STUDY QUALITY ACTION PLAN

Recommendations Based on the Study Quality Review

CONTENTS

ACKNOWLEDGEMENTS

ACRONYMS AND ABBREVIATIONS

ASQ Tool	Assessment of Study Quality Tool
BE2	Building Evidence in Education
CIES	Comparative and International Education Society
DEC	Development Experience Clearinghouse
DEEP	Data and Evidence for Education
ERC	Ethical Review Committee
IP	Implementing Partner
IRB	Institutional Review Board
IRC	Inter-Coder Reliability
IRR	Inter-Rater Reliability
RCR	Responsible conduct of research
USAID	United States Agency for International Development

BACKGROUND AND PURPOSE

In 2022, USAID commissioned a team under the Data and Evidence in Education Program, or DEEP, to revise the evaluation quality tool to capture updated international best practices in research and evaluation and updated USAID guidelines and policies. The commission included a mandate to revise the tool so that it is more broadly applicable to research and evaluation in education and the social sciences. It also included a second review of the quality of USAID funded studies in education. This document presents recommendations informed by the findings from the review. A brief overview of the revised Assessment of Study Quality (ASQ) Tool, the study quality review, and the top findings are included in Annex A.

RECOMMENDED ACTION PLAN

To inform USAID's actions regarding this tool, the DEEP team proposes the following action plan based on the results of the review and collected feedback.

IMMEDIATE ACTIONS (0-6 MONTHS)

TO IMPROVE USAID STAFF CAPACITY TO USE THE TOOL:

Socialize the ASQ Tool with USAID Staff, including Missions, Regional Bureaus, and other Operating Units:

- Share the ASQ Tool, the updated USAID evaluation report template, and the updated USAID evaluation design template with USAID staff, including Regional Bureau and Mission staff (education teams, program office, M&E teams).

Integrate the ASQ Tool into formal and informal processes used by USAID Operating Units to commission research and evaluations:

- Encourage USAID staff to include references to the ASQ Tool in formal requests for proposals or task orders:
    - Include the ASQ Tool One-Page Checklist, the definition of each Principle of Quality, and the Edu-Links hyperlink to the tool in RFPs or task orders;
    - Include a statement that the research or evaluation activity is expected to adhere to each of the Principles of Quality in the quality control section of RFPs or task orders;
- Encourage USAID staff to share the ASQ Tool with IPs when discussing expectations for research or evaluation studies:
    - Include a copy of the ASQ Tool in research or evaluation planning and kick-off correspondence;
    - Develop a slideshow about the ASQ Tool to include in kick-off meetings;
    - Include time during planning and kick-off meetings to discuss expectations around the ASQ Tool with the IP.

TO IMPROVE IP AND RESEARCH/EVALUATION PARTNER CAPACITY TO USE THE TOOL:

Socialize the ASQ Tool with IPs and other research and evaluation partners:

- Develop supporting materials to socialize the ASQ Tool and support understanding of the tool, including:
    - Infographics about when and how to use the ASQ Tool;
    - Short video(s), approximately 3-minutes long, that clarify expectations about when and how to use the tool.

- Share the ASQ Tool and these materials with existing networks for distribution through USAID network listservs and newsletters, such as LTLGP managed networks and the Basic Education Coalition.

- Encourage IPs to share the ASQ Tool with their partners and consultants prior to designing education studies.

- Encourage IPs to include the ASQ Tool or refer to the tool in research or evaluation scopes of work (SOWs).

SHORT-TERM ACTIONS (6-12 MONTHS)

TO IMPROVE USAID STAFF CAPACITY TO USE THE TOOL:

Integrate the ASQ Tool into existing research and evaluation trainings given to USAID Education Officers and FSNs in education.

TO IMPROVE USAID STAFF CAPACITY TO OVERSEE QUALITY RESEARCH AND EVALUATIONS:

Develop and hold a series of targeted trainings or guidance materials about the study process based upon the ASQ Tool for USAID staff. These can be synchronous trainings, asynchronous trainings which include videos or e-modules, or visual materials such as infographics or how-to documents.

1. **Improve USAID staff capacity to oversee the design and implementation of research and evaluations.** Develop standalone guidance materials such as infographics or short videos that provide a quick and basic overview of the key elements of study quality identified as the weakest in the review. Guidance materials could include:

- An infographic or short video which provides a quick overview about common types of biases and common mitigation strategies to improve USAID staff's ability to assess the robustness of a methodology and validity of a study proposed and implemented by the research/evaluation team. This could also be shared with IPs and research/evaluation partners.

- An infographic or short video about the purpose and logistical requirements of piloting data collection instruments with an emphasis on the additional time and effort required for IPs to use pilot results to adjust the data collection instruments to support USAID staff's ability to support the cultural appropriateness of a commissioned study.

2. **Develop and host trainings (synchronous or asynchronous) or guidance materials to improve USAID staff capacity for the use of research and evaluation findings:**

- Develop asynchronous training materials or e-modules about using the ASQ Tool to assess the quality of a report, especially the "Question Description and Source" table in Annex B. This could include updating the orientation materials used for the review or creating an annotated bibliography from the sources used in the ASQ Tool.

- Encourage USAID staff to use the ASQ Tool when reading a report to keep track of areas of strength and weakness. To facilitate this, develop a markable PDF version of the ASQ Tool.

In all guidance or training materials, include clear statements that this is an overview meant to support USAID staff general understanding and is not comprehensive.

TO IMPROVE IP AND RESEARCH/EVALUATION PARTNER CAPACITY TO PRODUCE QUALITY RESEARCH AND EVALUATIONS:

Develop and hold a series of targeted trainings or guidance materials about the study process based upon the ASQ Tool for IPs and other partners. These can be synchronous trainings, asynchronous trainings which include videos or e-modules, or guidance documents such as infographics or how-to toolkits.

1. **Develop and host trainings (synchronous or asynchronous) or guidance materials to improve IP capacity in designing and implementing research and evaluations, with an emphasis on the weakest items from the weakest principles of quality identified during the review.** Trainings could use an asynchronous format like the training provided to the study quality review participants, which includes short video modules along with links to external resources. Guidance materials could take the form of infographics with links to external resources or case studies of USAID-funded studies which met or exceeded the criteria for the weakest elements identified during the review.

- Share with partners the materials developed to enhance USAID staff understanding about biases and mitigation strategies with partners to support IP capacity in robustness of methodology.

- Develop infographics or briefs about appropriate sampling design to support IP capacity in robustness of methodology. For example, an infographic could present an overview about different types of sampling designs, how they support different study objectives, and strengths and weaknesses of different sampling approaches.

- Develop asynchronous training modules or how-to kits with a set of guidance notes about different ways to engage meaningfully with local stakeholders during the study design process to support IP capacity in designing a cultural appropriate study. For example, modules could include short videos or guidance notes which present potential ways to identify different relevant stakeholders (i.e., students, teachers, parents, Ministries of Education or other government officials, research partners), strategies for engagement at different stages of the study, such as consulting with community leaders about the appropriateness of data collection instruments, and holding interactive sessions with teachers (e.g.) to validate findings and generate recommendations.

2. **Develop guidance materials to support partner capacity in writing research and evaluation reports.** To ensure that report writing remains strong, develop guidance materials

such as briefs, infographics, or job aides about the low-scoring items under openness and transparency and cogency:

- Develop a brief which presents segments or blurbs from USAID-funded education study reports that discuss bias to support openness and transparency in reports written by partners. The brief would identify the strengths and weaknesses in each of the examples, along with writing tips or counterexamples to improve the highlighted weaknesses.

- Develop a job aid to present strategies to include evidence that each item on the ASQ Tool has been addressed in a study report, such as through the use of footnotes and annexes, to support partners' ability to demonstrate cogency in their reports.

TO REINFORCE USAID STAFF AND PARTNER UNDERSTANDING ABOUT ETHICAL REQUIREMENTS FOR RESEARCH WITH HUMAN SUBJECTS:

Incorporate ethical expectations in all USAID trainings about research and evaluation. Include in every guidance or training material a link to USAID's ethical requirements for research with human subjects. Clearly state USAID expectations regarding IRB training and RCR certification.

LONG-TERM ACTIONS (12+ MONTHS)

TO IMPROVE IP AND RESEARCH/EVALUATION PARTNER CAPACITY TO USE THE TOOL AND PRODUCE QUALITY RESEARCH AND EVALUATIONS:

Hold virtual interactive sessions about the ASQ Tool to engage with researchers and evaluators working in education subsectors:

- Sessions can provide a forum for researchers and evaluators to share their approaches to applying best practices for each principle of quality.

- Sessions can include informal brown bag discussions and interactive webinars where participants share their approaches to and discuss their challenges with research and evaluation using the ASQ Tool.

- Sessions can be hosted by an existing USAID mechanism, such as Share, LTL, Decode, etc.

Develop additional supporting materials to support understanding of the tool, such as:

- Annotated examples of report sections, based upon the updated evaluation report template, to illustrate the type of information required to demonstrate adherence to the questions under each principle of quality (see the example shared by a reviewer in Annex D).

- Case study briefs and infographics to enhance understanding of items in the tool that fall under strong principles of quality yet were weakly adhered to, according to the review findings. These include:
  - To ensure that conceptual framing does not continue to weaken, develop a case study brief about the difference between study questions and study hypotheses. A case study could present a real or hypothetical USAID-funded education study that included both study questions and study hypotheses. The case study would be presented as a brief that

explains why hypotheses were necessary to the study, how the study purpose, questions, and hypotheses connect, and how the study report presented the questions and hypotheses.

- o To enhance IP capacity in reliability, develop infographics which outline when and how to establish inter-rater reliability and inter-coder reliability.

TO IDENTIFY NEXT STEPS FOR ONGOING USAID STAFF AND IP AND RESEARCH/EVALUATION PARTNER CAPACITY BUILDING:

Conduct a rapid study of ASQ Tool socialization efforts to determine next steps to support for long-term tool use:

- Share a survey with USAID staff, including Missions, Regional Bureaus, and other Operating Units to collect data on whether and how frequently they incorporate the tool in task order requests and other RFPs, the extent to which they share the tool with IPs and partners, how the tool was shared with them, and other details.

- Share a survey with IPs and partners to collect data on whether and how frequently they used the tool in the past year, how frequently they refer to the checklist, how the tool was shared with them, and other details.

- Use the findings from the surveys to determine whether additional socialization is required and to develop a long-term plan to support ongoing use of the tool.

- If the survey indicates high levels of ASQ Tool use, consider another rigorous review of USAID-funded education studies to determine the impact of ASQ Tool use on the quality of reports.
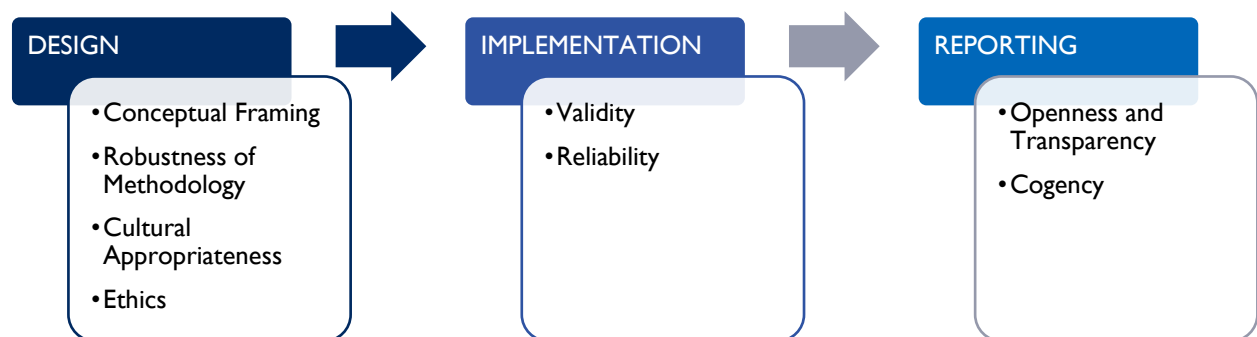
ANNEXES

ANNEX A: OVERVIEW OF THE ASSESSMENT OF STUDY QUALITY (ASQ) TOOL, STUDY QUALITY REVIEW, AND TOP FINDINGS

ABOUT THE ASSESSMENT OF STUDY QUALITY TOOL

In collaboration with the USAID Center for Education, the DEEP team revised the Evaluation Quality Assessment Tool using the same Building Evidence in Education (BE2) framework to assess the principles of quality used for the previous tool. The revised tool was renamed the Assessment of Study Quality (ASQ) Tool. The revised ASQ Tool is comprised of 35 items under eight principles of quality that are grouped the principles by study phase.[1] Exhibit 1 shows the principles of quality by study phase.

Exhibit 1: Principles of quality by study phase

DESIGN
- Conceptual Framing
- Robustness of Methodology
- Cultural Appropriateness
- Ethics

IMPLEMENTATION
- Validity
- Reliability

REPORTING
- Openness and Transparency
- Cogency

ABOUT THE REVIEW

In January 2023, the DEEP team conducted a review of the quality of 122 USAID-funded education research and evaluation studies published on the Development Experience Clearinghouse (DEC) between January 2018 and 2022. A group of 92 volunteer reviewers based in 26 countries reviewed the reports using an online version of the ASQ Tool. The DEEP team analyzed the reviewer responses to determine the percent of reports that met "Minimum Adequacy" for each principle of quality, to identify the strongest and weakest principles of quality, to identify item-level trends across reports, and to determine changes since the 2017 review.[2,3]

The purpose of the review was to support USAID's assessment of the quality of education evaluations and research commissioned by USAID since 2017, get feedback about the tool to help address any

---

[1] Ethics was added as a principle of quality in the 2023 tool.
[2] The methods are detailed in the full report, "The State of Quality in USAID-Funded Education Research and Evaluations: Results from the 2023 Study Quality Review".
[3] Comparisons with the 2017 review should be taken lightly as the reviews used different methodologies, different tools, and included different data sources.

outstanding issues and continue to improve the tool for future use, and give reviewers the opportunity to practice using the tool, giving them insights for their own research and reporting practices. In addition, the review provided an opportunity for reviewers to share ideas about how USAID can encourage continued use of the ASQ Tool and continue to support the quality of education studies.

REVIEW OF FINDINGS AND CONCLUSIONS

This section provides a brief overview of the key findings which informed the recommended action plan. Detailed findings and conclusions from the review are presented in "The State of Quality in USAID-Funded Education Research and Evaluations: Results from the 2023 Study Quality Review".

- Most reports reviewed met "Minimum Adequacy" for at least one out of the eight principles of quality, and most reports met "Minimum Adequacy" for multiple principles of quality.
  - 86.1 percent of reports reviewed (n=105) met "Minimum Adequacy" for at least two principles of quality.
  - Over half of the reports (n=64) met "Minimum Adequacy" for five or more of the eight principles.
- Reports received stronger scores under the Report Writing phase.
  - Cogency is the strongest principle of quality despite anecdotal evidence to the contrary: 86.1 percent of reports met "Minimum Adequacy" based on reviewer scores. However, consumers of reports who did not participate in the review, such as decision makers and practitioners, often indicate that the reports are not easy to understand or useful.
  - Openness and transparency is the third strongest principle of quality: 60.7 percent of reports met "Minimum Adequacy".
- Reports received the lowest scores under the Study Design phase.
  - The weakest principles are ethics (met by 36.9 percent of reports), cultural appropriateness (met by 40.2 percent of reports), and robustness of methodology (met by 40.2 percent of reports).
- The two principles related to study implementation, validity and reliability, are stronger than robustness of methodology, cultural appropriateness, and ethics, three of the four principles under study design.
  - This could mean that study implementation is stronger than study design, or that report writing is not accurately capturing each of the items under all principles of quality.
- Reports have improved in all but two principles of quality since 2017.
  - Cogency remains the strongest principle of quality for both reviews.
  - Cultural appropriateness slightly improved, from the weakest principle in 2017 to the second weakest principle in 2023.
  - Validity and openness and transparency experienced the strongest gains, by 14.7 percentage points and 11.7 percentage points respectively.
  - Robustness of methodology and conceptual framing decreased, by 24.8 and 3.9 percentage points respectively.
- At the item-level, items related to biases received the lowest scores in all three study phases.
  - This seems to point to a pattern that planning for and mitigating biases is challenging for those conducting education studies as is reporting on bias.

- o This also highlights how connected all study phases are.
- At the item-level, reviewers indicated there is uncertainty about when inter-rater reliability and inter-coder reliability are necessary.
- Each finding may be related to report writing.
  - o All the items under each principle of quality may have been addressed by those who designed and implemented the studies, but if there is no evidence provided in the report, the reviewers' judgment may not reflect the study design and implementation.
  - o This could be an issue of reviewer participation bias, as those who participated in the study are experts, researchers, evaluators, and those who understand the technical jargon.
  - o The reviewers' experience may have allowed them to make connections and assumptions about the items in the ASQ Tool that are not clearly reported, raising the possibility that reviewer participation bias exists for all principles of quality.

REVIEWER FEEDBACK

Following the review of the reports, reviewers were asked to provide feedback on the ASQ Tool. Reviewers were asked some closed ended questions using a Likert scale and they were able to provide written comments in their survey responses as well, to inform final revisions of the tool. Reviewer responses to the closed questions and their additional comments were downloaded into tables included in Annex B. Most reviewers found the tool relevant and easy to use to inform the research and evaluation process, and most indicated that they would use the tool in their own research and evaluation work. In addition, most reviewers indicated that they would share the tool with their colleagues.[4]

In February 2023, the DEEP team collected feedback about the ASQ Tool following a presentation of the tool and the preliminary findings at the 2023 Comparative and International Education Society (CIES) Annual Conference. The feedback was captured in a Google Document included in Annex C. Additionally, the DEEP team collected qualitative feedback from reviewers through two findings workshops held in March 2023. Participants shared their experience using the tool during the review process and offered suggestions to improve the tool, encourage the use of the tool, and improve the quality of USAID-funded education research and evaluation. Many of these suggestions were taken into consideration during the final ASQ Tool revision process. The suggestions were captured in a Google Document, which is included in Annex D.

The following themes emerged from the collected feedback:

- The tool is complex.
  - o It is long and takes a lot of time to use to review a report.
  - o It includes extensive details which support reviewer understanding but which makes it challenging for non-researchers/evaluators to understand.

---

[4] For more details, refer to the study report, "The State of Quality in USAID-Funded Education Research and Evaluations: Results from the 2023 Study Quality Review".

- Different forms of the tool should be readily available, including the version with extensive guidance and a short version that may be easily referenced by those implementing a study.
- The tool should inform additional materials to support quality of USAID-funded education research and evaluation, such as:
  - A short version of the tool centered on the "critical questions".
  - Annotated reports which provide exemplary, good, and bad examples of the items in the tool in existing reports.
  - Use cases to clarify when to use the ASQ Tool instead of other tools.
- The tool should be socialized widely with different audiences:
  - Research and evaluation networks.
  - Academics and graduate programs.
  - Missions and Implementing Partners (IPs).
  - Project teams and report writers.
- Various socialization approaches were proposed:
  - Develop materials which include infographics and videos.
  - Provide mentorship and specialized trainings.
  - Hold webinars and share updated versions of the existing orientation materials.
- There should be future reviews:
  - Sustained use of the tool will provide greater clarity on how to refine the tool.
  - Inter-rater reliability should be established.
- There should be clarity about USAID's expectations regarding this tool:
  - Is it required?
  - Will it dictate future funding eligibility?
- Acknowledge and address existing challenges:
  - Time constraints, which make ethical clearances challenging.
  - Intended audience (i.e., what to include in briefs vs. research reports).

ANNEX B: FEEDBACK TABLES AND COMMENTS FROM THE REVIEW

On a scale of 1-5, how easy was it to use the tool? (with 1 being easiest, 5 being hardest) (n=115)

AVERAGE SCORE	1 (EXTREMELY EASY)	2 (MOSTLY EASY)	3 (SOMEWHAT EASY)	4 (MODERATELY HARD)	5 (EXTREMELY HARD)
2.5 (Mostly – Somewhat Easy)	32.2% n=37	24.3% n=28	14.8% n=21	18.0% n=21	10.0% n=12

COMMENTS

A bit long, and some of the explanations of each question might have been a bit clearer. Had to draw on a lot of expertise and understanding outside of what was explained in the questions, but perhaps that was intended!

As a researcher I used many similar tools and I find this to be easy and self-explanatory.

As someone who has used and participated in developing similar types of quality assessment tools, and also as an experienced researcher and evaluator, I feel comfortable with these kind of reviews in general. What made the tool manageable, was that the questions were organized by sections which helped me navigate around a 159-page report. However, my experience is that this tool is extremely cumbersome to use and would not be realistically deployed at an implementation level. I think many of the questions are difficult to assess even with clear guidance. Especially for local implementers. I struggled with the meaning of a couple of questions (I noted in the comments). I think the number of questions and the time it takes is simply too long for anyone to want to use this tool in practice. It may be a useful tool for research purposes to assess quality of evaluations. But I would not use this tool for implementers to assess quality of their evaluations. I would take this tool and transpose it into a "writing template" so that way in each report section, you can have the guidance what to include based on the questions in this tool. For example, guidance that there is a section on stakeholder engagement in the evaluation, would encourage more submitters to include this in the reports more clearly, or even to design their evaluations with this in mind in the first place.

Clear instructions and guidelines are provided to explain each question and corresponding sections

Easy and user friendly.

I am not familiar with some of the statistical analysis questions.

I found the 'Research and evaluation quality assessment tool' to be very good and well elaborated. I was also able to refer to it easily during the review. The fact there is direct referencing from the review questions to the items in the assessment tool, makes it user friendly, and the review process is made easy.

I found the tool readily accessible with only minimal review of the quality assessment tool PDF beforehand.

I missed the sessions. Maybe with the above sessions. Review quality will improve

I think the tool is quite simple to use, at least for a qualitative study. It helps that key terms are provided in the tool to search for in the document.

illustrations are useful - especially for assessing the "partial" cases

In general, using the tool was easy. However, while assessing particular study in some questions it was a bit difficult to decide between yes/no/partial because some issues were implicit in the study. For example, it was obvious that authors addressed validity of findings and next steps, but transferability is something to be read between the lines. Hence, it's not no because it contains suggestions for overall transferability. It's not really yes because suggestions are implicit. And it's not partial because partial score in the tool is described as only if some findings are transferrable.

It is a user-friendly tool. Very well organized and it follows the structures of reports

It is clearly stated and easy to follow.

It is easy to complete this survey because I read first the report.

It was difficult for the first time

it was easier to use on a mixed methods / quantitative tool than a qualitative tool.

It was easy to use this tool. However, I needed to scroll back and forth through the sample report (>100 pages) a lot to find the answers which was a little bit tedious. If projects follow the USAID report guidelines very carefully so their report includes the required components, the report would be easier to evaluate with this tool.

It was fairly easy to use. Having the content listed in the survey questions was helpful.

It was fairly straightforward, and the instructions were clear, but it was a little long.

It was lengthy. It took a significant amount of time to look for different elements, as the order did not match the order of the questions. It may be an easier tool to use for self-assessment.

It was quite easy, but some theories / concepts were new for me that I had to understand them first.

Nicely designed and easy to use.

Not all the questions are relevant for this study.

Not friendly usable

Questions are clear and explained in this tool that it has been easy for me to complete the survey after reading the assigned report. However, some concepts and theories (in the explanation) were new for me and required further reading to understand them.

Seems straightforward, I was initially a little confused about the different order of questions in the online tool compared to the how-to document.

simple

The evaluation was clear and provided supporting explanation to assist with my responses

The explanatory texts of the questions are certainly useful but are sometimes perhaps too extensive.

The reports are long, like this Zambian ECE research. The tool is clear.

The resources, listening session, video, google questions support ease of access, understanding and use of the tool

The tool has excellent explanatory notes, making it user friendly

The tool is easy to use because if brings clarity to the entire research process by providing practical examples what to look for in each specific item

The tool is easy, but long, I would suggest less questions, some of them were repetitive.

The tool is friendly

The tool is very detailed and self-explanatory

The tool made the review process easier, but there are still degrees of subjectivity. P.S. The wording of this scale question may cause confusion among respondents.

The tool was easier to use with the pure qualitative research report. My previous report was a mixed methods study.

The tool was fairly easy to use, but in some cases I would have appreciated more guidance. For example, one of the questions was to check if the study has a hypothesis if applicable. But I wasn't sure if the type of study I was reviewing needed to have a hypothesis.

The tool was very straightforward and easy to use. The only challenge I ran into is that this particular study is a summary of other studies, including both quasi-experimental and qualitative studies. That was not a clear choice when I was asked which type of study it was. Similarly, some of the questions are targeted toward impact evaluations. I struggled with whether I should answer those since this was not an impact evaluation but contained analysis of impact evaluation data.

The use of the tool was made easy by the fact that I watched the video and read the pdf document before undertaking the task to read the report. In addition, the report takes into account various possible answers/scenarios (yes, no, partial, NA)

This tool was easy to use except for the fact that this was a baseline study rather than an evaluation. So, some of the questions were not applicable but will become applicable once the post-intervention study is done.

User-friendly

Very time consuming, but generally not difficult to use.

Very useful and helpful

On a scale of 1-5, how relevant do you find the tool to research and evaluation? (with 1 being the least relevant, 5 being completely relevant) (n=113)

AVERAGE SCORE	1 (NOT RELEVANT)	2 (SLIGHTLY RELEVANT)	3 (MODERATELY RELEVANT)	4 (MOSTLY RELEVANT)	5 (COMPLETELY RELEVANT)
4.3 (Mostly – Completely Relevant)	2.7% n=3	2.7% n=3	6.2% n=7	35.0% n=39	54.0% n=61

COMMENTS

At the start, I needed assistants from Administration to identify the survey and studies. It was complicated and it gotten better as I progress, but I'm not sure if I did my best for the first time.

Good questions and considerations in assessing the quality of evaluations and research studies.

Great tool!! I suppose I could use a bit more training on it. I would be interested in becoming more of a regular evaluator, but there should probably be some compensation for the time it takes and level of expertise.

Highly relevant.

I don't think it goes sufficiently in-depth, but this depends on the ultimate purpose. As a superficial, general evaluation it terrific.

I find it more relevant to research than evaluation. I also find it heavily orientated toward impact and quasi/experimental approaches.

I read the document to ensure a comprehensive understanding of what was to be evaluated. The tool is excellent in that not only did it give a general reminder of the context for the evaluation parameter, but it also included specifics around it. The approach reduced the ambiguity of expectations.

I spent more time finding answers to the questions than actually considering the quality of the findings and recommendations. For someone who wants to put more energy into understanding the use of the report (than whether the report ticks all the boxes), the tool could be frustrating.

I think this depends on the intended audience. The Q&A states there are many different audiences. However, I don't feel that this tool is at a level that is reachable for many audiences. For USAID as an audience, they may want to see this level of detail around approaches/methods, etc. But for most peer reviewed research journals, this is overly detailed. A journal will allot 10-20 pages for research report. So a lot of this would not be required to demonstrate a sufficient rigor. Some areas I think are very relevant: stakeholder engagement, how tools were developed, IRB and ethics requirements (do no harm and protections/confidentiality). Other areas I think =are extremely detailed and cumbersome include some of the sampling details, some of the questions/details around inter-reliability measurements, stratifications, and alternative interpretations of findings.

I would use this tool and share it with my organization for all research activity going forward. It clearly and comprehensively covers all expectations for USAID research and present standards and guidelines for quality research in general

In general, tool is relevant but reviewed study shows that some requirements in the tool may be almost inadequate for the specific studies.

It has given me a different view of items

It helps quality check of most aspects of research.

It is relevant as an evaluation tool but not necessarily as a research tool. Those are usually more open and explore new questions.

It is relevant to research and evaluation as it touches important aspects of the research process.

It is relevant. However, it assumes that the reports follow a format that allows for assessing most of the aspects. In my case, I was assigned a summary report, which does not include all relevant aspects included in the tool.

It's applicable to all types of research activities and has clarity on where which type of research is applicable

it's relevant to evaluation quality (not so much research)

Relevant since it includes specific criteria for evaluation reports assessment.

Some of the questions did not apply for this particular type of study. Maybe add a "skip" function.

Some variations might be included to accommodate for various research methods (qualitative, quantitative, mixed, or other less common methods)

Sometimes using structured tools like this to assess diverse types of reports is challenging. It is also hard to judge a final evaluation report without additional context about the study's insights.

The components assessed by this tool should be included in all evaluation/research reports.

The relevance in the tool lies in the composite array of elements for overview, questions, descriptors, links to further resources - which provide teams with broader and deeper guidance on what makes high level research

The tool covers all elements of research report

The tool is very relevant and covers the most important topics.

The tool is very relevant because it explored all the aspects of research to guide both researchers and research quality checkers.

The tool provides a surface level review of the evaluation. It touches upon key elements that are important is set up as basically a yes, no, somewhat response to each question. Although it allows for a quick assessment of the evaluation/research project, it is difficult to provide an in-depth critique of the report.

There should be more guidance for qualitative. There is some, but studies often struggle with this type of inquiry.

This is a helpful tool to review and assess the quality of a study in a consistent manner.

this is a relevant

This is a very helpful tool, especially for those who are not specialists in evaluation.

This is an odd question - something is missing - it is relevant to evaluation quality (I think research should be taken out of this tool) since research serves a different purpose, though uses similar techniques as eval.

This tool is very similar to an exercise I was asked to do in grad school to evaluate research studies. So, I think it is very relevant.

very relevant and useful for assessing research

Very relevant to evaluate a research work

Very relevant. It has provided me with a good understanding on how to design and evaluate research in a systematic and logical manner.

Very strong approach to reports quality assessment.

Was expecting to get details on different item of the evaluation form

On a scale of 1-5, how likely are you to use this tool in your work? (with 1 being highly unlikely, 5 being extremely likely) (n=109)

AVERAGE SCORE	1 (HIGHLY UNLIKELY)	2 (SLIGHTLY UNLIKELY)	3 (MODERATELY LIKELY)	4 (LIKELY)	5 (EXTREMELY LIKELY)
3.9 (Moderately – Likely)	4.6% n=5	8.3% n=9	20.2% n=22	21.0% n=23	46.0% n=50

COMMENTS

as a reminder on what to include in a study

As above I consider the tool serves as a resource and guideline that can inform research design and ensure inclusion of critical elements for quality and robustness

Despite my answer to the previous question, I found the tool helpful, and I will reflect on how to use some of its insights for similar evaluation exercises.

I am teaching a postgraduate course in Applied Qualitative Research, and I have incorporated large parts of this tool in the teaching units.

I certainly would use it to ensure my reports cover all important sections

I design and implement research. This tool aligns with best practices expected by my own institution, and it is pleasing to see that USAID is implementing these expectations as well.

I found the tool interesting not only for assessing the quality of studies by others, but also as a reminder to me of what one needs to do when carrying out such evaluations.

I think this tool is helpful in organizing elements of a project to be sure to include during the development, implementation, and analysis phase. Also, it would be helpful in reviewing the tool while writing reports to make sure all elements are included.

I will adopt several components of this tool for some of our research work.

I will certainly introduce this too for reviewing our research papers in my Faculty

I will use this tool in my own academic writing, writing consultancies and when supporting students writing assignments.

I would like to use such a tool in my work with few changes about items by removing those which are similar.

I would use alternative tools, unless there is a requirement to provide an assessment of a report using this tool by the client. There are other quality assessment tools that take much less time and still get at the core areas of quality. The Movement for Community Led Development has a great Quality assessment Tool that is easier to use than this, and much quicker, especially for those who are not technical yet reviewing these kinds of study reports for quality/rigor. Some of the questions in this tool looked VERY similar to the MCLD tool.

I would use the tool if the project completing the evaluation had had access to the guidelines in advance. It was tedious to scroll through the 100 pages of the report in search of different components.

I would use this tool in my work and would certainly recommend it to those with less experience in research and evaluation.

I'd love to use it to assess evaluation reports we are receiving.

If I had to assess or evaluate study quality, this would be a helpful and relevant resource.

If I had to assess quality of evaluations or studies, I think this would be a valuable tool for that purpose.

If I need to do a thorough review of a handful of studies, I will definitely use this tool. But the sheer number of questions to get through makes this tool less practical for reviewing a large number of studies unless you have a team of reviewers.

I'm sharing with my team now to use in assessing our own evaluations.

In academic setting that I belong to, this tool is very helpful in writing different academic reports, papers, these, and supporting students in their writing. So, I will make use of it and share it with colleagues.

It may serve as a good reminder of what elements of a study should be in a write-up.

Maybe more as a framework or checklist, but yes!

Maybe some of the questions but not all.

Maybe some of the questions.

My work does not pertain to this, but if it did, I would use the tool.

Not because the tool content isn't useful. It's because the focus is mostly standard practice. With a few exceptions, such as cost analysis.

not useful for what I typically do.

The tool will be a reference going forward for research projects

The USAID evaluation report template and guidance provides a good sense of what should be in a good evaluation report. I mostly review evaluation reports and I am not sure I would use this tool to assess quality - I think it is too focused on the technical aspects of statistical analysis and quality and less focused on whether these reports are in fact useful to anyone or used in any way to help in strategic decision-making.

This is a helpful way to review studies against USAID criteria and best practices.

This seems very specific to USAID's type of projects, and not all the questions are relevant for all studies.

This survey tool is much more different than other evaluation committees. Once I see more of this survey tool, I'm quite sure I will be able to utilize this tool in my work

This tool will be helpful in explaining quality considerations for studies to my colleagues at USAID, to guiding implementing partners in what we are looking for when we ask for quality studies, and to evaluate studies we fund or use for strategy and design.

This tool would be useful to provide to evaluators to make sure they are considering each of the characteristics captured by the tool.

Though I may not use the tool as it is, I will include some elements to consider while evaluating reports or PhD theses

Too much of a stand-alone survey. I think it'd be more valuable and meaningful if it can be integrated into the report being reviewed rather than a stand-alone survey.

Tool is helpful for evaluation, but likeliness of usage will depend on type of evaluated studies and reports because not all segments of the tool are adequate.

Useful for other contexts.

Will use it in my consultancy work and also in lectures with my students

On a scale of 1-5, how likely are you to recommend the tool to your colleagues? (With 1 being highly unlikely, 5 being extremely likely) (n=109)

AVERAGE SCORE	1 (HIGHLY UNLIKELY)	2 (SLIGHTLY UNLIKELY)	3 (MODERATELY LIKELY)	4 (LIKELY)	5 (EXTREMELY LIKELY)
4.1 (Likely)	3.7% n=4	7.3% n=8	15.6% n=17	26.0% n=28	48.0% n=52

COMMENTS

as a reminder on what to include in a study

As said, tool is helpful and can be recommended but likeliness for recommendation will depend on needs and purposes.

At this time, no one until I know how to instruct this survey tool.

I can recommend it to Faculty members as at tool for reviewing research articles and as a checklist

I did not initially consider it but will share it with my colleagues.

I do like this tool and find it help as an organizing or checklist type of tool.

I do not usually use these types of surveys in my work.

I have already shared the tool amongst several networks of professional colleagues

I have already shared it with colleagues

I think this could be a very helpful tool to share with the M&E Practice in PPL as well as with M&E POCs across sectors.

I will certainly introduce this too for reviewing our research papers in my Faculty

I will contribute to discussions on report or thesis evaluations

I will recommend it because of its easy use.

I will share the tool with colleagues who need guidance to assess their research

I will share with organization teams as a key reference for MERL in project design

I would definitely recommend this tool to colleagues, particularly in PPL and M&E POCs.

I would not because it is too cumbersome and technical. Many of my colleagues working in the programs would be overwhelmed.

If a colleague is looking for a guide on how to write up a study, it can be helpful.

I'm currently in advocacy, so this is too technical for that audience, but in research and academia, I would recommend it.

I'm going to summarize areas of potential deficiency with our reports based on what is required by USAID.

I'm sharing with my team now to use in assessing our own evaluations.

It's a great tool for evaluating our scientific research document

My colleagues will find this tool useful when designing proposals for USAID as well as other funding organizations.

probably not, but they don't use tools like this, that is why.

The tool systematizes basic elements of applied research and evaluation in an easy-to-use manner.

This is a tool for informing Research, Evidence, Accountability, Monitoring, and Learning (REALM) design elements of an organization's intervention.

Yes

ANNEX C: CIES FEEDBACK

Feedback and suggestions for Donors:

Topic	Suggestion
Why are standards of quality not met?	Lack of awareness on existing guidance, templates, trainings resources, and related expectations/recommendations
	lack of clear guidance on how to conduct a quality evaluation
	lack of capacity (not every project has a statistician etc.)
	lack of knowledge of the checklist requirements
	knowledge of awareness of importance
	not having a full appreciation of the standards and having structures to support them
	USAID wants reports to be accessible
	USAID expectations aren't clear
	perceived expectations from donors
	conflicting / potentially conflicting donor expectations / preferences
	lack of understanding on evidence for report (appropriate level of technical detail)
	lack of time, capacity, bandwidth, staffing resources to prioritize in-depth reporting on topics
	time to pay attention to these standards
	incentives from bosses
	incentives from USAID
	structural issues impede localization
What can donors do?	Provide more resources (time and money
	Change report timelines based on evidence of how long these take
	Provide more time for proper analysis and report writing
	longer timelines and budgets to support staff resources
	provide examples
	templates / examples
	highlight good examples
	think of impact evaluation from the design phase of the intervention not when it's already in the picture
	build networks with people who do this already
	increase openness to alternatives to reports that are simpler and more accessible
	communicate standards clearly from the start of the research / eval
	articulate requirements, expectations, suggestions, and options for consideration
	communicate / disseminate the tool widely
	more dissemination and socialization activities for existing resources including hands-on training, office hours
	provide trainings (webinars etc.)
	be clearer / provide trainings
	impose requirements from RFP stage, with tools

|	make these checklists and report templates a requirement for research submission

Feedback and suggestions for Researchers:

Topic	Suggestion
Why are standards of quality not met?	cost
	financial constraints
	laziness
	access
	time
	project timelines
	timelines
	tight timelines
	variation in the study purpose
	lack of awareness
	misconceptions
	lack of awareness / capacity
	lack of experience
	unclear reporting guidelines
	competing requests from nuanced concept, rigor, but simplicity
What can researchers do?	help each other -- communities of practice etc.
	look at good quality examples
	learn about what makes for a high-quality study
	training / capacity strengthening
	Use it! Incorporate into templates/planning
	encourage budgets that enable quality
	assume USAID wants the same rigor as a journal, then negotiate
	we need clear guidelines
	GUIDELINES

ANNEX D: FINDINGS WORKSHOP FEEDBACK

Instructions to Participants: Please add your feedback to any of the three questions below. Please be respectful and do not delete any comments from your colleagues; however, we hope that you will be inspired by your colleagues' feedback and expand upon it.

1. **Reflect upon the review process**. We would like to know how it went for you. What was challenging? What worked well? You may consider any part of the process, including communication and support, the orientation materials and listening sessions, the report assignment, the timeframe, the tools used, etc.

   - To improve the tool, I would include annotated reports in the final version. These annotated reports would include outstanding, good, regular, and bad reports. See here for an example, taken from the American Psychological Association
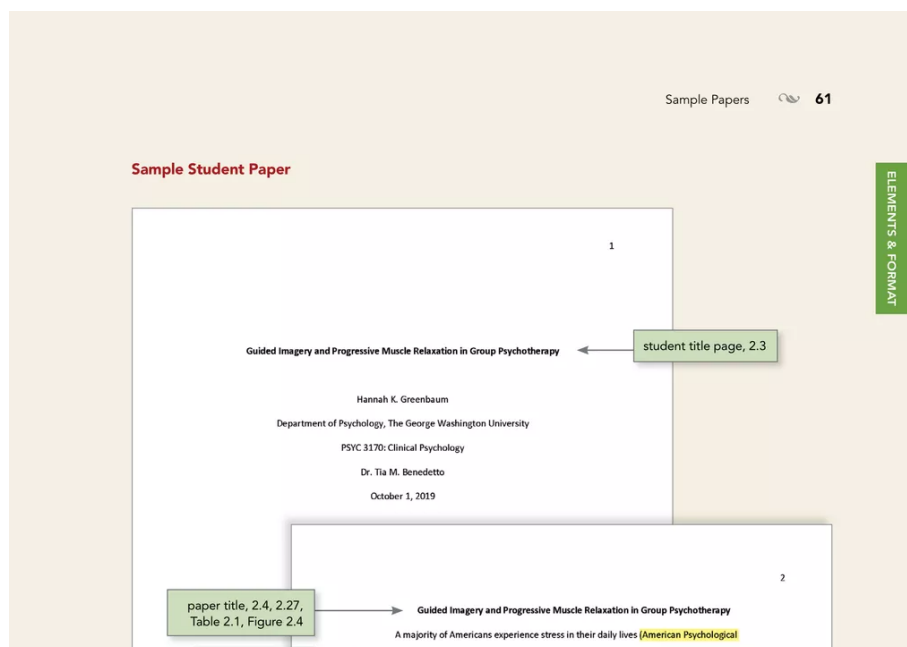
   - The glossary could be better organized by providing examples of items that fall within these categories. This of course could create some biases by steering individuals to think in a certain way.
   - A standard of evidence would need to be included, similar to the What Works Clearinghouse, Department of Education
   - Was unfamiliar with some of the concepts so challenging to review study while learning about concepts.
   - For studies with attribution of causality, the tool may be insufficient.
   - I reviewed one qualitative study. Not clear if the writers had had guidance on what to include in the report. I did a lot of scrolling through the (long) report to find pieces of the tool included in different parts of the study (time consuming). Not clear what the purpose of the tool—why are the reports being given a score? Should there be any point added to the tool regarding useability/ accessibility/ clarity of the evaluation (for example, meets the needs of different

audiences—such as those accessing the DEC; USAID education officers; other donors; implementers)?

- Shorten the tool (without compromising the quality :))

- The tool appears to be for specialized public and obviously cannot be easily understood by every other audience interested in education research. (1 like)

- The audience of the particular product matters, and audience would influence what elements a writer does and does not include. Should we add "audience" as a category? Should we figure out some way that the audience influences what other elements you look for?

- It is challenging and time consuming to fully run through the tool on a lengthy report (2 likes)

- The materials shared were excellent

- We need to use it more before deciding what to change--with more use, we might be able to refine it even more.

- It will be helpful to share the tool with projects to ensure they follow criteria when they develop the report (1 like)

- I personally have never used the tools. I would suggest that the USAID raise the awareness of stakeholders, especially researchers and academicians, on the tools and their use. Probably through trainings.

- The tool itself if complex.

- I learned of this program a bit later than expected and wished I had known earlier so that I could have gotten the full training.

- I found the orientation material provided with the tool to be very helpful.

- Articulate the purpose of the tool--to improve usability to those accessing it in the future (such as on the DEC); eligibility for future funding; a requirement of the USAID program closeout…

- Aimed for use by only a specific user audience (with a background in research/evaluation design). Those without expertise in research in Education would find it difficult to apply this tool.

- The reports were not organized to meet the elements listed on the tool (1 like)

- Next review – two reviewers per report to see how different people read the same report. Likely have variability. Maybe not every item, just selected items for double review.

- The presenters said that for analysis, they identified "critical questions" which if the answer was not Yes, then the study couldn't meet adequate quality. Could USAID be clearer how "critical questions" are important and why the tool could not be shortened/simplified with just the critical questions.

- Test the inter-rater reliability of the tool.

- The cut off of 66% for quality seems arbitrary.

2. **Think about the future of education research and evaluation and the Assessment of Study Quality (ASQ) Tool**. How can USAID support improving the quality of future studies? How can we support you to use the tool? What strategies should USAID take to encourage the use of the tool?

- link tool to USAID's evaluation report template

- Perhaps AOR/COR and TOCOR should receive additional training on evaluation methodologies, so they can order better studies and profit more from reports.

- Encourage Missions to share it with IPs engaging in studies.
- It would be possible to coordinate with Missions what is desirable to include in a report?
- Be clear how this tool compares to other existing tools. (1 like) + Could USAID provide use cases to clarify when to use this tool, as compared with alternative quality assessment tools, such as are linked on USAID's Learning Lab https://usaidlearninglab.org/resources/community-led-development-cld-tools.
- Share the tool more broadly - for example with graduate school programs to have their students try out.
- Video on YouTube, to help people understand the tool (1 like)
- Interactive Infographic or short vids (1 like)
- Audience for use and applicability is going to be very important. Different audiences will have different interests in the report. As mentioned, what USAID expects in terms of quality, vs stakeholders who want to understand how to apply the findings more directly, for example. (1 like)
- Show the value of using the tool - give researchers/ implementers incentive to use it (and produce high quality reports)
- Based on the findings, focus on guidance for writing on the weakest areas.
- Trainings
- Train the stakeholders on the research's process and integrated it in the local universities in different countries.
- Identify more interested groups through specialized platforms like education research platforms and ensure widespread webinars that expose these groups to the tools. More like orientation sessions.
- If programs are meant to create studies that align with this tool, provide access/ guidance to ethics training and possibly other types of intensive mini research trainings.
- USAID can promote the tool by providing more capacity building (Training) on the use of the tool.
- I agree with what most have commented about providing more training opportunities like workshops and professional development. The tool is a great checklist to unify and systemize best practices in research.
- Share and training to researchers.
- USAID can train projects team on the tool and elements to fulfill and provide guidelines to ensure that they follow.
- Could be useful as a guideline for writing research, and to understand the design features that USAID is looking for in their reports.
- Consider having a one-on-one/ small group training or mentoring for organizations that would be "required" to use the tool (it's complex for those will less technical background)
- Require it be used for all USAID-funded research and evaluation.
- Possible to frame as guidance/guidelines ahead of study so that by time in writing they've already designed in alignment with USAID expectations what they want to see in education evaluations. Need to be clear on audience (researchers/experienced). (2 likes)
- Provide training on how to use the tool.

- Additional training in use of the tool, socializing the tool and its components with those working on research studies and evaluations so studies/reports can be planned to respond to necessary requirements.
- Disseminate the tools in advanced and treat it as a kind of learning tools or guidelines rather than solely for assessment or review criteria.
- Provide the research team to use the tools and learn collect data.
- Train the stakeholders on the research's process and integrated it in the local universities in different countries.
- Coordinate with missions on reporting needs: tension with reporting everything in tool vs simplified documents for the missions (annexes for technical items) – tension between user types
  - Tool in light of how missions use it and how other stakeholders use it – tool for technical report, then also produce brief. Standard errors are not included in the briefs (multiple products out of each study) – have a paper trail (know this exists somewhere)
- Report format for usability different than classical evaluation approach. End users are not necessarily trained to be familiar with academic reports. In terms of communication and dissemination, that could be the thing.
- Dissemination: share tool with projects, provide orientation sessions to the tool to understand it better, provide examples because not everything is included in the report and not everyone who develops report is familiar with requirements to include different elements. Maybe YouTube video

3. **Reflect upon the findings of the review about the quality of research and evaluations that were presented**. How do these findings resonate with you? What do you think they mean? Do you interpret any of the findings differently than they were presented?

- Yes, the findings kind of make sense. These findings resonate with past experience.

- It wasn't clear if the writers of the report I reviewed had received guidance on report contents.

- I am glad that cultural Appropriateness is a factor in this survey. It is a critical addition as most researchers come from the global north studying the global south and at times may not be as informative of the local context as they should be. (1 like)

- It's hard to determine reliably an entire study quality from what's written in the report, vs what may have actually been done or considered but not written about. (1 like)

- I was surprised to see the ethics category rated so low. I could see that perhaps this was just not mentioned in reports, but I had thought ethical review was standard. (1 like)

- Yes, the cultural appropriateness and ethics findings being weak resonates because this is an area we need to improve in general.

- The report I reviewed had components spread throughout the report. It was difficult to rate. Not sure if others had similar situations-->maybe some of the ratings (such as mine) were a little "off" (1 like)

- Yes, it did.

- Yes, especially on the Ethics and Cultural appropriateness. I think sometimes it's not that researchers don't consider ethical aspects, but it's because of the rarity of ethical applications or guidelines in their country context to orient their practice. For cultural appropriateness, there should be genuine intentions from researchers to understand and take into account in their research rather than just verbal manifestation.

- Evaluators are pressed by deadlines – do not have time for procedure such as ethics. This is very time consuming and may take weeks or months to get proper authorization. Timelines do not allow for this – this may be part of the ethics weakness. Proper approach to consent/assent is not the same across contexts.

  - Time constraints are not appropriate reasons to not adhere to regulations/standards on protection of subjects in US funded research. That said, this tool could reinforce some consistency in USAID's expectations for ethics procedures and approval bodies.

- PMs at missions are interested in looking good. PM has authority to order evaluation and is also responsible for project success – this is complicated. Maybe this should be part of recs – clean cut separation between all evals and all management of projects (COIs). Most people do not make distinction between types of evaluations.

- Missing data/nonresponse item, we have not thought about this during design or when writing reports. This is interesting.

4. **Please share any additional thoughts**.

- You Need to train faculty members in different countries on USAID research to make capacity building and applied sustainability strategy for USAID assessment tools in Academic faculty.

- Train the research team on the planning and collect data because I faced challenge in Amman and Egypt when we applied EGRA with accessories researchers who collect data from the field, so you need to establish local research centers refer to USAID to use these tools.

- Develop research priorities based on Countries and USAID needs in all development areas and establish different research design for each kind of researchers to publish research results with decision makers.

- You could also provide some samples for guidance supported with infographic that explains/shows how the criteria was fulfilled in the report.

- Does USAID have expectations on when evaluations meet the criteria for "research involving human subjects? Important because the federal Common Rule applies to research but not evaluations. Does USAID expect evaluations to adhere to the same standards of conduct, and therefore required IRB processes, as research, even if they are not "research" by federal definition?