How Relevant is Relevance? A Comparative Analysis of How the DAC Criterion Is Understood and Applied Across Evaluation Agencies

Xiaomen Wong, Diana Dadzie, Maudline Okeyere, Pan Bei, and Howard White

REC Working Paper No. 1
June 2026

The Research and Evaluation Centre

Promoting evidence-based policy and practice

REC undertakes high-quality, policy-oriented research and evaluation in areas of social and economic policy. We work with a network of Research Associates and Policy Associates across the world, with a focus on Africa and Asia. REC is committed to supporting the career development of early career professionals, especially those from disadvantaged backgrounds.

We support the production and use of evidence to inform policy and practice. We contribute to building the evidence architecture through: evaluation, evidence synthesis, knowledge translation, evidence-based advice for policy and practice, training and mentoring, and strategic advice and institutional support.

Suggested Citation

Wong, X., Dadzie, D., Okeyere, M., Bei, P., and White, H. (2026) How Relevant is Relevance? A Comparative Analysis of How the DAC Criterion Is Understood and Applied Across Evaluation Agencies. REC Working Paper No. 1. London: The Research and Evaluation Centre.

How Relevant is Relevance? A Comparative Analysis of How the DAC Criterion Is Understood and Applied Across Evaluation Agencies

Xiaomen Wong,[a] Diana Dadzie,[b] Maudline Okeyere,[b] Pan Bei,[a] and Howard White[a,b]

a Centre for Evidence-Based Social Sciences at Lanzhou University

b The Research and Evaluation Centre

Abstract

The criterion of relevance is one of the cornerstones of the OECD Development Assistance Committee (DAC) evaluation framework, guiding international development agencies in determining whether interventions are consistent with beneficiaries' needs, country priorities, and global frameworks. However, despite its conceptual importance, "relevance" has often been seen as a low bar—equated with superficial policy alignment rather than a substantive assessment of design responsiveness or adaptability. This article revisits the DAC relevance criterion through a comparative analysis of four major agencies—the International Labour Organization (ILO), the United Nations Development Programme (UNDP), the United Nations Children's Fund (UNICEF), and the World Bank. We consider three dimensions of evidence claims: relevance to whom, specificity of the evidence claim, and the evidence basis for the claim. We find that the application of the relevance criterion has improved between 2013 and 2023. Nonetheless, some deficiencies remain, notably with unsubstantiated evidence claims. Our main conclusion is a methodological one. The evaluation community should adopt the practice of assessing the confidence we can have in the evidence claims made in evaluations. As an example of this approach, we propose a checklist to be used to guide evaluators and to assess relevance claims.

1. Introduction

The evaluation criteria established by the OECD Development Assistance Committee (DAC) have defined the standard criteria for development evaluation for over three decades. When the DAC first introduced five core criteria in 1991—relevance, effectiveness, efficiency, impact, and sustainability—it provided a shared vocabulary for assessing development interventions (OECD DAC, 1991, 2002). These criteria have been adopted across the UN system, by the World Bank and many other agencies. Within this framework, relevance can be seen as a gateway criterion. If an intervention is not seen as relevant – that is, it is not doing the right thing—its success against the other criteria does not seem to matter as much (OECD DAC, 2019).

Despite its clear conceptual definition, the application of relevance in practice has often proved challenging, most notably because it sets a low bar. Consultations undertaken for the revision of the DAC criteria in 2018 emphasized key perceived weaknesses. Feedback from practitioners indicated that relevance assessments were often restricted to citing policy documents without examining whether the intervention addresses a policy issue in the way that is appropriate to the context and preferred by stakeholders or whether the intervention remains relevant as conditions change (OECD DAC, 2019).

Poor reporting of the evidence base used to support relevance claims further complicates efforts to distinguish between well-supported evaluation findings and those based on simple assumptions. Hence, whilst the relevance criterion is meant to anchor the evaluation in a solid understanding of contextual needs and priorities, it risks becoming a simple procedural formality, offering little insight into the actual justification of an intervention.

This issue has become more pronounced as the global development agenda has expanded. The transition from the Millennium Development Goals (MDGs) to the Sustainable Development Goals (SDGs) dramatically multiplied development priorities—from 8 goals and 21 targets to 17 goals and 169 targets (Nilsson et al., 2016).

A review of World Bank's Country Partnership Frameworks (CPFs) found they had a large number of, often very broad, objectives, which allowed a wide array of interventions to be judged as relevant (World Bank IEG, 2022). Having such a broad range of development objectives – with many countries simply adopting the MDGs rather than customizing to their context – undermines the relevance criterion's analytical value by making it difficult to distinguish issues of genuine strategic importance and strong local demand from those which are more peripheral (World Bank IEG, 2017).

Motivated by these concerns, this study presents a new framework for assessing relevance. We apply this framework to investigate how four development agencies have applied the relevance criterion in their evaluations. We examine whether practices vary significantly across institutions or have changed over time.

The analysis is guided by three main questions:

1. What criteria can be used to assess the strength of "relevance claims"? An evidence claim is a statement based on evidence, with a relevance claim being a specific type of evidence claim.
2. Using these criteria, how do major international agencies assess relevance in practice?
3. To what extent has the assessment of relevance evolved over time, particularly in relation to the 2019 DAC revision?

These questions provide a framework for exploring both the strengths and the limitations inherent in current evaluative practice in assessing relevance.

This study contributes to the evaluation literature in several ways. By conducting a cross-agency comparison of the International Labour Organization (ILO), the United Nations Development Programme (UNDP), the United Nations Children's Fund (UNICEF), and the World Bank, it offers a comprehensive view of how relevance is operationalized across distinct institutions. The analysis highlights recurring conceptual and methodological gaps, including inconsistent treatment of the stakeholder groups to whom relevance is attributed.

Following this introduction, Section 2 details our assessment framework, outlining the core concepts of the three-dimensional coding approach. Section 3 reports our findings with a discussion in the concluding Section 4. The annex includes a short checklist to be used by evaluators to make or assess relevance claims.

2. Assessment Framework

2.1 Background

The DAC evaluation criteria were first introduced in 1991 and have since become the global standard for assessing development interventions (OECD DAC, 1991; 2019). Initially, the relevance criterion was defined as the "extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities, and partners' and donors' policies." This definition implicitly assumed policy coherence as a proxy for relevance; that is, alignment was equated with appropriateness.

By the time of the 2019 revision, the DAC Working Party on Development Evaluation recognized the need to broaden this understanding. The revised definition emphasized not only alignment but also adaptiveness and responsiveness to changing circumstances. It explicitly called for evaluators to consider how interventions remain relevant in a changing context, particularly in fragile states and climate-sensitive regions. This shift marked a change from static to adaptive relevance (OECD, 2019).

Consultations preceding the 2019 revision revealed that "relevance" was among the most frequently addressed criteria, but it was also among the least deeply assessed. Evaluations often equated relevance with formal alignment to national strategies or UN frameworks, neglecting the deeper question of whether the design, logic, and timing of interventions actually addressed beneficiaries' needs or adapted to emerging conditions. In other words, the criterion remained conceptually rich but methodologically thin in practice.

The growing recognition of the superficiality in how relevance was operationalized across major agencies was one factor in the revision of the DAC criteria. The 2018 DAC consultation confirmed that the relevance criterion was often treated as a check-box exercise, leading to uncritical conclusions such as “the project is relevant to national priorities.” Such statements lack analytical depth and provide little evidence of real alignment or responsiveness.

In light of these concerns, the new definition is “The extent to which the intervention objectives and design respond to beneficiaries’, global, country, and partner/institution needs, policies, and priorities, and continue to do so if circumstances change” (OECD DAC, 2019: 6). The main changes are adding the word “design” and “continue to do so if circumstances change”. A final change is that the word donor has been dropped, with reference now to “partners/institutions” which “includes government (national, regional, local), civil society organisations, private entities and international bodies involved in funding, implementing and/or overseeing the intervention”.

The motivation for undertaking this work arose from the direct experiences of one of the authors (White) which are listed here as they illustrate the importance of some of the points being made.

The first example is anecdotal from one of the author’s experiences on the government side for negotiations between an African government and the EU delegation. The delegation, presented with a long list of requests covering every ministry, stated they wanted to fund priority areas. The Permanent Secretary for Planning replied “everything is a priority”. As mentioned above, the SDGs provide an official stamp of approval for the “everything is a priority” view, making high-level assessments of relevance meaningless.[1]

---

[1] The same point regarding the SDGs is made in Lowcock and Dissanayake’s (2024) history of DFID.

The second example is the support of the European Union for youth employment in Scotland under the EU's Youth Employment Initiative (Scottish Government, 2016). The Scottish government first proposed support to the apprenticeship programme, but this was rejected by Brussels as it was nationwide whereas the funds could only be used in the most disadvantaged areas identified by the European Union (Scottish Government, 2016). The government then proposed supporting programmes in schools to keep children in school at age 16, where they would be exposed to a range of programmes to prepare them for pathways into employment. This was also rejected as the European funds had to be used for youth who were already NEET (not in employment, education or training), not preventive interventions for those at risk of becoming so. Only a third attempt was accepted, though funds were under-utilized as the EU rejected the Scottish government's proposed costing model in favour of one which was unduly onerous for many prospective implementing partners. Despite all these issues, the intervention would still be judged as relevant using a high-level policy assessment (youth employment is a priority, though rather less so by the time the project was finally approved) despite clearly not fitting with the government's preferred programmatic approach. Hence the need to consider what we call the specificity of the relevance claim – is it at a general policy level or does it refer to the specific intervention design?

Third, White was an advisor for the World Bank's Results and Performance (RAP) report for 2022 which assessed performance at the country level. The report identified a recurring failure of country policy frameworks to prioritize among competing objectives, leading to what the authors termed "a lack of selectivity" (World Bank IEG, 2022: xix). The report also identified the issue of weak adaptiveness, that is "insufficient preparedness to respond to changes in country conditions, government commitment, or Bank Group priorities" (ibid).

These anecdotes underline the central question: How relevant is relevance itself? If relevance continues to be assessed superficially, its potential to guide decision-making and be a basis for assessing accountability will remain unrealized.

2.2 The Three-Dimensional Coding Framework

The DAC guidance "Applying Evaluation Criteria Thoughtfully" lays out several issues to consider when assessing relevance. These are organized around four elements to be considered: relevance to beneficiaries and stakeholders, relevance to context, relevance of quality and design of the intervention, and whether relevance has been maintained over time. The first of these is considered the most important: "the most important element for analysing relevance is the assessment of the extent to which an intervention addresses beneficiaries' needs and priorities" (OECD DAC, 2021: 39). Our framework incorporates this first element mentioned in the guidance, but our main focus is on what we believe are neglected aspects of specificity and the evidence basis for the relevance claim. Hence the framework employs three dimensions to analyse the assessment of relevance: (i) relevant to whom?, (ii) specificity of the evidence claim, and (iii) basis for the relevance claim.

Dimension I: Relevant to Whom? (Institutional Accountability)

This dimension identifies the stakeholder referenced in the relevance claim. A single evaluation report may make relevance claims with respect to more than one stakeholder. This dimension of relevance assesses how accountability is being judged in the evaluation. These stakeholders are:

- Country or government: Alignment with national policies, which indicates top-down, political accountability.
- Beneficiaries: Alignment with the direct needs of the target group (those intended to be engaged with the intervention), which indicates bottom-up, social accountability.

- Partners/institutions: Alignment with the executing partners' institutional mandates or with broader civil society, which is horizontal accountability between development partners.
- Funding Agency: Alignment with specific donor priorities, which is internal accountability for the funder itself. In the revised DAC definition of relevance, this is included under partners/institutions, but we find it useful to separate it out.
- Global: Alignment with high-level frameworks (e.g., SDGs), which we call general accountability as it does not correspond to relevance to any specific group, but is explicitly allowed for in the OECD guidance.[2]

Dimension II: Specificity of the Claim

This dimension assesses how specific the statement making the relevance claim is. We identify three categories:

- General assertion of alignment: Unsupported statements that the intervention is aligned with national or beneficiary priorities with no specific policy or document named to support the claim.
- Cite policy or strategy aligned with: explicit reference to a named policy or policy document, or some other means of identifying need.
- Give specific details of alignment: Provide details that connect the intervention design or components to specific parts of a policy or need. The 2018 revision to the DAC criteria added the word "design" to the criterion "to deepen the analysis of relevance" (OECD DAC, 2019: 6).

Dimension III: Basis for Relevance Claim (Evidential Rigour)

This dimension examines the type of evidence used to support the claim, which is critical for assessing the strength of the relevance claim:

- Not Stated: No source or evidence documented, which is a failure in evidence transparency.

[2] "Relevance can also be considered in relation to global goals such as the Sustainable Development Goals (SDGs)" (OECD DAC, 2021: 38).

- Documents: Evidence drawn from project or policy documents.
- Interviews / Non-Comparative: Stakeholder confirmation that the project was useful.
- Interviews / Comparative: Explicit evidence from stakeholders prioritizing the project above alternatives.

The last two categories distinguish between comparative and non-comparative evidence claims. The latter is simply a statement that the intervention is relevant for that stakeholder. A comparative claim is a statement that it is more important than alternative interventions or intervention designs tackling the same issue. The DAC guidance implies that relevance should be defined in a relative sense (the best thing to do, not just a good thing to do), and recognizes that perspectives on this may vary between stakeholders (OECD DAC, 2021: 39-40).

An example which illustrates comparative versus non-comparative assessment is from an evaluation of Social Funds, which are community-driven projects (World Bank, 2002). Grants are managed by the community to implement a small-scale project identified by the community such as school or health clinic construction or rehabilitation, local water supply or feeder roads. The evaluation found that project selection was normally made by a small group rather than the community as a whole. A survey of households in beneficiary communities showed that, whilst they appreciated the project (non-comparative relevance), only a minority saw the selected project as being a priority. Indeed, it was not in the top three priorities of the majority of respondents (lack of comparative relevance).

We give some examples of the application of the coding framework (Table 1). In the first example, from an ILO evaluation, there are two separate relevance claims in the quoted text. The first sentence is a general assertion of relevance to beneficiaries but no source is cited. The second sentence names a government policy document but, despite giving a quote from the document, does not provide specific details of intervention design linked to specific activities in the National Development Plan.

The second example, from UNDP, is a general assertion of both global and national relevance, but with no source (basis) cited. The same is true of the third example from the World Bank. The UNICEF example names the national programme for which the intervention is relevant, though fails to provide a reference to any basis for the claim, such as a policy document or programme strategy.

Table 1 Examples of applications of the three dimensions of assessing relevance

Example	Who?	Specificity	Basis
ILO: The design of the programme is well motivated based on the prevailing socio-economic situation in the country, the opportunities for meaningful benefits from cultural tourism and the need to strengthen the cultural stakeholders in their pursuit of sectorial effectiveness. The JP also rightfully anchored in Namibia's national planning efforts (NDP 3 and NDP 4) that recognise "the need for utilisation of natural and cultural resources through the priority sector of sustainable tourism".	Beneficiaries Government	General assertion Cite policy aligned with	Not stated Documents
UNDP: Project's broader goal of improving WRC conservation generally in China is highly relevant to world and China, in light of both continuing biodiversity losses and potential value of WRCs to food security.	Global Country	General assertion	Not stated
World Bank: The project was also, in line with the expressed request of the then newly and democratically elected Government for Bank's support in the emergency rehabilitation of critical infrastructure nodes.	Government	General assertion	Not stated
UNICEF: The programme has been relevant since its start and compatible with national priorities established at the National AIDS Programme (NAP)	Government	Cite policy (programme) aligned with	Not stated

Retrospective Comparative Analysis of Data

This study uses a rigorous retrospective comparative analysis based on systematically coding evaluations. The sample includes 160 reports, 40 from each of the four selected multilateral agencies—ILO, UNDP, UNICEF, and the World Bank—chosen for their diverse mandates. Reports are equally divided between 2013 (late MDG-era, pre-revision of DAC criteria) and 2023 (SDG-era, post-revision of DAC criteria) to track methodological evolution. We used the relevant evaluation repository for each agency to create a sample frame of all evaluations for each agency for each of the selected years, randomly selecting 20 reports for each agency/year by using the random number generator in Excel. In the case of the World Bank we consider Implementation Completion and Results Reports.

Data were manually extracted with team members each assigned one agency: ILO (PB), UNDP (DD), UNICEF (XW), World Bank (MO). The coding form was revised after piloting. Most notably, we realized that coding had to be at the level of the individual relevance claim, as most reports contain several claims, the coding for each of which may vary. Analysis is presented both at the level of the relevance claim, and at the level of the report.

3. Findings

The amount of text devoted to assessing relevance can range from a few paragraphs to several pages, with UNICEF reports in particular often having an extended discussion of relevance. Hence the number of relevance claims per report shows great variation, being notably lower for UNDP and the World Bank than for ILO and UNICEF. There is a modest increase in the number of claims per report over the ten years from 2013 to 2023 for three of the agencies, the exception being ILO, though it remains clearly in 'top place'.

Table 1: Average number of relevance claims per study

Agency	2013	2023	N (studies)
ILO	14.6	12.2	20/20
UNDP	3.0	3.2	17/18
UNICEF	8.7	9.9	20/20
World Bank	2.8	3.4	20/19

Note: N (studies) refers to the number of reports in the sample for each agency and period. In the case of UNDP the sample included four evaluations in French, which we excluded. One UNDP report and one World Bank report did not assess relevance.

Relevant to whom?

We analyse the data both at the level of the individual relevance claim - of which there are 572 claims in total in 2013 in 77 reports and 551 claims in 2023 also in 77 reports - and at the report level (154 reports). A report is coded against a code if it has at least one relevance claim with that code.

Across all agencies the largest share of claims are of relevance to the country or government. This is generally a bit over one-third of all claims, though in the case of the World Bank it is over half (Figure 1a). At the report level, virtually all reports make at least one claim of relevance to country or government (Figure 1b).

For ILO and UNICEF the second most important named stakeholder is a partner or civil society agency; at the report level the great majority of reports include this category. This is less so for UNDP and the World Bank. This difference is explained by the differing nature of the operations between the agencies. ILO works on a tripartite basis, so most of its programmes are intended to include government, trade union and private sector agencies (e.g. chambers of commerce), with the latter two types of agency falling in this category. UNICEF implements its programmes through national implementing partners, whereas World Bank programmes are most commonly "government executed" and many UNDP programmes are with government agencies.

There has been a slight increase in the proportion of evidence claims that are made at the global level (except for the World Bank), most notably for UNDP at the report level. This supports the argument that the broadening of the development agenda from the MDGs to the SDGs has made it easier to claim global relevance, though the effect is modest.

Figure 1a

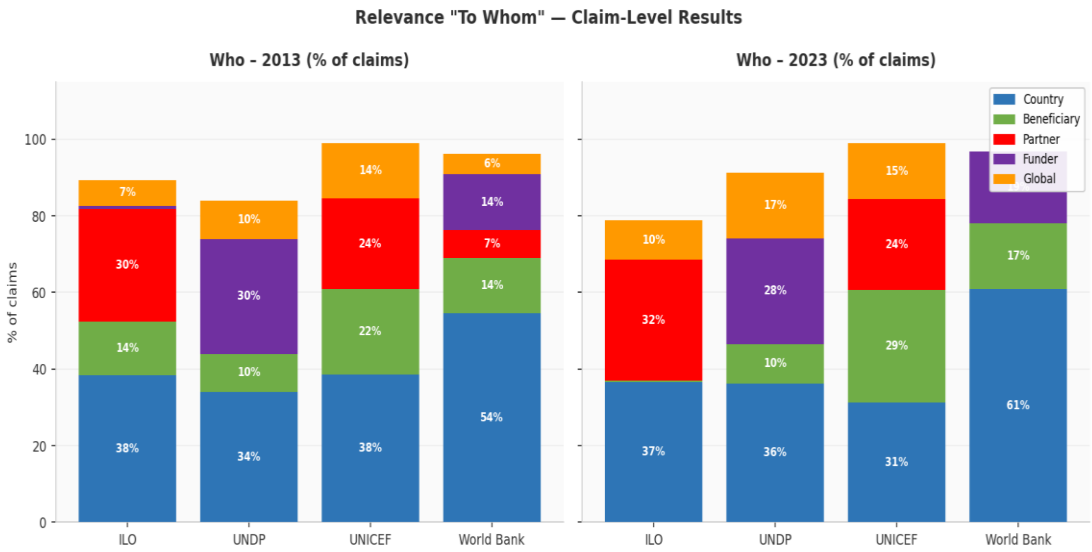
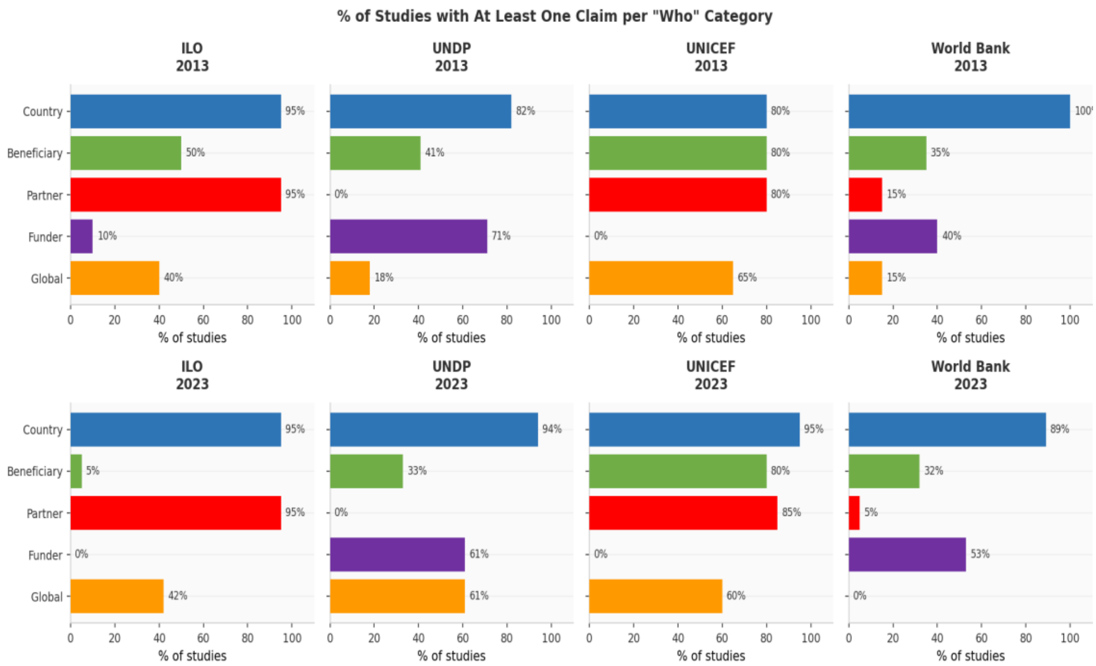


Figure 1b



Source: Authors' dataset

Finally, both UNICEF and the World Bank claim relevance to themselves, which ILO and UNDP almost entirely do not do. Relevance to the funder is not included in the DAC definition and it is open to discussion whether this is a good practice or not. The

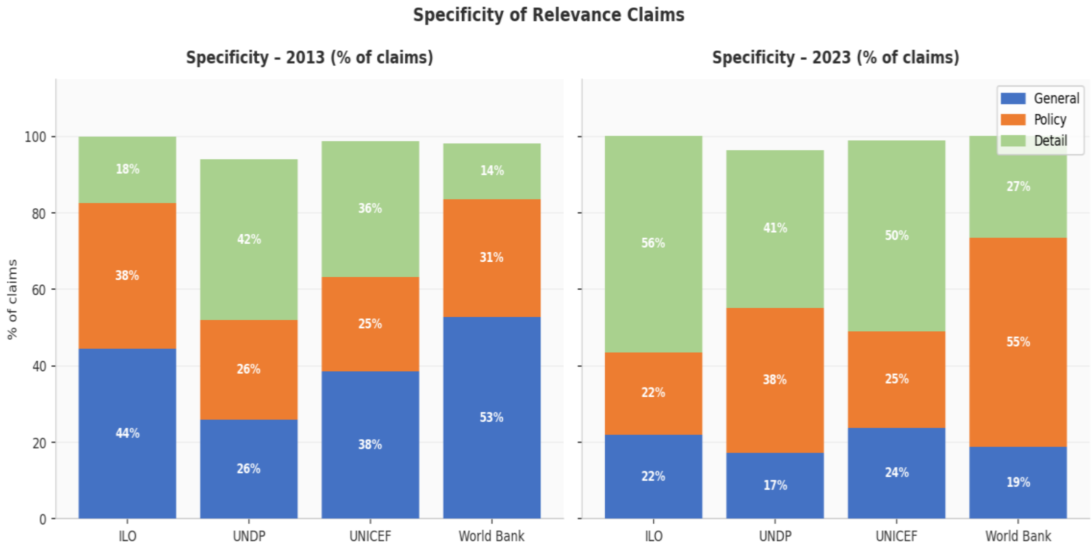
World Bank's review of country programmes argued that there was a "line of sight" from individual projects to country priorities provided projects were aligned with the Bank's CPF and the CPF was itself aligned with country priorities, though observed that this line of sight referred to high-level objectives (IEG World Bank, 2022).

Specificity of the evidence claim

There has been a shift over time for all agencies except UNDP from general relevance claims to more detailed claims (Figure 2). This shift represents a clear improvement, since general relevance claims are the easiest to make given the "everything is relevant" problem described above.

UNDP had the highest share of detailed claims in 2013, though in 2023 this position was taken by ILO. For the World Bank and UNDP a greater share of relevance claims are based on named policies, which in 2023 was the case for over half of all claims for the World Bank.

Figure 2 Specificity of evidence claims (analysis at claim level)



Source: Authors' dataset

At the study level, virtually all reports contain at least one claim at each specificity level.

UNICEF is notably consistent: 85–95% of studies include general, policy, and detail claims in both periods. The World Bank stands out for reaching 100% of studies having at least one policy citation in 2023, while its use of general assertions fell sharply (70% to 32%).

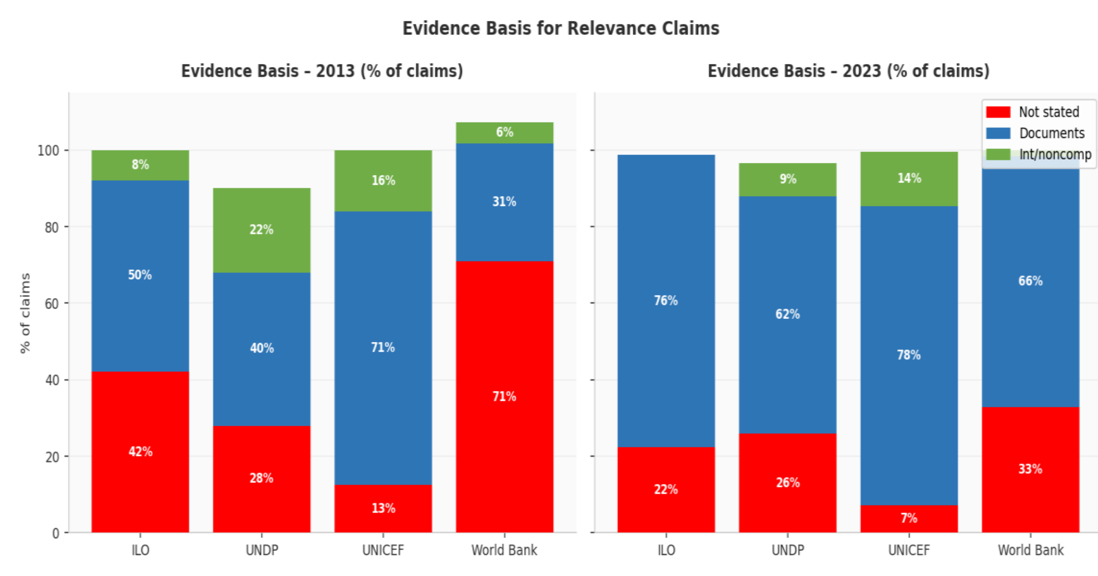
Basis for the relevance claim

The period between 2013 and 2023 saw a consistent and substantial shift from "not stated" to documentary evidence as the cited basis (Figure 3). This fall is particularly strong for the World Bank, for which over 70% of claims were unsupported in 2013. Documentary evidence is now the dominant basis for relevance claims across all agencies and both periods, accounting for 50–78% of claims in 2023. At the study level, documentary evidence is now near-universal: 89–95% of studies across all agencies make at least one document-backed relevance claim in 2023.

Nonetheless, the proportion of studies with any unsupported ("not stated") relevance claims remained high: even in 2023, 42–56% of studies contain at least one claim with no stated basis. While documentary evidence has become more prominent, evaluators continue to include some unsubstantiated assertions.

Interviews are rarely cited as a basis for relevance claims. Since key informant interviews are usually the main type of data collection for process evaluations, this likely reflects underreporting of the source of claims, possibly for reasons of confidentiality. However, transparency of reporting matters for the credibility of any evidence claim, so this approach needs reconsideration. The interview sources are all used for non-comparative relevance claims. We found just one exception in which a comparative claim was made in the evaluation of UNV's International Year of the Volunteer: "Although UNV's efforts were directly aligned with its Corporate Priorities, some key informants felt that other priorities, specifically mobilization, suffered because of the extensive focus on IYV+10" (Kallick Russell Consulting, LLC, 2013: 11).

Figure 3 Evidence based for relevance claims (claim level)



Source: Authors' dataset

Discussion and conclusion

We have used three criteria to assess the application of the DAC relevance criterion in the evaluation of development interventions. There are some marked patterns in our comparative analysis across time and between agencies.

Relevance is mostly claimed as being relevant to the country or national government. Next most important for ILOand UNICEF are national partners, whereas for UNDP and the World Bank the second most common relevance claim is relevance to themselves as funders. There has been a slight increase over time in claiming global relevance.

The most marked change in the data is a shift from general relevance claims to those based on a named policy (UNDP and World Bank) or providing specific detail of how intervention design aligns with government activities (ILO and UNICEF). There has also been a reduction in the share of unsubstantiated relevance claims for which the evidence base is not identified. The main source of support for evidence claims is reports; interviews are rarely used for this purpose.

Our final reflection is a methodological one. The analysis we have performed here - that is examining the basis for the relevance claim in evaluations - is currently not common practice in evaluation. This situation needs to change. To support such a change the appendix includes a proposed checklist to be used by evaluators in preparing relevance sections, and for evaluators using evaluations either as part of a strategic or thematic evaluation or as a user.

The use of reporting standards, as contained in the checklist, is common practice in studies of health interventions and for systematic reviews, but not more generally in the social sciences or the evaluation community. Use of the checklist can ensure more comprehensive coverage by the evaluation and more rigorous use of evidence. Without such an assessment we do not know how much confidence we can place in relevance claims.

Paul Mosley long ago identified the macro-micro paradox in aid effectiveness that whilst project evaluations find 80% or more of projects to be satisfactory the macro literature finds no relationship between aid and growth (Mosley, 1986). The new version of this paradox is that the figure of 80%+ continues to be reported based on process evaluations but impact evaluation findings support the 80% rule that the majority of interventions in both developing and developed countries have little or no effect (White, 2019).

The main explanation for this paradox is that process evaluations have no business making causal claims. To the extent that they do it is largely based on a before-versus-after data for outcomes of interest which is likely to be biased as it attributes to the intervention welfare improvements which were caused by other factors. But another part of the reason is a possibly weak evidence base for any evidence claims, including causal claims, made in the evaluation. Our main conclusion is therefore a call for a concerted effort to develop sound evaluation standards for evaluations of development interventions.

Appendix Checklist for assessing relevance

This checklist combines considerations from the OECD document *Applying Evaluation Criteria Thoughtfully* (OECD, 2021) with the three dimensions proposed in this paper. It may be used either by an evaluator conducting an evaluation or for an assessment of how an evaluation has considered relevance.

No.	Item	Relevance claim 1	Relevance claim 2
1.	Are the objectives of the intervention (i) adequately defined, (ii) realistic and (iii) feasible?		
2.	Is there a clear statement as to how the intervention is relevant to the context?		
3.	Has possible relevance been considered to each of (i) country or government, (ii) beneficiaries, (iii) partners and institutions (other development partners engaged in the intervention and civil society organizations), (iv) funding agency, and (v) global priorities?		
4.	Does the intervention reach those most in need?		
5.	Consider whether the intervention is designed in a way to meet stakeholder priorities and needs?		
6.	If relevance is claimed with reference to a policy or a strategy, is that policy or strategy explicitly named and referenced, with a link to the policy or strategy document where available?		
7.	If relevance is claimed based on intervention design or a specific component, is the design feature or component clearly identified, and is it clear how that supports the relevance claim?		
8.	Is the evidence basis for the relevance claim clearly stated, with references provided as appropriate?		
9.	Consider whether relevance has changed in the light of changes in context or priorities.		

References

Dissanayake, R., & Lowcock, M. (2024). *The rise and fall of the Department for International Development*. Center for Global Development. https://www.cgdev.org/sites/default/files/rise-and-fall-department-international-development_0.pdf

Kallick Russell Consulting, LLC. (2013). *The summative and forward-looking evaluation of the marking of the tenth anniversary of the International Year of Volunteers: Final report*. United Nations Volunteers Programme.

Mosley, P. (1986). Aid-effectiveness: The micro-macro paradox. *IDS Bulletin*, *17*(2), 22–27.

Nilsson, M., Griggs, D., & Visbeck, M. (2016). Policy: Map the interactions between Sustainable Development Goals. Nature, 534(7607), 320–322. https://doi.org/10.1038/534320a

OECD DAC (2019). Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use. Paris, OECD Publishing. https://www.oecd.org/en/publications/2019/12/better-criteria-for-better-evaluation_f7a307eb.html

OECD. (2021). Applying evaluation criteria thoughtfully. OECD Publishing. https://doi.org/10.1787/543e84ed-en

OECD. (2002). Evaluation and aid effectiveness no. 6: Glossary of key terms in evaluation and results based management (in English, French and Spanish). OECD Publishing. https://doi.org/10.1787/9789264034921-en-fr

Scottish Government. (2016). *Youth Employment Initiative (YEI): Evaluation report for the 2014–2020 Scottish Operational Programmes*. Scottish Government. https://www.gov.scot/publications/yei-evaluation-report-2014-2020/

White, H. (2019). The twenty-first century experimenting society: the four waves of the evidence revolution. *Palgrave Communications*, *5*(1), 47.

World Bank (2002). *Social funds: Assessing effectiveness*. World Bank. https://ieg.worldbankgroup.org/sites/default/files/Data/reports/social_funds.pdf

World Bank (2017). Evaluation of the World Bank Group's Country Partnership Frameworks. Washington, DC: Independent Evaluation Group, World Bank.

World Bank IEG. (2022). Results and Performance of the World Bank Group, 2022. Washington, DC: Independent Evaluation Group, World Bank.