provide a clear and balanced description of overall program progress within the country/regional context and aid policy framework?</b>
<b>Subquestion 1b</b>	Does the APPR respond to and reflect Australia's aid policy?
<b>Subquestion 1c</b>	Does the report identify key mutual obligations for the partner government and Australia that contribute to shared development goals?
<b>Subquestion 1d</b>	Does the report contain performance benchmarks that provide a reasonable basis for assessing program performance?
<b>Subquestion 1e</b>	Does the discussion of the program progress reflect on how well the program strategy is working to achieve the program objectives?
<b>Subquestion 1e</b>	Does the report make clear Australia's contribution and the proportion and significance of Australia's aid compared to other donors and national budgets, and the program and strategy implications of that?

This section looks at how well the APPRs:

- › described progress within the country or regional context
- › reflected Australia's aid policy
- › identified mutual obligations for the partner government and Australia
- › included reasonable performance benchmarks
- › discussed how well the program strategy is working
- › explained the size and significance of Australian aid within the country or regional context.

Without exception, the reports produced an adequate or good description of program progress within the country or regional context and aid policy framework. Quality was generally consistent with last year's review; the majority (15 out of 26 reports) exceeded minimum standards with a 'good' rating in their discussion of program progress within the country or regional context.

Better reports described the context and strategic orientation of the program by:

- › **Reflecting the Australian aid policy, with evidence of rationalisation and adjustments to align the program with that of the Australian aid policy.** This specifically included reflecting progress in promoting innovation, private sector engagement, gender equality, the empowerment of women and girls, and the promotion of disability-inclusive development. With the introduction of Aid Investment Plans, programs appear to have responded better to the Australian aid policy than last year, having integrated processes for more effective aid and translated Australian priorities into practice through the program.
- › **Clearly identifying key mutual obligations for the partner government and Australia that contribute to shared development goals.** Better reports demonstrated how Australia and the partner government have collaborated towards shared development outcomes and how the aid program has aligned itself to the partner government's development priorities. Nonetheless, there was a spread of results against this criterion. Examples of good practice are Bangladesh, Philippines, and Vanuatu. Three reports (Nauru, Pakistan and Sri Lanka) failed to explicitly state the mutual obligations. Reporting of mutual obligations is discussed more fully in Chapter 3.
- › **Establishing reasonable performance benchmarks designed to reflect the main focus of the program and also to assess program performance.** In 2014–15, all performance benchmarks met standards, with half providing good, rather than adequate, quality performance benchmarks that look at a range of performance measures and are built to realistically assess progress rather than being constructed to be easily achievable. Performance benchmark quality is discussed more fully in Chapter 3.

## Box 1 Bangladesh: Clear articulation of mutual obligations

Mutual obligations for the Bangladeshi government and Australia's contribution towards these are clearly described in the report. For example, Australia's efforts to align itself to Bangladesh's development priorities are described.

*'We have strategically deployed our limited resources to focus on key areas which align with the Government's Sixth Five Year Plan, Australia's aid policy and our new country program goal of "investing in human development, skills and productivity".'*

It is evident from the report that the Australian aid program has a mature understanding of the needs of the Bangladeshi government and how it will contribute to these. Indeed, a Statement of Commitment on development cooperation was signed by both governments in 2013 and it aligns with the Australia–Bangladesh program strategy.

- › **Clearly discussing how well the strategy is working to achieve the program objectives in terms of the underlying logic of what Australian aid does and how this contributes to achieving key outcomes.** Best practice is demonstrated by those programs that articulate a clear understanding of the 'problem' the program is addressing, its strategic orientation, and a mature understanding of the context and any changes in the context over the reporting period. Where appropriate, reports reflected on the AIP and the strategic thrust therein. All but one program (Tonga) provided an adequate or good discussion of the strategy, helping to inform the overall context and strategic orientation of the program.

## Box 2 Myanmar: Shifting strategic orientation

According to the APPR, it is a year of transition for the program, from one strategy to another (AIP). The report is open and frank about how the strategy is changing, why, and the implications of this on program progress.

*'Australia's aid program is in transition as we respond to the rapidly changing context in Myanmar and implement a new strategy, the Myanmar Aid Investment Plan 2015–19. While our ability to respond flexibly is central to successful aid delivery at this juncture, that flexibility creates challenges for reporting against the program's performance assessment framework (PAF). The Australia–Myanmar Aid Program Strategy 2012–15 PAF guided performance reporting at a high level, but some indicators are no longer relevant where our programs have shifted focus. There is a consequent disconnect between the PAF and programming realities. Further, in some instances we have set performance benchmarks or PAF indicators that pivot on the actions of the Myanmar Government rather than Australia's actions or contributions. Australia will aim to improve performance reporting in 2015–16.'*

- › **Articulating the contribution of Australian aid** in comparison with other donors, providing the ratio of Australia's ODA to partner government GDP where possible. Most reports discussed the role of Australian aid within the national context and provided a clear description of the donor landscape. Programs provided stronger narratives about the contribution of Australia aid over the last year. Two reports however, both regional programs (Pacific and Sub-Saharan Africa), did not adequately set out and contextualise Australia's contribution to the regions.

### Box 3 Vanuatu: Clear explanation of Australia's role

Australia's role as lead donor in Vanuatu is clearly explained, including expenditure, disaggregated by objective, and the relative contribution to total GDP. The relatively limited presence of other key donors is also narrated, indicating their sector contribution. The report clearly describes the donor landscape of Vanuatu, framed within the wider context of Australia's strategic interests there.

*'We remain by far the largest provider of aid, equivalent to around 30 per cent of total public spending and to around 6 per cent Vanuatu's GDP. Australia is the largest source of foreign direct investment (\$169 million in 2013–14) and tourists (over 200,000 a year) and the 10th ranked destination for Vanuatu exports. We are the lead donor in the health, education, governance, and justice and policing sectors. Australia was the first and largest contributor to the cyclone response, committing \$50 million for humanitarian, early and long-term recovery aid. Vanuatu has few other long-term country partners. New Zealand supports infrastructure, education, water and tourism. China and Japan have a growing presence in infrastructure. The World Bank and Asian Development Bank have increased their engagement in Vanuatu following the cyclone; their contributions to recovery, through both existing and new activities, will be important.'*

## 2.3 Progress assessment

	<b>Does the APPR narrative explain and justify progress ratings against each program objective?</b>
<b>Subquestion 2a</b>	Are the objectives sufficiently framed at the outset or in the narrative as statements of intent that are measurable (quantitatively and/or qualitatively)?
<b>Subquestion 2b</b>	Does the narrative explain and justify the progress ratings against each objective?
<b>Subquestion 2c</b>	Does the report assess progress made by the partner government and Australia against the mutual obligations identified in last year's APPR?
<b>Subquestion 2d</b>	Does the report provide evidence about progress against performance benchmarks, and where benchmarks have not been achieved, discuss why this is the case and what management actions will be taken?
<b>Subquestion 2e</b>	Where appropriate, is the policy dialogue engagement with partners clearly explained?
<b>Subquestion 2f</b>	Is there an appropriately open, balanced discussion of the program's successes, failures and challenges?
<b>Subquestion 2g</b>	Has relevant evidence been drawn from an appropriate range of internal and external sources given the program context?
<b>Subquestion 2h</b>	Does the narrative justify the progress and results in promoting gender equality and empowerment of women and girls? Are trends and planned follow-up actions identified?

In the APPRs, programs rate progress towards achieving objectives using a green, amber, red 'traffic light' system.\* This requires a clear framing of the intent of objectives, a narrative that explains the progress ratings, an assessment of progress towards mutual obligations, discussion of policy dialogue engagement, frank and balanced discussion of performance, use of evidence from a range of sources, discussion of the progress and results in promoting gender equality, and evidence about progress against performance benchmarks.

\* In which green = 'Progress is as expected for this point in time and it is likely that the objective will be achieved. Standard program management practices are sufficient'; amber = 'Progress is somewhat less than expected for this point in time and restorative action will be necessary of the objective is to be achieved. Close performance monitoring is recommended'; and red = 'Progress is significantly less than expected for this point in time and the objective is not likely to be met given available resources and priorities. Recasting objective may be required.'

In 2014–15, a strong majority (69 per cent) of APPRs provided good quality progress assessments. While there were no best practice ratings this year, more APPRs were rated good on justification of progress assessments. The overall picture is similar to last year.

Better reports provide a well-evidenced progress assessment through:

- › **Explicit framing of each objective.** Well-framed objectives suggest that there is a good foundation for a strong discussion of progress. More than three quarters of APPRs (77 per cent or 20 APPRs) were deemed to be good quality when assessed for how well program objectives were framed and measurable in terms of outcome. This is a small improvement from last year, when 71 per cent were found good quality or better on this subquestion.
- › **Clear justification of progress ratings against objectives.** In general, as in the 2013–14 review, most of this year’s APPRs clearly explained the basis for the ratings, often in an opening statement in the performance discussion under each objective. While there was a spread of ratings in terms of justifying the traffic light ratings against each objective, from inadequate to best practice, a high number (6) were considered best practice. A number of programs provided excellent narratives written at a ‘portfolio’ level, describing systemic changes (rather than progress against project level outputs), which were then validated. These APPRs also provided an excellent critical reflection on what worked well, why, and what worked less well and why (Box 4).

#### Box 4 Best practice in justifying ratings

**Pacific Regional** APPR ratings were justified in the narrative with a nuanced assessment of changes from the 2013–14 APPR and a clear diagnosis of the reasons for the assessment. For example, in Objective 1 the narrative states: *‘The rating on economic growth increased from a very high amber rating to a low green. The small improvement tipped it over to the next level and was largely the result of improvements in monitoring and evaluation of investments and improved efficiency.’*

**South Asia Regional** APPR displayed a strong justification and supporting statements with evidence of the interim steps towards outcomes for each of the countries within the region. An in-depth understanding of the process required for achievement of the objectives was demonstrated in the narrative. Major outputs were contextualised by how well they contribute to overall performance with reference to the PAF’s articulation of performance assessment.

**Sri Lanka’s** APPR narrative provides a clear and explicit basis for understanding the progress ratings against each objective, referencing the targets that have been identified through the PAF. Progress is explained with reference to the underlying program logic, and then the health, education and economic development targets that have been identified as performance evidence.

- › The basis for the ratings in the better APPRs was **backed up with evidence that went beyond project evidence** to contextualised evidence and results that allowed the reader to understand their significance.
- › **A comprehensive discussion of progress towards mutual obligations.** Fourteen of the 26 APPRs fell into the good or best practice categories. See full discussion of mutual obligations in Chapter 3.
- › **An in-depth description of policy dialogue with the partner government** discussing how this has influenced progress towards objectives, describing successes and challenges. APPRs were not judged on the extent of progress in this area, but rather on how clearly and extensively they discussed progress, influence and strategic engagement through policy dialogue with a relevant understanding of context (political, economic, social, cultural etc.). Excellent examples included a detailed understanding of the role of advocacy and policy dialogue in the overall program, how it shaped progress against intended outcomes and Australia’s role in key achievements (Box 5). There was a general improvement in the strength of the policy dialogue discussion compared with last year’s APPRs. A strong majority

of the APPR portfolio (69 per cent) was rated good or best practice in 2014–15, compared to less than half (42 per cent) in 2013–14.

## Box 5 Bangladesh and Vietnam's policy dialogue

**The Bangladesh** APPR provides a detailed account of policy dialogue with key program partners, clearly discussing the intended outcome of these engagements (e.g. to address child marriage, promote social protection policy, etc.). Australia's level of influence and role in policy processes is also clearly set out in the report. For example, Australia played a lead role in influencing the policy and management priorities for primary education sector reform in 2014–15. Australia encouraged collaboration across the primary education sector between government and donor partners, including by taking a lead role in a study tour to Tamil Nadu in India to study inclusive techniques for quality teaching.

**Vietnam** APPR clearly sets out Australia's policy dialogue with key partners, the Vietnam Government and the World Bank in particular. Explanation of the purpose of the policy dialogue is provided with examples of the kind of engagement and what it achieved. This is an example of best practice in terms of the clarity and level of attention given to this area.

- › **Performance benchmarks that were demonstrably cohesive with the program strategy and tied to the AIP.** At this early stage of integration into the APPR, this was relatively rare. Other rare good practice features included a strong gender orientation in the performance benchmarks and a commitment to baselines and targets. A fuller discussion can be found in Chapter 3.
- › **Open and balanced discussion of progress** with a frank assessment of performance. Stronger reports demonstrated critical thinking on failure and success, clearly articulating delays, challenges or successes of the program, and how Australia's role has helped or hindered progress. Most reports maintained good quality in this area, providing a frank and balanced assessment of performance. This is comparable to the 2013–14 assessment which found '*[m]ost APPRs provide a frank and balanced assessment of performance, with only a few painting an overly rosy picture*'.
- › **Use of a range of relevant sources of evidence.** Overall, the APPR portfolio provides better use of evidence than it has previously. APPRs generally verify and validate results from both external and internal sources.

### 2.3.1 Gender Assessment

In 2014–15, half (50 per cent) of APPRs provided good quality progress assessments on gender and 4, or 15 per cent, provided best practice assessments. In summary, it was found that there is a discussion of progress and results in promoting gender equality and empowerment of women and girls across the majority of APPR narratives with some considerations of gender in future actions and programming. However, this high standard is not consistent across the APPR portfolio, with several programs only meeting minimum standards and two programs not providing adequate gender narratives.

Better reports provided a discussion of gender dimensions that:

- › **integrated gender** throughout the program progress narrative and through the discussion against all objectives
- › **included gender analysis** of how change happens and, from this, identified actions to take in the following year. A good practice example is presented in Box 6.

## Box 6 Timor-Leste example of gender narrative

There has been progress made in the **Timor-Leste program** and evidence of results in promoting gender equality and empowerment of women and girls are discussed against each of the eight program objectives. The APPR also clearly articulates the contextual challenges relating to gender and how these are targeted through the program objectives.

Gains made in promoting gender equality are described in detail, with an explanation of the starting point and level of achievement through the program interventions. The articulation of the efforts of the program to contribute to gender equality are commendable, particularly in the disaggregation of program elements for gender dimensions (e.g. development of gender policy, appointment of female leaders, consideration of gendered aspects of health and education.) The program's efforts to support results for women and girls is evidenced by the introduction of Objective 9 to reduce violence against women and improve access to support services for women and children.

## 2.4 Management responses

### Subquestion 3a

**Are the management responses clearly linked to the ratings narratives and the identified risks to future progress?**

Is there a logical connection between the management responses section and the risks to future progress flagged in the other sections of the APPR?

### Subquestion 3b

Are the specific management responses framed so they are relevant to program strategy and management and are also actionable within the reporting period? Are they also prioritised?

APPRs are meant to provide evidence-based proposals for management action and signal management responses to significant issues. These are framed as 'management responses' in the reports and are interpreted as the consequences that the evidence reported has had on management decision-making. Assessment of the management responses considered how well the reports provided evidence of linkage between management responses and risk to progress and the relevance of the management responses.

Overall the quality of this year's management responses marks a slight improvement from last year. Programs improved their articulation of risks and linked this to management responses, but there was some variation in effectively setting out management responses. Most reports effectively used the management responses section to address risks to progress and to draw out performance issues highlighted in the text.

Better reports demonstrated that the APPR was being used as a management tool through:

- › **Linking management responses and risks to progress** in terms of alignment with the issues raised in the wider narrative. Effective management responses address significant or emerging challenges to achieving the program objectives and opportunities or risks associated with the program or with partners in the forthcoming 12 months. One program (ANCP) provided an example of best practice by including a considered analysis of risks to progress (Box 7). For learning purposes, the risk matrix has also been included below.
- › **Relevant management responses** that set out SMART (Specific, Measurable, Achievable, Relevant, Timebound) management responses. Better reports reflected a strong understanding of the nature of the program and its context, and made recommendations that were appropriately targeted and actionable. All programs met standards in this area with a number demonstrating best practice in their use of the APPR as a management tool. These examples of good practice came from ANCP, Cambodia, Nepal, Myanmar (Box 8) and Indonesia.

## Box 7 ANCP: Best practice example of risk management

Risks associated with the delivery of ANCP projects are largely managed through accreditation and consultation with Posts and through thematic areas on proposed new and higher risk projects. DFAT's internal audits target a number of Australian NGOs each year and head agreements and funding orders outline the terms and conditions of ANCP funding, helping to mitigate fraud risks. Where fraud cases do arise, these are actively managed by DFAT's NGOs and Volunteers Branch (NVB) in consultation with the Department's Fraud Control Section. DFAT webinars on fraud and corruption are held annually for Australian NGOs.

The Aid Quality Check found that while risk is generally managed well in the ANCP, there are areas for improvement, specifically the need for systematic capture and review of risks. NVB has since introduced a quarterly review and update of ANCP's risk register to capture any emerging risks.

A pervasive risk is the limited resources available to DFAT to appropriately manage the ANCP, including DFAT's ability to review evidence, undertake site visits, verify NGOs' self-reports and conduct data analysis. NVB will continue to review resources and the allocation and prioritisation of work to mitigate these risks.

Key risks	What actions were taken to manage the risks over the past year?	What further actions will be taken to manage the risks in the coming year?	Risk rating (low, medium, high, very high)
A fluctuating aid budget poses challenges to ANCP's funding policy, which was made more complex in an environment of rapidly expanding funding	Budget updates; new accreditation application guidelines shared with the ANGO sector. Consultations on funding principles in September 2014–15 at Annual Reflections	Development of transparent funding policy based on identified principles. Clear communication about the process and outcome	High
Inability to effectively analyse ANCP data because of resource constraints	Senior management matched priorities to resources and was clear about limitations	Continue to clarify the role of NVB and engage with Posts, NGOs and ACFID on resource constraints and identifying opportunities to share data analysis	High
Poor data quality affecting ability to provide accurate picture of ANCP	Working with relevant ANGOs to transition to July–June reporting period. Webinars providing examples of good reporting	Continue regular webinars on annual development plans and performance reporting. Updates to guidance and ANCP Online forms	Medium
Lack of alignment between ANCP funding and DFAT country and sector priorities	All new and higher risk projects sent to relevant Desk/Post for comment prior to approval. Desks/Posts notified when annual performance data becomes available. NGO roundtable meetings held at Post	Development and distribution of guidance for Desk/Post regarding the ANCP and their role. NVB will encourage Posts to hold regular roundtables with NGOs in country	Medium

## Box 8 Myanmar: Good practice in clustering management responses

The responses in this APPR are categorised as strategic or operational. This greatly helps the framing. All are actionable and by reading them, it is clear that while they are not numbered, they are prioritised. From 2014–15, Australia’s aid program will:

### Strategic

- › Mitigate political risk by working closely with the next Myanmar Government to implement Australia’s aid program.
- › Closely monitor and report on developments in the peace process and recalibrate our investments as necessary.
- › Build closer links with the private sector, including by using design and evaluation processes as opportunities to identify entry points.
- › Continue to strengthen program performance on gender equality and women’s economic empowerment by including gender as a key evaluation question for all reviews in 2015–16.

### Operational

- › Improve program performance reporting by developing a targeted performance assessment framework for the Myanmar Aid Investment Plan 2015–19 that explains Australia’s contribution, sets targets to assess adequacy of progress, and is reviewed six-monthly to maintain relevance.
- › Explore opportunities to balance humanitarian assistance with a greater development focus on education and livelihoods opportunities, particularly in Rakhine State.
- › Support the Ministry of Education’s monitoring and evaluation systems to improve data collection and its use in planning, including through the design of a new investment.
- › Work with implementation partners of livelihoods programs to improve program quality and performance and revisit progress against review recommendations in six months.

## 3 Key observations for learning

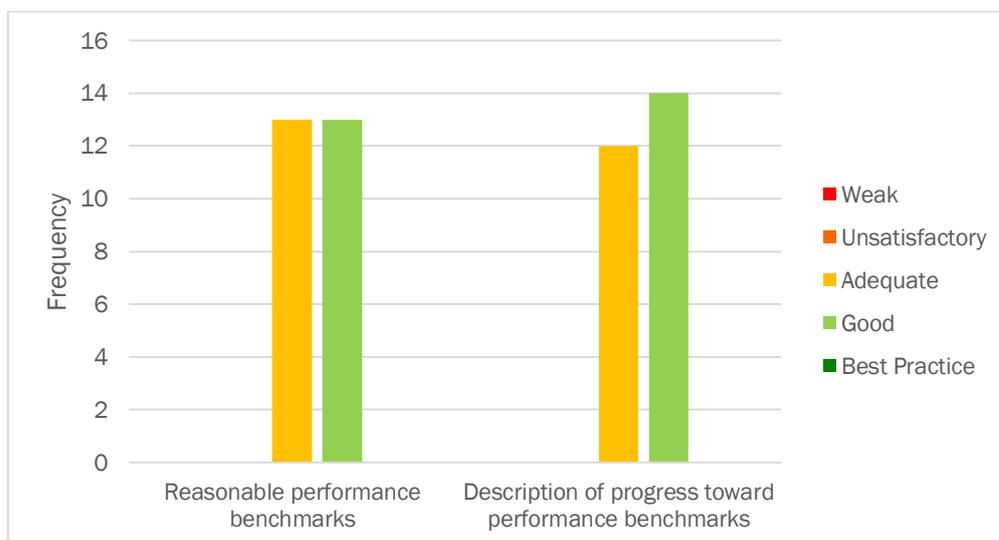
This chapter sets out further analysis and conclusions against four areas: performance benchmarks, mutual obligations, performance assessment frameworks, and compliance with guidance.

### 3.1 Performance benchmarks

2014–15 is the first year APPRs include an assessment of progress against performance benchmarks, therefore it is important to look more closely at the results. In 2013–14, when performance benchmarks were piloted, it was found that they needed to improve to become more robust indicators of program performance and to be useful for informing budget decisions.

In 2014–15 there was an improvement, with all programs meeting reporting requirements for performance benchmarks. Half (13) of the APPRs set good quality performance benchmarks and just over half (14) provided a considered discussion of progress against last year’s pilot performance benchmarks (Graph 3).

Graph 3 Frequency of ratings for performance benchmark questions



Performance benchmarks need to challenge a program and be set at an appropriate level. They should include outcomes as well as outputs. The stronger performance benchmarks (in terms of providing a reasonable basis for assessing program performance) had the following characteristics:

- › Clear and SMART (Specific, Measurable, Achievable, Relevant, Timebound)
- › Reflective of sector spend or priority areas for the country program
- › Pitched at a level that enables assessment of DFAT’s contribution
- › Well balanced by type, weighted to outcome indicators

- › Drawn from the performance assessment framework, which links them to a trajectory of change over time

In some cases, performance benchmarks were well integrated with the program strategy and tied to the AIP, but this was relatively rare. Other rare good practice features included a strong gender orientation in the performance benchmarks and a commitment to baselines and targets. By contrast, the weaker performance benchmarks lacked specificity, were set at too high or too low a level, often related to output or activity level rather than outcomes, were too numerous and difficult to measure and were not aligned to sector priorities or country program objectives.

While there is an improvement in the pilot performance benchmark narratives from last year, the performance benchmarks are not yet at a level to confidently assess the performance of each program in meeting the strategic priorities set out in its AIP. A significant issue is the lack of integration of the performance benchmarks with program objectives. Drawing performance benchmarks from the PAF is one way to help increase confidence that they are reflective of the country program and also helps to integrate them more clearly with progress against objectives.

Graph 3 also illustrates that just over half (14) programs provided a good report on last year's performance benchmarks while 12 reports were adequate in this regard. Programs in the adequate rating tended to provide minimal discussion and some had difficulty in justifying progress or providing evidence that benchmarks had been achieved. One manifestation of this was a mismatch between the objective ratings and the performance benchmark report in the same sector. Another example was a lack of consistency between the articulated steps to rectify performance benchmark progress and the real performance of programs discussed elsewhere in the APPR.

The APPRs that demonstrated poorer quality pilot performance benchmarks last year have all met minimum standards in reporting this year. Considering the brevity of the report, Kiribati's detail of progress against performance benchmarks was good. Assessment included well-considered and balanced judgements of progress with supporting evidence.

Table 2 Kiribati APPR: Extract from assessment of progress against performance benchmarks

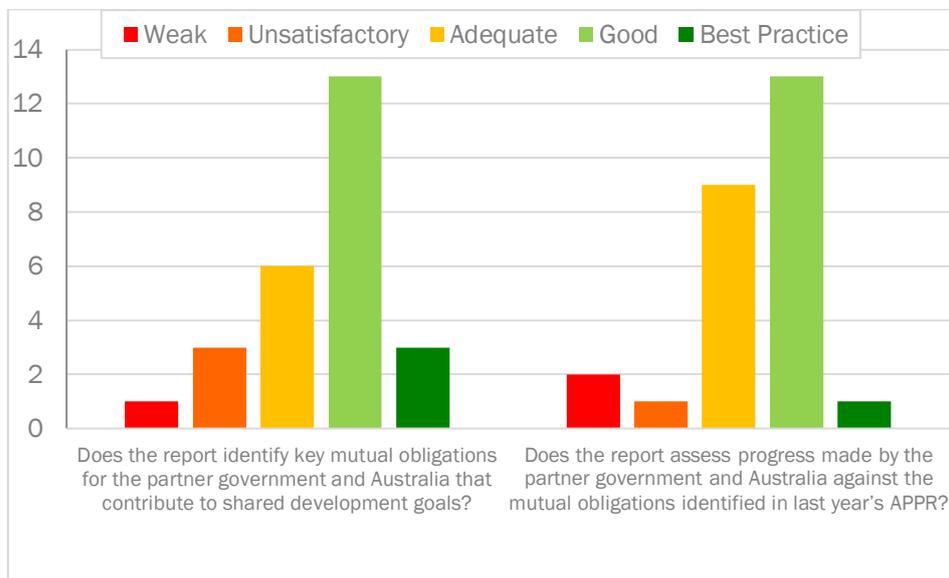
<p><b>Priority Outcome 4: Infrastructure</b></p> <p>Improve infrastructure services through increasing access to telecommunications and sanitation services and improving the main road.</p> <p><i>Note: Very heavy rainfall associated with an El Niño year (2015) has contributed to the delay in some infrastructure projects.</i></p>	<p>Improve telecommunications service on South Tarawa</p>	<p>Partly achieved</p>	<p>Achieved: Australia has supported the successful privatisation of Telecommunications Services Kiribati Limited and finalisation of the Universal Access Plan and ICT Policy, which it anticipates will improve telecommunications access and service across the county. Australia's financial involvement in the project is now completed, however, the World Bank anticipates that further funding will be needed to expand access to the outer islands.</p>
	<p>Improve the main road network on South Tarawa and strengthen road finance and maintenance capacity</p>		<p>Partly achieved: The construction work completed to date has reduced travel time and provided people with a safer road network. However, the project remains behind schedule.</p>
	<p>Enhance access to sanitation services by rehabilitating the existing sewage network</p>		<p>Partly achieved: Work to rehabilitate the dilapidated sewage network on South Tarawa commenced in 2014–15. However, the project is behind schedule.</p>

### 3.2 Mutual obligations

One year on from the introduction of mutual obligations and 2014–15 APPRs are more effective in identifying clear mutual obligations than they were in assessing progress towards them (as seen in Graph 4). All programs included mutual obligations, which is an improvement on the

2013–14 APPRs where three APPRs failed to include them.\* Better practice included programs that had specified mutual obligations, often sector-specific, beyond simply sharing development objectives with the partner country. There were some examples of good practice in articulating clear, evidenced mutual obligations (Bangladesh, Vanuatu and the Philippines).

Graph 4 Frequency of ratings for assessing mutual obligations



Three programs provided a weak description of mutual obligations (Nauru, Pakistan, Sri Lanka); one program (South East Asia Regional Program [SEARP]) identified the program’s mutual obligations with the partner governments but did not provide an assessment of progress. Weaker practice in mutual obligations was characterised by reporting alignment with partner governments but failing to articulate mutual obligations between the two governments. The Philippines presented an example of best practice in reporting on mutual obligations (see Box 9).

\* Mutual obligations were not applicable for ANCP and were therefore not included.

## Box 9 Good Practice: the Philippines assessment of progress towards Mutual Obligations

The Philippine Government articulates its development agenda in the Philippine Development Plan (PDP). The PDP 2011–16 is focused on addressing the fundamental impediments to inclusive growth—low economic growth, high income inequality, and governance challenges, by seeking to generate inclusive growth, uplift the poor and protect the vulnerable. The PDP was the anchor for the bilateral development agreement between Philippines and Australia, through the Statement of Commitment (SOC) signed in March 2012. The Philippine Government is making significant progress in its eight commitments under the SOC:

Philippine Government commitments	Progress
Increase public expenditures, through the General Appropriations Act, in basic education and support to local government units through performance-based mechanisms.	The Philippine Government has increased the education budget year on year. 2011: PHP207.5 billion, 2012: PHP238.4 billion, 2013: PHP292.2 billion, 2013–14: PHP307.9 billion, and 2014–15: PHP364.7 billion.
Accelerate implementation of the basic education reform agenda.	The K–12 program was approved in 2012 through the Enhanced Basic Education Act. The new K–12 curriculum has been completed and is being implemented in nine grade levels (K-4 and 7–10).
Continue investing in human resource management and development for all levels of government bureaucracy to improve service delivery and public sector accountability.	The Civil Service Commission (CSC) is implementing the ISO 9001:2008 Certified Program to Institutionalise Meritocracy and Excellence in Human Resource Management. The CSC has also been certified by Investors In People, an international quality standard accrediting agency for excellence in people management.

### 3.3 Performance assessment frameworks

Performance assessment frameworks (PAFs) articulate the logic for how aid interventions will contribute to program objectives. They align program objectives with outcomes, indicators and baselines and thus enable tracking of progress towards objectives. As shown in Table 3 below, in 2014–15, those programs with PAFs tended to be at the upper end of the ratings table with few exceptions, which is consistent with the findings from the 2013–14 review. Reports that drew upon a PAF in the APPR were of good or better quality. Although having a PAF does not appear to be a determinant on the overall quality of the APPR (six APPRs were overall high quality and did not draw on a PAF), all programs with a PAF clearly framed objectives at the outset. Using the PAF also helped in setting good quality and well-integrated performance benchmarks, discussed above.

Table 3 Ranking of APPR quality by total cornerstone question scores for each APPR.

	Country/region	Cornerstone ratings only	2a	2b
High quality	<b>*Myanmar</b>	13	4	5
	<b>*Afghanistan</b>	12	4	4
	<b>*Bangladesh</b>	12	4	5
	<b>*Pakistan</b>	12	4	4
	Philippines	12	4	5
	Solomon Islands	12	4	4
	<b>*South Asia</b>	12	4	3
	SEARP	12	4	4
	Sri Lanka	12	4	5
	Timor-Leste	12	4	4
	Vanuatu	12	4	3
	<b>*Vietnam</b>	12	4	4
	Good quality	<b>*Cambodia</b>	11	4
Indonesia		11	4	3
Pacific Regional		11	4	5
Samoa		11	4	4
<b>*ANCP</b>		10	4	3
<b>*Laos</b>		10	4	3
Kiribati		10	3	2
Fiji		10	3	4
Nepal		10	3	3
Palestinian Territories		10	4	4
Sub-Saharan Africa		10	3	4
PNG		9	3	2
Adequate quality		Tonga	9	3
	Nauru	8	2	2

**Key:**

\* = Programs with a PAF

2a = Are the objectives sufficiently framed at the outset or in the narrative as statements of intent that are measurable (quantitatively and/or qualitatively)?

2b = Does the narrative explain and justify the progress ratings against each objective?

### 3.4 Compliance with guidance

All reports complied, in large part, with guidance as set out in the APPR template and DFAT's good practice note *Preparing Aid Program Performance Reports 2014*. Compliance with guidelines has improved since 2013–14, particularly in the inclusion of the appropriate section headings for mutual obligations and/or partner assessment.

Table 4 APPR requirements

<b>Section headings in place</b>	All reports included the required section headings, with the exception of two programs (Kiribati, Nauru) which did not include a risks section
<b>Annexes in place</b>	All reports included Annexes A-D with the exception of two programs (Kiribati, Nauru)

However, more than half of reports did not comply with guidance and exceeded the maximum page limit for APPRs as determined by the size of programs.\* Reports' page length does appear to have some bearing on quality, notably those reports that went under the page limit (Fiji, Kiribati, Nauru, South Asia) were in the lower quality tier and those reports that exceeded their maximum page length were, with a single exception, in the upper quality tier. A further observation is that those reports that went over the maximum page length (by three pages or more) benefited from a stronger narrative and more extensive analysis of progress, with the exception of PNG. Based upon this, length of the report appears to be an asset.

---

\* Specifications for page limits are set out in Table 1 of the DFAT good practice note, which specifies a maximum page limit for APPRs (excluding annexes and title page) of 20 pages (more than \$300 million total ODA in 2013-14), 15 pages (\$50 million to \$300 million) or 10 pages (less than \$50 million).

## Appendix 1: Review Terms of Reference (excerpt)

### Background

1. Established in 2006, the Office of Development Effectiveness (ODE) in the Department of Foreign Affairs and Trade (DFAT) builds stronger evidence for more effective aid. A core function of ODE is to quality assure the robustness of the aid program's internal performance systems and reporting.
2. The aid program's Aid Program Performance Reports (APPRs) are self-assessments by program areas of their progress toward objectives set under country or regional strategies. APPRs were established in 2007 to serve three purposes:
  - a) strengthen program management and improve effectiveness;
  - b) ensure accountability and transparency at the country/regional program level; and
  - c) strengthen relationships with partners, especially partner governments and multilateral agencies.
3. Bilateral and regional programs are critical elements of the Australian aid program, accounting for 64 per cent of the DFAT aid budget. The department requires high-level oversight of country-level spending, results and alignment of aid, trade and foreign policy objectives.
4. In June 2014 the Minister for Foreign Affairs released *Making Performance Count*, the new performance framework for the Australian aid program. In addition to setting out 10 agency-level targets, it also requires country/regional programs to establish 'performance benchmarks' against which their yearly performance will be assessed and budgeting decisions made. The program-level benchmarks were set out in the 2013-14 APPRs and will be reported against in 2014-15 APPRs.
5. In 2015, 23 APPRs will be produced. Under simplified reporting arrangements, APPR size and the numbers of program-level performance benchmarks will vary according to program size (see table).
6. Since 2008, ODE has reviewed annually the quality of APPRs. ODE's reviews are overseen and endorsed by the Independent Evaluation Committee. A response to review recommendations is sought from DFAT Executive. The review is published.

### Objectives

7. In quality assuring and analysing the 2014-15 APPRs, ODE's 2015 Review of APPRs will:
  1. Assess and report on the reasonableness of the performance narratives and the use of evidence in the 2014-15 APPRs; and
  2. Assess and report on the reasonableness of the performance narratives against country and regional program-level performance benchmarks proposed in the 2013-14 APPRs.

## Audience

8. The key audiences of ODE's Review of APPRs are the Office of the Minister for Foreign Affairs, DFAT Executive and senior managers, staff with responsibility for delivering country and regional programs, APPR authors, and aid program performance and quality managers.
9. The review should be written with a view to it being published.

## Approach

10. Evidence for the 2015 Review of APPRs will be gathered from the following sources:

1. The 2014-15 APPRs
2. The 2015 Aid Investment Plans
3. The most recent APPR guideline and template
4. Previous ODE APPR Quality Reviews and working papers
5. Interviews with a sample of APPR authors, if required.

11. The review consists of two main tasks:

1. Assess the 2014-15 APPRs, including the performance benchmarks, against the following quality criteria, consulting with ODE on the assessment method:
  - Does the APPR provide a clear and balanced description of overall program progress within the context of the new aid policy framework?
  - Does the APPR narrative explain and justify, in a proportionate way, the progress ratings against each program objective?
  - Does the APPR provide a proportionate, balanced discussion of the contribution of partner governments and delivery partners?
  - How well have new requirements been integrated into the APPR reporting process, namely performance benchmarks and mutual obligations?
  - Are the management responses clearly linked to the ratings narratives and the identified risks to future progress?
2. Write the 2015 Review of APPRs.

## Appendix 2: Method

The **objective** of the 2015 Quality Review of APPRs is to assess and report on the reasonableness of the **performance narratives** and the use of evidence in the 2014–15 APPRs; and the reasonableness of the performance narratives against country and regional program-level **performance benchmarks** proposed in the 2013–14 APPRs.

To inform the review, **evidence** has been gathered from all 26 of the 2014–15 APPRs (see Table 1), 25 2015 Aid Investment Plans, the most recent APPR guideline and template, previous ODE APPR quality reviews and working papers and ongoing dialogue with ODE.

Catherine Currie (CC) and Judith Friedman (JF) were contracted as Team Leader and Team Member respectively for IOD PARC to undertake the quality review of 2014–15 Aid Program Performance Reports (APPRs) with oversight and technical advice from Jo Hall (JH).

Table 1 2014–15 APPRs assessed in this quality review

APPR type	Program
Country APPRs	Afghanistan, Bangladesh, Cambodia, Fiji, Indonesia, Kiribati, Laos, Myanmar, Nauru, Nepal, Pakistan, Palestinian Territories, Papua New Guinea, Philippines, Samoa, Solomon Islands, Sri Lanka, Timor-Leste, Tonga, Vanuatu, Vietnam
Regional APPRs	East Asia, Pacific, South Asia, Sub-Saharan Africa
Other APPRs	Australian NGO Cooperation Program (ANCP)

The quality criteria were put into a template including cornerstone questions, which, as in previous years, provide summary ratings for each section, and subquestions.

1. **Context and strategic orientation:** Does the APPR provide a clear and balanced description of overall program progress within the country/regional context and aid policy framework?
  - a. Does the APPR respond to and reflect Australia’s aid policy?
  - b. Does the report identify key mutual obligations for the partner government and Australia that contribute to shared development goals?
  - c. Does the report contain performance benchmarks that provide a reasonable basis for assessing program performance?
  - d. Does the discussion of the program progress reflect on how well the program strategy is working to achieve the program objectives?
  - e. Does the report make clear Australia’s contribution and the proportion and significance of Australia’s aid compared to other donors and national budgets, and the program and strategy implications of that?
2. **Progress assessment:** Does the APPR narrative explain and justify progress ratings against each program objective?
  - a. Are the objectives sufficiently framed at the outset or in the narrative as statements of intent that are measurable (quantitatively and/or qualitatively)?
  - b. Does the narrative explain and justify the progress ratings against each objective?
  - c. Does the report assess progress made by the partner government and Australia against the mutual obligations identified in last year’s APPR?
  - d. Does the report provide evidence about progress against performance benchmarks, and where benchmarks have not been achieved, discuss why this is the case and what management actions will be taken?

- e. Where appropriate, is the policy dialogue engagement with partners clearly explained?
  - f. Is there an appropriately open, balanced discussion of the program's successes, failures and challenges?
  - g. Has relevant evidence been drawn from an appropriate range of internal and external sources given the program context?
  - h. Does the narrative justify the progress and results in promoting gender equality and empowerment of women and girls? Are trends and planned follow-up actions identified?
3. **Management responses:** Are the management responses clearly linked to the ratings narratives and the identified risks to future progress?
- a. Is there a logical connection between the management responses section and the risks to future progress flagged in the other sections?
  - b. Are the specific management responses framed so they are relevant to program strategy and management and are also actionable within the reporting period? Are they also prioritised?

## Analysis

Ratings were compiled and analysed. The analysis focused, at the portfolio level, on identifying common themes and patterns in the assessment narratives, as well as identifying good practice and/or common weaknesses by analysing the ratings. The portfolio of APPRs was then analysed by:

- › objectives: comparing the objectives of 2013–14 to those of 2014–15
- › region: the particular characteristics of regional programs in 2014–15
- › country: small and large country programs in 2014–15
- › 2014–15 priorities: mutual obligations, performance benchmarks and gender.

## Changes to Assessment Tool Section A 2014–15

In 2013–14 there was one cornerstone question and four subquestions under section A. In 2014–15, two subquestions were deleted and three were added.

Deletions from the 2013–14 assessment tool:

- › 1b, *'Does the context analysis explain the key development challenges/needs of the country/region, with an emphasis on changes that have occurred during the reporting period? Is this analysis used to explain the Australian program objectives?'* was deleted in light of the introduction of the Aid Investment Plans for country and regional programs which provide in-depth contextual analysis of the country or region. Therefore, the requirement that the APPRs provide an in-depth context analysis as a backdrop for the program became less of a focus for the review.
- › Question 1c, *'If there is a stated strategy or "theory of change" does this explain how program objectives will be, and does this explanation reference the country/regional context analysis? In the absence of a stated strategy or theory of change, is some other explanation for achieving program objectives, given the context, provided?'* was deleted, but a new question (1d in the current assessment tool) focuses on how well the program strategy is working to achieve the objectives.

The three new additions for 2014–15:

- › Does the report identify key mutual obligations for the partner government and Australia that contribute to shared development goals? (1b)
- › Does the report contain performance benchmarks that provide a reasonable basis for assessing program performance? (1c)

- › Does the discussion of the program progress reflect on how well the program strategy is working to achieve the program objectives? (1d)

Therefore, there remain four questions that are directly comparable between 2013–14 and 2014–15.

2013–14	1	1a	1c	1d
2014–15	1	1a	1d	1e

### Changes to Assessment Tool Section B in 2014–15

In 2013–14 there was one cornerstone question and seven subquestions under section B. In 2014–15, two subquestions were deleted and three were added. The three new additions are:

- › Does the report assess progress made by the partner government and Australia against the mutual obligations identified in last year’s APPR? (2c)
- › Does the report provide evidence about progress against performance benchmarks, and where benchmarks have not been achieved, discuss why this is the case and what management actions will be taken? (2d)
- › Does the narrative justify the progress and results in promoting gender equality and empowerment of women and girls? Are trends and planned follow-up actions identified? (2h)

Therefore, there remain six questions that are directly comparable between 2013–14 and 2014–15.

2013–14	2	2a	2b	2c	2d	2e
2014–15	2	2a	2b	2e	2f	2g

As a result of feedback from 2013–14 process, the following changes were made to section B questions:

- › The question, ‘*Are the sources of evidence referenced properly?*’ (2g) was deleted because it was deemed not to carry the same weight as the other subquestions. However, the 2014–15 APPR assessment tool retained an understanding of the importance of referencing.
- › Question 2f, ‘*Is the results evidence appropriately contextualised and contribution of DFAT clear?*’ was also deleted in 2014–15 but focus on contribution was retained through subquestion 1d and also in terms of how the APPR articulated the ‘problem’ the program was addressing.

### Changes to Assessment Tool Section C 2014–15

Section C remains mostly consistent with previous years, with one small change:

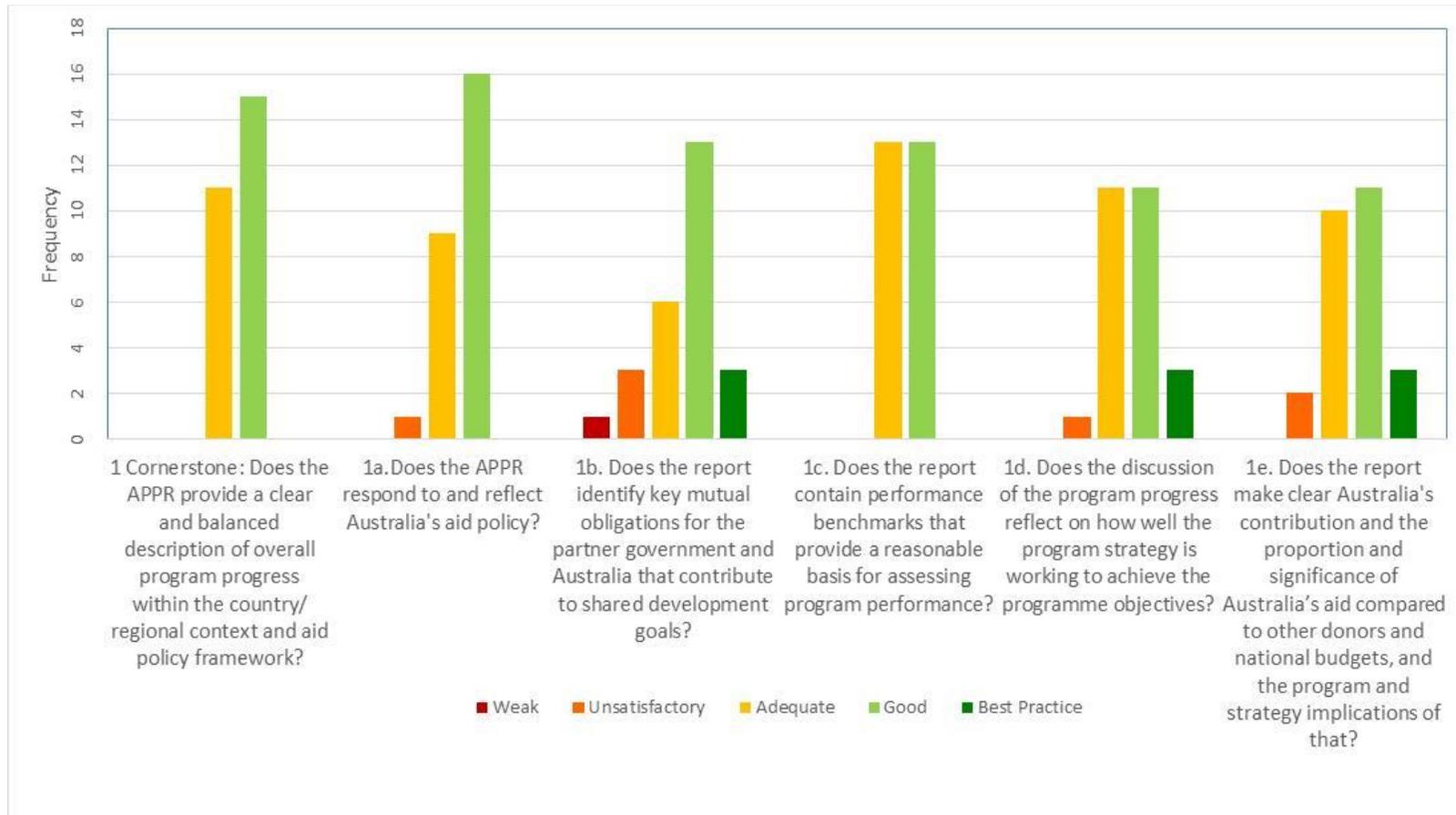
- The question, ‘*Are lessons from the experience of implementing the previous year’s management responses noted and acted upon in this year’s consequences section?*’ (1c) was deleted because it was found to be less critical than the other management issues.

In 2013–14 there were three subquestions and one cornerstone under section C. In 2014–15, one subquestion was deleted. Therefore, there remain three questions that are directly comparable between 2013–14 and 2014–15.

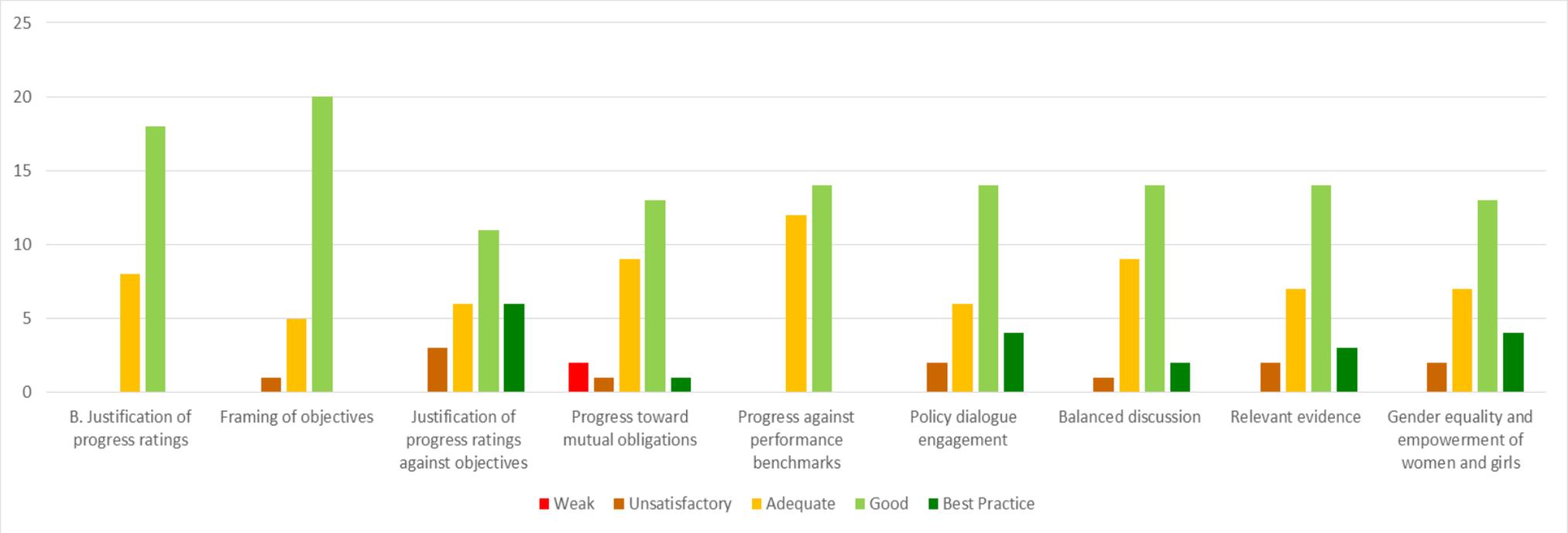
2013–14	4	4a	4b
2014–15	3	3a	3b

## Appendix 3: All ratings assessment criteria

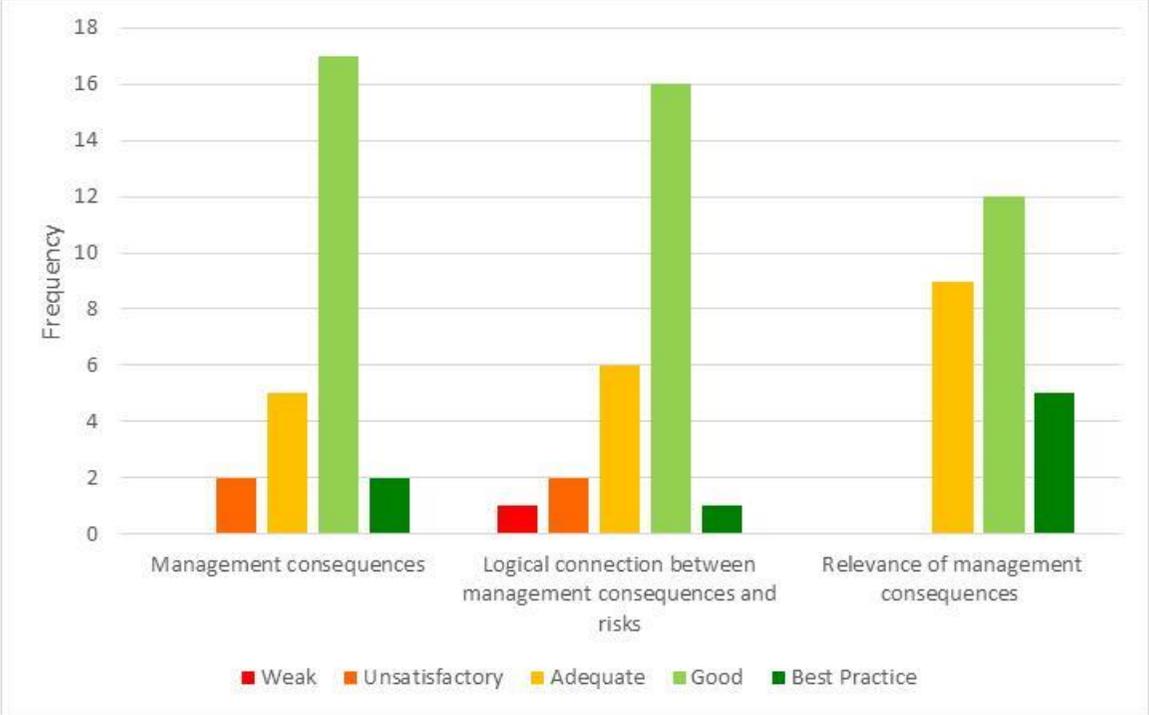
### Context and strategic orientation



# Progress assessment



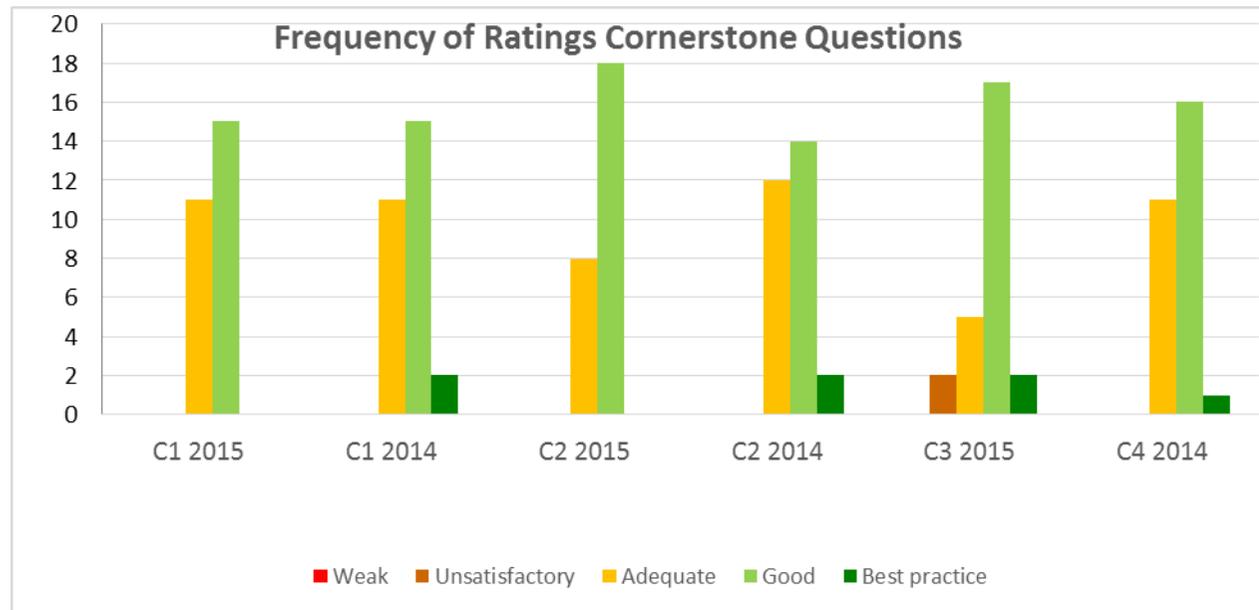
# Management responses



## Appendix 4: Comparison between 2013–14 and 2014–15

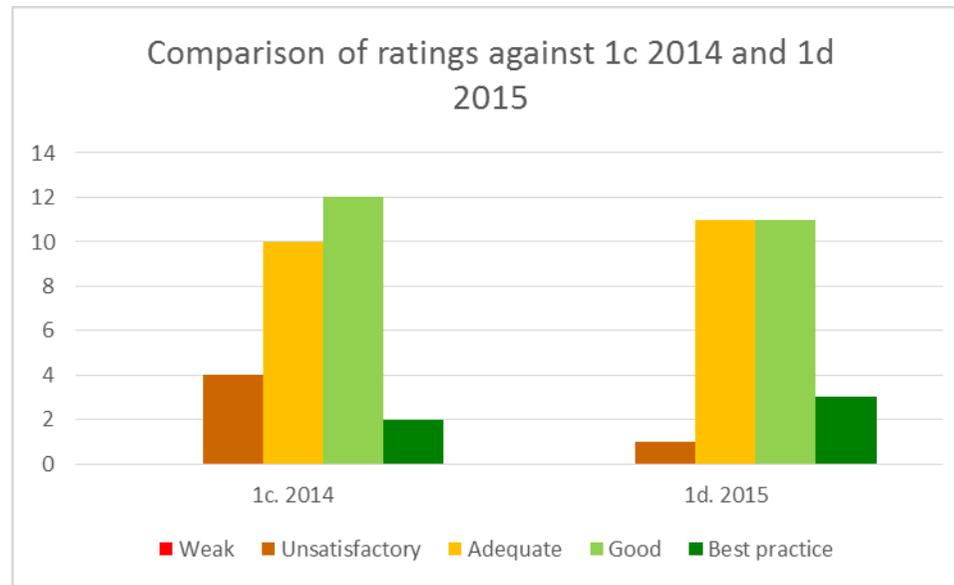
### Cornerstone questions comparison: frequencies 2013–14 and 2014–15

1. Does the APPR provide a clear and balanced description of overall program progress within the country/regional context and aid policy framework?
2. Does the APPR narrative explain and justify progress ratings against each program objective?
3. Are the management responses clearly linked to ratings, narratives and the identified risks to future progress?



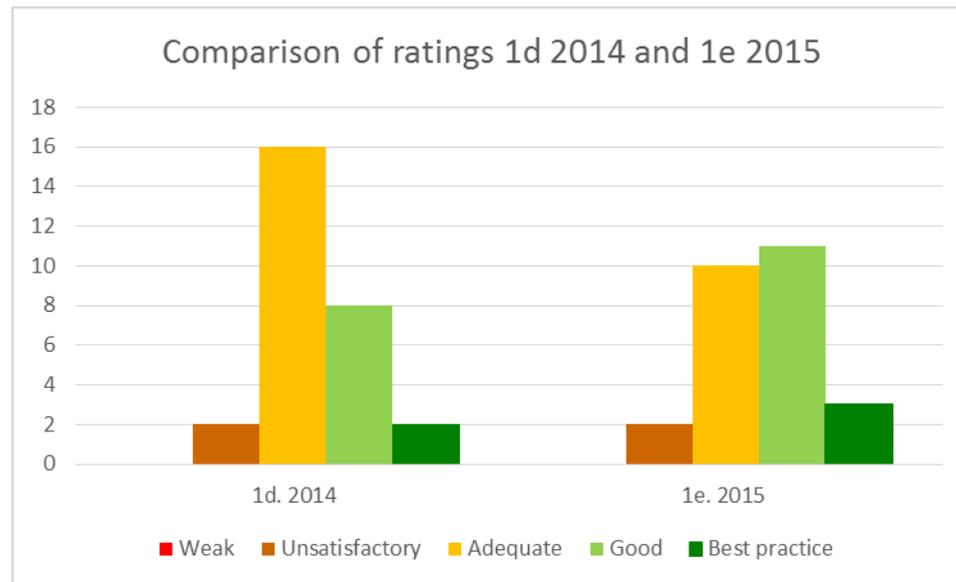
Comparison of core question ratings on program progress: 1d (2015) with 1c (2014)

1. Does the discussion of the program progress reflect on how well the program strategy is working to achieve the program objectives?



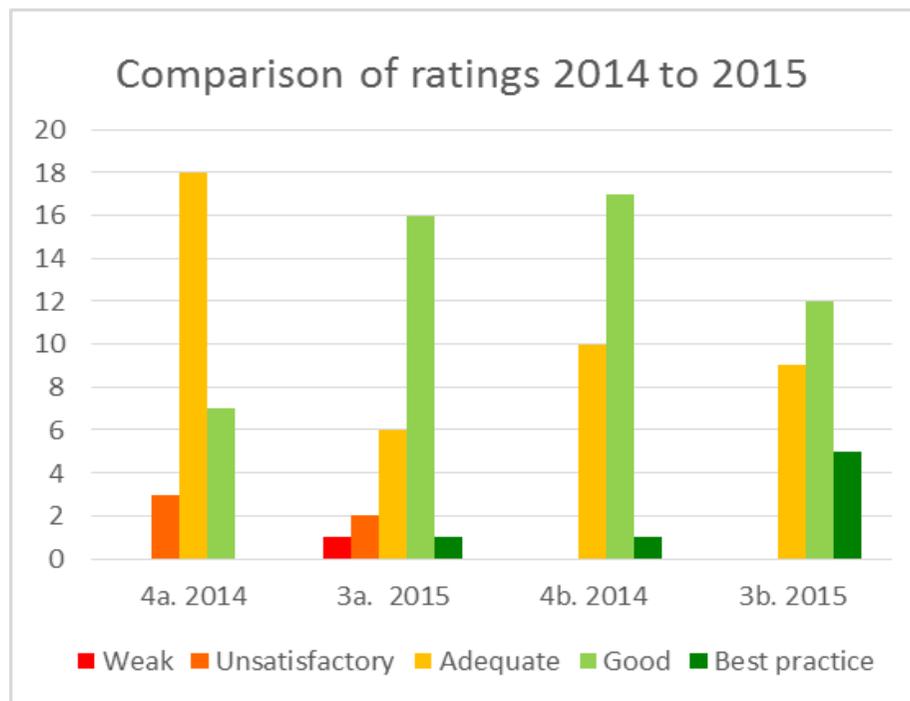
Comparison of core question ratings on program progress: 1e (2015) with 1d (2014)

1. Does the report make clear Australia's contribution and the proportion and significance of Australia's aid compared to other donors and national budgets, and the program and strategy implications of that?



### Comparison of ratings on management responses: 3a and 3b (2015) with 4a and 4b (2014)

- a. Is there a logical connection between the management consequences section and the risks to future progress flagged in the other sections?
- b. Are the specific management consequences framed so they are relevant to program strategy and management and are also actionable within the reporting period?  
Are they also prioritised?



---

<sup>1</sup> <http://dfat.gov.au/aid/how-we-measure-performance/ode/odepublications/Pages/policy-dialogue.aspx>