



OPEN LETTER

Challenges and opportunities in evaluating programmes incorporating human-centred design: lessons learnt from the evaluation of Adolescents 360 [version 1; peer review: 2 approved, 1 approved with reservations]

Aoife M. Doyle ¹, Emma Mulhern², James Rosen³, Gabrielle Appleford ², Christina Atchison¹, Christian Bottomley¹, James R. Hargreaves¹, Michelle Weinberger ³

¹London School of Hygiene & Tropical Medicine, London, WC1E7HT, UK

²Itad Ltd, Hove, BN31RE, UK

³Avenir Health, Washington, DC 20012, USA

v1 First published: 24 May 2019, 3:1472
<https://doi.org/10.12688/gatesopenres.12998.1>

Latest published: 25 Sep 2019, 3:1472
<https://doi.org/10.12688/gatesopenres.12998.2>

Abstract

Adolescents 360 (A360) is a four-year initiative (2016–2020) to increase 15–19-year-old girls' use of modern contraception in Nigeria, Ethiopia and Tanzania. The innovative A360 approach is led by human-centred design (HCD), combined with social marketing, developmental neuroscience, public health, sociocultural anthropology and youth engagement 'lenses', and aims to create context-specific, youth-driven solutions that respond to the needs of adolescent girls. The A360 external evaluation includes a process evaluation, quasi-experimental outcome evaluation, and a cost-effectiveness study. We reflect on evaluation opportunities and challenges associated with measuring the application and impact of this novel HCD-led design approach. For the process evaluation, participant observations were key to capturing the depth of the fast-paced, highly-iterative HCD process, and to understand decision-making within the design process. The evaluation team had to be flexible and align closely with the work plan of the implementers. The HCD process meant that key information such as intervention components, settings, and eligible populations were unclear and changed over outcome evaluation and cost-effectiveness protocol development. This resulted in a more time-consuming and resource-intensive study design process. As much time and resources went into the creation of a new design approach, separating one-off "creation" costs versus those costs associated with actually implementing the programme was challenging. Opportunities included the potential to inform programmatic decision-making in real-time to ensure that interventions adequately met the

Open Peer Review

Reviewer Status

	Invited Reviewers		
	1	2	3
version 2 (revision) 25 Sep 2019		 report	
version 1 24 May 2019	 report	 report	 report

1. **Isaac Holeman** , Medic Mobile, Seattle, USA
University of Washington, Seattle, USA
2. **Melissa Gilliam**, University of Chicago, Chicago, USA
3. **Kristien Michielsen**, Ghent University, Ghent, Belgium

Any reports and responses or comments on the article can be found at the end of the article.

contextualized needs in targeted areas.

Robust evaluation of interventions designed using HCD, a promising and increasingly popular approach, is warranted yet challenging. Future HCD-based initiatives should consider a phased evaluation, focusing initially on programme theory refinement and process evaluation, and then, when the intervention program details are clearer, following with outcome evaluation and cost-effectiveness analysis. A phased approach would delay the availability of evaluation findings but would allow for a more appropriate and tailored evaluation design.

Keywords

Human-Centered Design, User-centered Design, Design Thinking, Evaluation, Adolescents, Contraception, Africa

This article is included in the [International](#)



[Conference on Family Planning](#) gateway.

Corresponding author: Aoife M. Doyle (aoife.doyle@lshtm.ac.uk)

Author roles: **Doyle AM:** Conceptualization, Investigation, Methodology, Writing – Original Draft Preparation; **Mulhern E:** Conceptualization, Investigation, Methodology, Writing – Review & Editing; **Rosen J:** Conceptualization, Investigation, Methodology, Writing – Review & Editing; **Appleford G:** Conceptualization, Investigation, Methodology, Writing – Review & Editing; **Atchison C:** Conceptualization, Investigation, Methodology, Writing – Review & Editing; **Bottomley C:** Conceptualization, Methodology, Writing – Review & Editing; **Hargreaves JR:** Conceptualization, Methodology, Writing – Review & Editing; **Weinberger M:** Conceptualization, Investigation, Methodology, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This study was supported by the Bill & Melinda Gates Foundation (OPP1134172) and the Children's Investment Fund Foundation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Doyle AM *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Doyle AM, Mulhern E, Rosen J *et al.* **Challenges and opportunities in evaluating programmes incorporating human-centred design: lessons learnt from the evaluation of Adolescents 360 [version 1; peer review: 2 approved, 1 approved with reservations]** Gates Open Research 2019, **3**:1472 <https://doi.org/10.12688/gatesopenres.12998.1>

First published: 24 May 2019, **3**:1472 <https://doi.org/10.12688/gatesopenres.12998.1>

Rationale

In the field of public health, human-centred design (HCD), also known as design thinking, is increasingly being used during intervention development, and sometimes implementation, to design innovative solutions to complex problems¹. HCD is a creative approach to programme design that focuses on building empathy with the population of interest, and generating and iteratively testing many potential solutions to complex problems. The approach has a high tolerance for both ambiguity and failure¹⁻³. HCD shares some characteristics with other methods used to design programmes such as traditional socio-behavioural research but tends to be more ‘nimble/flexible/iterative’ and less protocol driven⁴.

Intervention evaluation is an essential component of public health research and programming, and detailed guidance on evaluation approaches and methodologies exist^{5,6}. However, the iterative and flexible nature of HCD potentially poses some unique challenges for evaluation^{1,3,7}, and examples of evaluations of HCD public health interventions have so far been limited^{7,8}. There are some parallels between evaluations of HCD interventions and evaluations of interventions incorporating adaptive management or quality improvement where the intervention can also change over time⁹⁻¹⁴. Also, programmatic interventions, as opposed to researcher-led interventions, often involve an element of flexibility in intervention implementation, as implementers strive to develop more context appropriate interventions, and respond to programmatic targets. In such interventions, the location and timing of intervention implementation is guided primarily by programme targets or local government priorities, and less so by a need to maximise the quality of evidence of impact^{15,16}. Despite some overlap with other evaluation scenarios, we have identified some unique challenges to the evaluation of HCD-led interventions.

Purpose of the letter

In this letter we discuss the evaluation opportunities and challenges identified during the ongoing A360 evaluation and make recommendations for future evaluations of programmes designed using HCD.

Adolescents 360

The intervention that is under evaluation is Adolescents 360 (A360), a multi-country initiative which aims to increase the voluntary use of modern contraceptives among adolescent girls aged 15–19 years. The A360 initiative is led by Population Services International (PSI) in collaboration with the Society for Family Health (SFH) in Nigeria, design partner IDEO.org, and the Centre for the Developing Adolescent at the University of California. A360 is co-funded by the Bill & Melinda Gates Foundation and the Children’s Investment Fund Foundation.

A360 set out to take a transdisciplinary approach to intervention development which involves the different discipline ‘lenses’ working jointly to create new innovations that integrate and move beyond discipline-specific approaches to address a common problem (Figure 1).

A360 aims to create context-specific, youth-driven solutions that respond to the needs of young people. Through an iterative



Figure 1. The A360 approach.

design process, unique solutions were created in each of the A360 countries. In Ethiopia, Smart Start uses financial planning as an entry point to discuss contraception with newly married couples¹⁷. In southern Nigeria, 9JA girls provides physical safe spaces (branded spaces in public health clinics) for unmarried girls¹⁸, and in northern Nigeria, Matasa Matan Arewa targets married girls and their husbands with counselling and contraceptive services using maternal and child health as an entry point¹⁹. In Tanzania ‘Kuwa Mjanja’ (be smart) delivers life and entrepreneurial skills training alongside opt-out counselling sessions and on-site contraceptive service provision to both married and unmarried adolescent girls²⁰.

External evaluation

The independent external evaluation aims to understand whether the A360 approach leads to improved reproductive health outcomes; generate evidence about A360’s cost-effectiveness in relation to other programming approaches; and capture how A360 is implemented in different contexts, and the experience of the participating young people and communities. It has four components (Figure 2)²¹:

1. Outcome evaluation which has a quasi-experimental design with before-after cross-sectional surveys in all four settings (Oromia, Ethiopia; Mwanza, Tanzania; Ogun and Nasarawa, Nigeria) and a comparison group in the two Nigerian settings²².
2. Process evaluation incorporating both traditional qualitative methodologies and youth participatory research²³.
3. Costing study to understand the drivers of cost and a more extensive cost-effectiveness study within the outcome evaluation study geographies.
4. Engagement and research uptake.

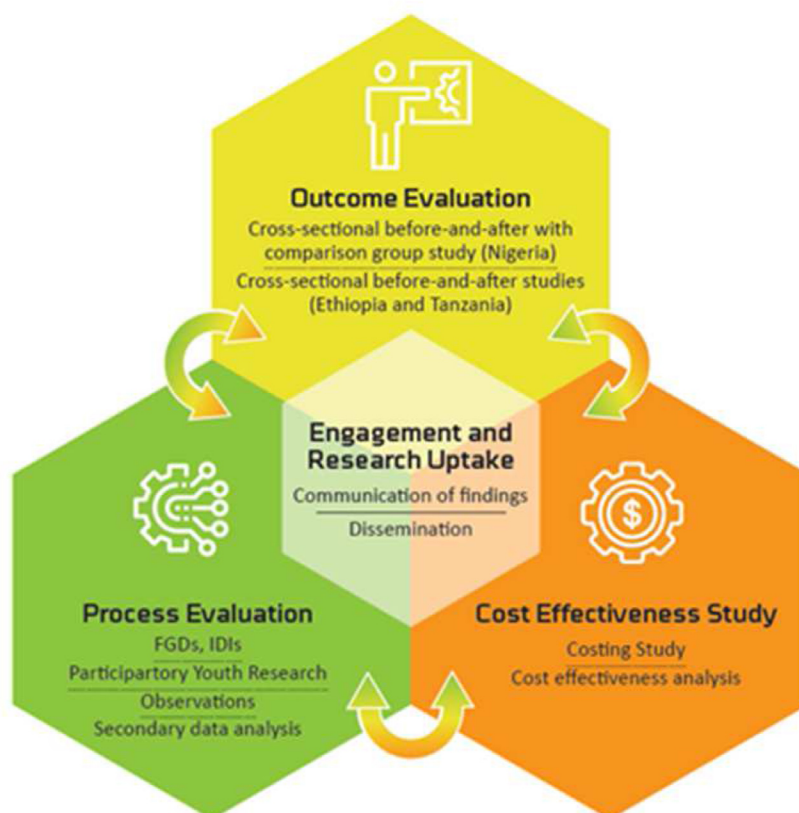


Figure 2. A360 multi-component external evaluation.

HCD process and timeline

HCD takes a phased approach to intervention development. In the ‘inspiration phase’ the designers immerse themselves with girls and her influencers to understand the issues. In the ideation phase they attempt to make sense of what has been learnt, identify opportunities for design and conduct an iterative cycle of prototyping of possible solutions and strategies. The most promising prototypes are then piloted. In the case of A360, implementation at scale involved a further ‘optimisation’ period where the selected solutions were further modified to maximise scalability and affordability. In A360, the inspiration and ideation phases took place over approximately 12 months (Sept 16-Aug 17) with piloting for a further 3-4 months (September-December 17). Optimisation and scale-up started towards the end of 2017/early 2018. Among other things, this lengthy intervention design period presented some challenges for the evaluation, which we explore in the sections that follow.

Outcome evaluation

The outcome evaluation aimed to measure the impact of A360 on the voluntary use of modern contraceptives among adolescent girls aged 15–19 years in the study geographies. The outcome evaluation was designed when the implementers were at the ‘inspiration phase’ with final study protocols submitted for ethical approval when prototyping was still ongoing and before the intervention had been finalised (Figure 3). This outcome evaluation

timeline was necessary in order to conduct baseline surveys in the four study settings in 2017 prior to scale-up of the finalised interventions. The implications of the concurrent intervention development and evaluation design were that key pieces of information about the intervention components and the theory of change were unclear and changed over the protocol development period. Detailed information on the intervention were needed to inform the setting for baseline surveys and the kinds of questions to be asked during the surveys. The OE study design was finalised before the interventions themselves so there remained a risk that the evaluation was not optimal.

A key consideration for the baseline survey design was who would be targeted with the A360 intervention in terms of geographical location and demographic characteristics. Early on in the intervention design process it became clear that A360 would be targeted towards only married or unmarried adolescent girls in some locations though uncertainty remained through early phases of the design process. For example, in Northern Nigeria, the outcome evaluation focused on married adolescent girls as the initial intention of A360 was to target only those who were married, however, 9ja Girls, the intervention that was designed in Southern Nigeria to target unmarried adolescent girls, was eventually also implemented in Northern Nigeria. In Ethiopia, initial HCD insights pointed towards a focus on unmarried adolescents in school but other programming pressures including donor

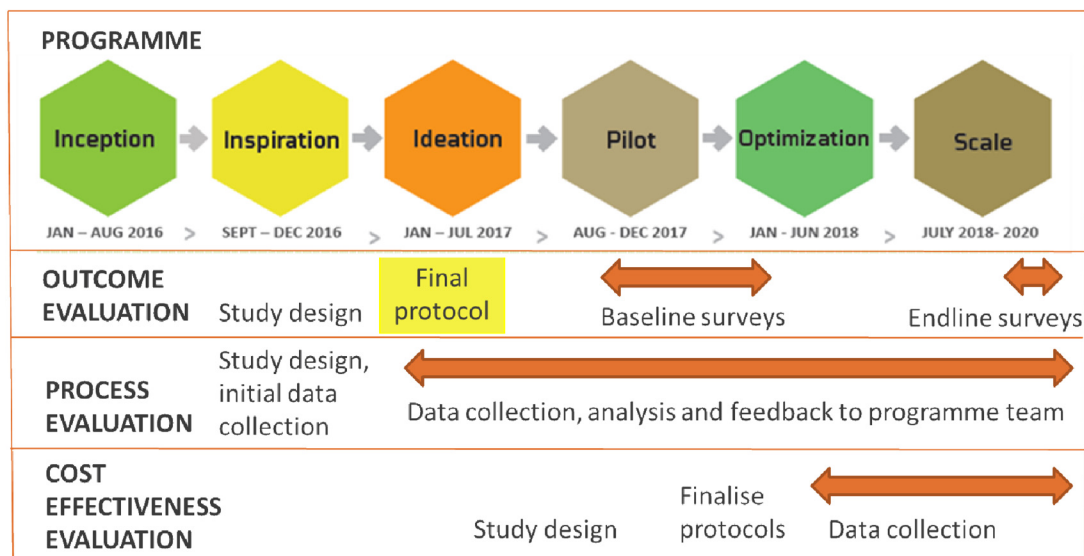


Figure 3. Timeline of A360 intervention development and implementation and the external evaluation.

preferences led to a later shift in focus to married rural women. The outcome evaluation was designed to target all married 15-19-year-old girls, but the final intervention focused on ‘newly married’ adolescent girls. While some change in target population might be expected with non-HCD programmatic interventions, the HCD process provided the implementers with more freedom and confidence to make additional changes. It was challenging to determine the geographies that would receive A360 as the implementers were in the middle of an intense phase of solution prototyping and were not yet thinking about larger scale implementation of A360.

The level of implementation of the intervention was also relevant for study design, e.g. would A360 be implemented at health facilities, in schools or in the community? Given the uncertainty as to how A360 would be implemented, we opted for a community-based survey, which turned out to be appropriate for the final A360 interventions. However, if A360 had been implemented primarily in health facilities or schools then evaluating the intervention at the broader community level might not have been the most efficient study design.

The uncertainties in key study design parameters meant that we had to develop multiple study design scenarios which were repeatedly revised as new information came to light. For example, initially the intervention was to have distinct implementation phases as it was rolled across Nigeria and we explored the idea of conducting a stepped-wedge trial. Following further discussions with the implementers, it became clear that there was no certainty that roll-out would be in a phased manner as they did not know yet what the intervention would comprise. In Tanzania, we explored whether a regression discontinuity design might be possible if the intervention were implemented in one or several of the existing demographic surveillance sites, however, given the uncertainty around the nature of the

intervention, the implementers were reluctant to commit to implementing in those areas. This prolonged study design process entailed some added costs and was at times frustrating for all involved.

Process evaluation

The process evaluation was aligned to the HCD-driven phases of the design process for A360 and had the primary objective of presenting a descriptive and analytical account of how the implementation of A360 has played out in relation to its Theory of Change²³. The process evaluation attempted to understand both the intervention development process and intervention implementation. During intervention development the process evaluation team faced challenges as the fast-paced, highly iterative HCD process meant that the decision making process, the ‘design energy’, often went undocumented⁴. This challenge was also noted in another evaluation of a HCD design process⁸. As a result, the process evaluation team needed to be flexible in order to align closely with the work plan of the implementers and methodologies such as direct participant observation were key to capturing the depth of the HCD process.

The potential for research fatigue was observed among target community members as their views were solicited by both the implementers designing the intervention and the process evaluation team interested in understanding the HCD process from the participants’ point of view. The process evaluation team, therefore, needed to balance the importance of capturing the views of community members with the potential for research fatigue.

During the design of the process evaluation, the intention was that the findings would feed into and inform the intervention design at key moments. However, there was limited uptake of process evaluation findings by implementers. For example, the process

evaluation highlighted the need for the programme to do more to address broader community and social norms, but this finding had a limited impact on intervention design. Poor uptake of findings was partly a result of the fast-paced nature of A360 and the resultant demands on the implementers' time which did not allow them sufficient time to pause and reflect on the process evaluation findings. Country implementing teams differed in how they engaged with external recommendations, with some teams receptive and willing to listen and adjust, while others were more protective of their 'solutions'. Designers add value to public health as they take a different approach when addressing an issue, however, in the case of A360, a preference to work from a blank slate free of the constraints of public health evidence may have impacted on an ability to respond to the PE findings. Uptake of the process evaluation findings improved when the evaluation team introduced participatory action research activities which focused on questions that were important for the country teams (e.g. health care provider attitudes in Ethiopia²⁴ and Nigeria²⁵).

Cost-effectiveness evaluation

The cost-effectiveness study faced similar challenges as the outcome evaluation in terms of complicated and delayed development of the cost-effectiveness study protocols. Although all costing exercises face some unknowns about the intervention to be costed, the HCD process added a layer of uncertainty. Furthermore, because proponents hypothesized that it was the design process itself that would be the key factor in producing an effective (and cost-effective) intervention, an important challenge was both measuring the total design cost and isolating the cost of HCD specifically as HCD activities of the design process were tightly intertwined with the other A360 'lenses'. In addition, because A360 was creating a new trans-discipline approach, there was also interest in separating out the costs associated with the 'creation' of the A360 approach from the costs associated with implementing and scaling A360 in countries. Interviews were held with intervention staff in order to get a sense of the distribution of activities (and associated costs) across the study 'lenses' and between efforts to create versus implement A360. The implementing agencies planned continuous changes or tweaks to the intervention once it was up and running and so more frequent cost data draws were required during the scale-up period to capture how those change might affect the "production process" and associated costs.

Lessons learnt

Robust evaluation of a new and promising approach such as HCD is warranted yet challenging. Some of the challenges faced are not unique to interventions designed with HCD and will be recognisable by those who have led evaluations of programmatic interventions. In comparison to research-led studies, evaluation of programmatic interventions is associated with a reduced level of control and increased uncertainty. A challenge for evaluators is to find ways to deal with this uncertainty while still retaining scientific rigor.

We have the following recommendations for others who would like to evaluate programmes incorporating HCD:

1. Implementers should allow adequate time to participate in the process evaluation, as well as work with the process evaluation team to ensure that findings are timed to feed into key decisions.
2. The process evaluation team should maximise secondary analysis of data collected by the implementers, and joint data collection could be considered where additional data collection is needed and participant research fatigue is anticipated.
3. Like HCD, an iterative and adaptive process evaluation approach is required. In A360, the process evaluation paused after the pilot phase and our team worked with the A360 implementers and donors to develop evaluation questions that reflected the solutions that were being developed, reflecting an iterated A360 Theory of Change.
4. For the process evaluation, methodologies, such as direct observations, are instrumental to capturing the fast-paced, highly-iterative HCD process, and to understand the 'design energy' i.e. how decisions were made at key points in the design process.
5. To avoid a time-consuming and resource-intensive design process, future HCD-based initiatives should consider a phased evaluation approach:
 - Conduct process evaluation during the HCD inspiration, ideation, and pilot phases.
 - Wait until the implementers have a better understanding of the emerging programme and have finalised the target geographies, target population, and intended outcomes before planning an outcome evaluation and cost-effectiveness study.
 - During the implementation phase, conduct a comprehensive process evaluation that can capture whether, how, and why the intervention changed during implementation.

The advantages of a phased approach need to be balanced against the disadvantages of lengthening of the time between implementation and the availability of evaluation findings.

Data availability

No data are associated with this article.

Grant information

This study was supported by the Bill & Melinda Gates Foundation (OPP1134172) and the Children's Investment Fund Foundation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We would like to thank Anne LaFond for helpful comments on an earlier draft of this letter.

References

1. Bazzano AN, Martin J, Hicks E, *et al.*: **Human-centred design in global health: A scoping review of applications and contexts.** *PLoS One*. 2017; 12(11): e0186744.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
2. Kolko J: **Design Thinking Comes of Age.** *Harvard Business Review*. 2015.
[Reference Source](#)
3. IDEO.org: **The Field Guide to Human-centered design.** IDEO.org/Design Kit; 2015.
[Reference Source](#)
4. Tolley EB: **Traditional Socio-behavioral research and human-centered design: Similarities, Unique Contributions and Synergies.** 2017.
[Reference Source](#)
5. **Evaluation in health and wellbeing.** 2018; (Accessed 08 Mar 19, 2019).
[Reference Source](#)
6. Webster J, Exley J, Copestake J, *et al.*: **Timely evaluation in international development.** *J Dev Effect*. 2018; 10(4): 482–508.
[Publisher Full Text](#)
7. LaFond A, Davis N: **Evaluating the Influence of Human-Centered Design on Health Programs.** The Pump; 2017.
[Reference Source](#)
8. Itad: **Evaluation of the Hewlett Foundation's Strategy to Apply Human-Centered Design to improve Family Planning and Reproductive Health Services in Sub-Saharan Africa.** Final Report. 2017.
[Reference Source](#)
9. Rogers P: **Does evaluation need to be done differently to support adaptive management?** *BetterEvaluation*; 2017.
[Reference Source](#)
10. Simister N: **How Adaptive Management is challenging the monitoring and evaluation of complex programmes.** intrac for civil society; 2018.
[Reference Source](#)
11. Valters C, Cummings C, Nixon H: **Putting learning at the centre. Adaptive development programming in practice.** London, UK: ODI; 2016.
[Reference Source](#)
12. Barry D, Kimble LE, Nambiar B, *et al.*: **A framework for learning about improvement: embedded implementation and evaluation design to optimize learning.** *Int J Qual Health Care*. 2018; 30(suppl_1): 10–4.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Øvretveit J, Gustafson D: **Evaluation of quality improvement programmes.** *Qual Saf Health Care*. 2002; 11(3): 270–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Waiswa P, Manzi F, Mbaruku G, *et al.*: **Effects of the EQUIP quasi-experimental study testing a collaborative quality improvement approach for maternal and newborn health care in Tanzania and Uganda.** *Implement Sci*. 2017; 12(1): 89.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Birdthistle I, Schaffnit SB, Kwaro D, *et al.*: **Evaluating the impact of the DREAMS partnership to reduce HIV incidence among adolescent girls and young women in four settings: a study protocol.** *BMC Public Health*. 2018; 18(1): 912.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Chandrasekaran P, Dallabetta G, Loo V, *et al.*: **Evaluation design for large-scale HIV prevention programmes: the case of Avahan, the India AIDS initiative.** *AIDS*. 2008; 22 Suppl 5: S1–15.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. A360: **Country Solutions- Ethiopia.** 2019; (Accessed 28 Mar 19).
[Reference Source](#)
18. A360: **Country Solutions- Southern Nigeria.** 2019; (Accessed 28 Mar 19).
[Reference Source](#)
19. A360: **Country Solutions- Northern Nigeria.** 2019; (Accessed 28 Mar 19).
[Reference Source](#)
20. A360: **Country Solutions- Tanzania.** 2019; (Accessed 28 Mar 19).
[Reference Source](#)
21. **Adolescents 360 external evaluation.** 2019; (Accessed 28 Mar 19).
[Reference Source](#)
22. Atchison CJ, Mulhern E, Kapiga S, *et al.*: **Evaluating the impact of an intervention to increase uptake of modern contraceptives among adolescent girls (15-19 years) in Nigeria, Ethiopia and Tanzania: the Adolescents 360 quasi-experimental study protocol.** *BMJ Open*. 2018; 8(5): e021834.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Mulhern E, Appleford G: **Adolescents 360 Evaluation: Process evaluation methodology.** Itad; 2018.
[Reference Source](#)
24. Itad: **Adolescents 360 Evaluation: How might we better meet the needs of adolescent couples with contraceptive counseling and services through Ethiopia's Health Extension Program?** 2018.
[Reference Source](#)
25. Itad: **Adolescents 360 Evaluation: What do service providers think about contraceptive service provision to 15–19 year old girls in Nigeria?** 2018.
[Reference Source](#)

Open Peer Review

Current Peer Review Status:   

Version 1

Reviewer Report 15 July 2019

<https://doi.org/10.21956/gatesopenres.14106.r27351>

© 2019 Michielsen K. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Kristien Michielsen

International Centre for Reproductive Health (ICRH), Ghent University, Ghent, Belgium

The article presents lessons learnt from an external evaluation of the A360 programme in three African countries. A360 uses the human-centred design and hence applied three types of evaluations to evaluate it: whether the programme leads to improved reproductive health outcomes (outcome evaluation), capture how it is implemented (process evaluation) and study its cost-effectiveness. Additionally engagement and research uptake.

The article honestly presents the known tension between effective implementation of a context-specific and flexible interventions versus an experimental evaluation design. It is refreshing to read an honest reflection on this, and as someone who has done similar evaluations, it clearly recognize the challenges. However, it would be even better if more specific recommendations would have been given on how to actually deal with this tension.

Limited uptake of the process evaluation results by the implementers: it would have been interesting to understand why some implementers were eager to take on recommendations and others were not, and how this affected the effectiveness.

Cost-effectiveness: since the effectiveness study was somewhat compromised, it is difficult to link it to the costs. It feels as if the cost-effectiveness study mainly mapped the costs, without linking it to the effectiveness.

Reflections:

- The lessons learnt are important, but stem from common sense and are not new. While it is often clear what should be done, the challenge remains on how to do it: how to convince donors and implementers of the importance of proper evaluations and its consequences for the programme design and implementation.
- Cost-effectiveness of the evaluation: a lot of time and energy was given to do three types of evaluations. However, it seems as if there value was not optimal. It may be interesting to reflect on the cost-effectiveness of the evaluation itself.

- How were the different evaluations tied together? They seem to be presented quite separately, while for a clear understanding of the programme, they should be tied together.
- How is the evaluation approach chosen better/worse than e.g. realist evaluation, which seems quite appropriate for this type of intervention/programme.

Is the rationale for the Open Letter provided in sufficient detail?

Yes

Does the article adequately reference differing views and opinions?

Partly

Are all factual statements correct, and are statements and arguments made adequately supported by citations?

Yes

Is the Open Letter written in accessible language?

Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Adolescent sexual health, monitoring and evaluation.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 12 Sep 2019

Aoife Doyle, London School of Hygiene & Tropical Medicine, London, UK

The article presents lessons learnt from an external evaluation of the A360 programme in three African countries. A360 uses the human-centred design and hence applied three types of evaluations to evaluate it: whether the programme leads to improved reproductive health outcomes (outcome evaluation), capture how it is implemented (process evaluation) and study its cost-effectiveness. Additionally engagement and research uptake.

The article honestly presents the known tension between effective implementation of a context-specific and flexible interventions versus an experimental evaluation design. It is refreshing to read an honest reflection on this, and as someone who has done similar evaluations, it clearly recognizes the challenges. However, it would be even better if more specific recommendations would have been given on how to actually deal with this tension.

Response: Thanks. We have added a new recommendation (1) and revised recommendation (4) to suggest that evaluators and implementers need to invest time upfront to become familiar with each other's methodologies and then to have more open discussions prior to the study on the potential challenges that they may face in the conduct of interdisciplinary research.

Limited uptake of the process evaluation results by the implementers: it would have been interesting to understand why some implementers were eager to take on recommendations and others were not, and how this affected the effectiveness.

Response: Yes, we agree that this is interesting. This issue is being explored through ongoing process evaluation and the findings will be made available at a later date.

Cost-effectiveness: since the effectiveness study was somewhat compromised, it is difficult to link it to the costs. It feels as if the cost-effectiveness study mainly mapped the costs, without linking it to the effectiveness.

Response: The cost-effectiveness and outcome evaluation studies are still ongoing. When the effectiveness results are available following endline surveys in 2020, the findings will be used in the cost-effectiveness analysis.

Reflections:

- The lessons learnt are important, but stem from common sense and are not new. While it is often clear what should be done, the challenge remains on how to do it: how to convince donors and implementers of the importance of proper evaluations and its consequences for the programme design and implementation.

Response: In recognition of this issue we have added the following sentence: 'An additional long-standing challenge, not specific to HCD, is to convince implementers that rigorous evaluation, while at times intrusive, will improve programme design and implementation.'

- Cost-effectiveness of the evaluation: a lot of time and energy was given to do three types of evaluations. However, it seems as if there value was not optimal. It may be interesting to reflect on the cost-effectiveness of the evaluation itself.

Response: The three evaluations are presented separately as they represent separate streams of the evaluation work but are very integrated and will be presented and interpreted together at the end of the project. This is an ongoing evaluation and so it is too early to reflect on the cost-effectiveness of the evaluation itself.

- How were the different evaluations tied together? They seem to be presented quite separately, while for a clear understanding of the programme, they should be tied together.

See previous response.

- How is the evaluation approach chosen better/worse than e.g. realist evaluation, which seems quite appropriate for this type of intervention/programme.

Response: Our evaluation approach focuses on several key process, outcome and cost measures so we are to some extent capturing the context- mechanism-outcome configurations that are described in the realist evaluation literature(1, 2). However, the outcome evaluation is powered to look at the average impact of the intervention across

the total population (Does it work?) and so we will have limited power to look at intervention impact in different sub-groups (What works for whom and where?), which is an essential part of a realist evaluation. It is possible that the iterative nature of HCD would necessitate more frequent revisits of the programme theory and CMO statements and would hence be relatively more expensive than a realist evaluation of a non-HCD programme.

1. **Pawson R, Tilley N. *Realistic Evaluation*. London: Sage; 1997.**
2. **Westhorp G. *Realistic Impact Evaluation. An Introduction.*: ODI; 2014.**

Competing Interests: No competing interests were disclosed.

Reviewer Report 15 July 2019

<https://doi.org/10.21956/gatesopenres.14106.r27411>

© 2019 Gilliam M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Melissa Gilliam

The Centre for Interdisciplinary Inquiry & Innovation in Sexual Health and Productive Health, University of Chicago, Chicago, IL, USA

In this article, the authors discuss challenges and opportunities associated with evaluating HCD interventions. In this case, the intervention was associated with A360.

Overall, the honesty of the authors is appreciated as rarely does one get to hear about problems that occurred during project implementation. Reporting on successes and failures is important as it provides an opportunity for others to learn. Nevertheless, the descriptions of the challenges that arose were a little cryptic and it was not always clear.

I was also disappointed that the authors did not reach another logical conclusion which is that there was a methodologic gap between the public health researchers/evaluators and designers which lead to the challenges. Conducting interdisciplinary research can be challenging and requires each team to be very familiar with the other's methods and processes so they can carefully design the collaboration. For example, when the study populations grow frustrated with the multiple inquiries it suggests that the data collection was not harmonized across teams. Similarly, the breakdown in communication emerges. These are the classic challenges of interdisciplinary research and may not only apply to HCD. The authors might discuss how they prepared to work with human centered designers or vice versa. What type of training took place and how familiar were the evaluators with the design process or were the designers with public health methods?

Specific comments are as follows:

1. It is interesting that the intervention design phase appears long. If one were to compare it to drug development or another multiphase intervention development process, then two years would be quite short.
2. It is not clear why baseline data were collected prior to developing the intervention. How does collecting data in the four settings, the explanation provided here, constrain when data are collected?
3. In looking at the diagram of the timeline, it seems that the baseline survey could have been moved up to July 2018 once all decisions were made regarding what the study would be about.
4. The process evaluation is very interesting as is the objective of including insights into the final intervention. However, I was not clear what was meant by including PAR improved uptake. Is the author saying that they collected data using PAR that was of interest to the designers? If so, that would suggest that a recommendation might be that each team needs more familiarity with the array of methods available for process research prior to starting the program so data collection could be designed in collaboration.
5. Cost effectiveness - it is not clear what the authors mean by a layer of uncertainty created by HCD.
6. Further, it is not clear why the cost of HCD was being assessed rather than the cost of the intervention.
7. Lessons learnt, item 5 is very important as is anticipating issues of interdisciplinary research from the beginning.

Is the rationale for the Open Letter provided in sufficient detail?

Yes

Does the article adequately reference differing views and opinions?

Partly

Are all factual statements correct, and are statements and arguments made adequately supported by citations?

Yes

Is the Open Letter written in accessible language?

Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Family planning, adolescent health, public health, gynecology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 12 Sep 2019

Aoife Doyle, London School of Hygiene & Tropical Medicine, London, UK

In this article, the authors discuss challenges and opportunities associated with evaluating HCD interventions. In this case, the intervention was associated with A360.

Overall, the honesty of the authors is appreciated as rarely does one get to hear about problems that occurred during project implementation. Reporting on successes and failures is important as it provides an opportunity for others to learn. Nevertheless, the descriptions of the challenges that arose were a little cryptic and it was not always clear.

I was also disappointed that the authors did not reach another logical conclusion which is that there was a methodologic gap between the public health researchers/evaluators and designers which lead to the challenges. Conducting interdisciplinary research can be challenging and requires each team to be very familiar with the other's methods and processes so they can carefully design the collaboration. For example, when the study populations grow frustrated with the multiple inquiries it suggests that the data collection was not harmonized across teams. Similarly, the breakdown in communication emerges. These are the classic challenges of interdisciplinary research and may not only apply to HCD. The authors might discuss how they prepared to work with human centered designers or vice versa. What type of training took place and how familiar were the evaluators with the design process or were the designers with public health methods?

Response: Thanks for highlighting this omission. We have added the following text to the 'Lessons learnt':

'At the start of the programme, there was a clear methodologic gap between the evaluation team who had a background in public health and the social sciences, and the design team at IDEO who were leading the HCD process. However, PSI and the other implementers, who were working closely with IDEO had a background in public health and the social sciences. The evaluation team had been thoroughly briefed about the HCD methodologies that would be used but they had only limited practical experience evaluating programmes incorporating human-centered design. Some of the challenges faced are not unique to interventions designed with HCD and will be recognisable by those who have conducted interdisciplinary research and/or led evaluations of programmatic interventions.'

Specific comments are as follows:

- It is interesting that the intervention design phase appears long. If one were to compare it to drug development or another multiphase intervention development process, then two years would be quite short.

Response: We agree that intervention development can often take a long time. We have edited the text to remove the word 'lengthy' and instead state 'the length of the intervention design period'.

- It is not clear why baseline data were collected prior to developing the intervention. How does collecting data in the four settings, the explanation provided here, constrain when data are collected?

Response: The baseline surveys needed to be conducted prior to the scale-up of the interventions i.e. so that we had baseline values for our primary and secondary outcomes. In the sentence in question, we have moved the words 'prior to scale-up' so that they appear before 'in the four study settings'. We hope that this is now clearer.

- In looking at the diagram of the timeline, it seems that the baseline survey could have been moved up to July 2018 once all decisions were made regarding what the study would be about.

Response: In hindsight, yes this would have made sense but the scale-up of the interventions were originally due to start in January 2018. The optimization phase was introduced as an additional phase only towards the end of 2017 and the extension of the evaluation period into 2020 (previously was due to end in 2019) was agreed in 2018.

- The process evaluation is very interesting as is the objective of including insights into the final intervention. However, I was not clear what was meant by including PAR improved uptake. Is the author saying that they collected data using PAR that was of interest to the designers? If so, that would suggest that a recommendation might be that each team needs more familiarity with the array of methods available for process research prior to starting the program so data collection could be designed in collaboration.

Response: Thanks, yes this is a good observation. We have added this to recommendation 4:

4. For the process evaluation, methodologies such as participant action research may identify questions that are more relevant for the implementers, and for the evaluators direct observations can be instrumental to capturing the fast-paced, highly-iterative HCD process, and to understand the 'design energy' i.e. how decisions were made at key points in the design process.

- Cost effectiveness - it is not clear what the authors mean by a layer of uncertainty created by HCD.

Response: We have replaced this phrase with 'the flexible and iterative HCD process increased the number of unknowns'.

- Further, it is not clear why the cost of HCD was being assessed rather than the cost of the intervention.

Response: The cost of the design process is being collected as is the cost of implementing the intervention itself. Within the design costs, the cost of HCD is being estimated. A360 is based on the idea that using an integrated HCD process to design the interventions will result in more impactful interventions. Therefore, the cost-effectiveness evaluation is interested in capturing the design costs as well as the intervention cost in order to benchmark against the costs and cost-effectiveness of a more traditional design process.

- Lessons learnt, item 5 is very important as is anticipating issues of interdisciplinary research from the beginning.

Response: we have added an additional recommendation on the issue of interdisciplinary research:

'1. Evaluators and implementers/designers should take time to familiarise themselves with the methodologies used by the different disciplines and have open discussions about the potential challenges of interdisciplinary research, and how they will be addressed/mitigated against.'

Competing Interests: No competing interests were disclosed.

Reviewer Report 26 June 2019

<https://doi.org/10.21956/gatesopenres.14106.r27240>

© 2019 Holeman I. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Isaac Holeman 

¹ Medic Mobile, Seattle, WA, USA

² University of Washington, Seattle, WA, USA

Overview

This letter offers reflections on the authors' experience evaluating an HCD-led global health program called A360. The letter is timely in that it raises questions that many global health programs have grappled with of late. I particularly appreciate the authors' frankness as they describe disconnects between design/implementation work and research. On the whole I am very enthusiastic about this letter moving forward in the publication process. I would like to offer a few reflections of my own in the hope of helping the author's improve their letter and of helping our research community identify better solutions to these deeply important challenges.

Major comments

This letter seems to focus on the relationship between health research and design practice, and my primary question has to do with how the authors' experiences and conclusions might have differed if their project had emphasized linking health research and design research. To be clear, I'm not referring to the way that practitioners refer to doing 'design research' or 'user research' to investigate user experiences and contexts. By design research, I'm referring to the program of empirical study that has been carried out by human-centered design researchers in the fields of design, information systems, computer supported cooperative work (CSCW), human computer interaction, and design anthropology (and beyond) in the last few decades. These fields draw heavily on ethnographic, participatory, and action research methods that are familiar to many process evaluators in global health (no doubt including the authors of this letter). Design researchers in these fields also often draw on domain-specific craft skills that enable visual expression and prototyping, intervention development and implementation, and iteration based on 'hands-on' engagement with end users. That is to say, in fields with a more established tradition of design research, the researchers often are designers. For methods of contemporary importance, see e.g., action design research ¹ and research through design ². For a paper that is older but perhaps more relevant to my comment here, see Spinuzzi's ³ argument that

participatory design is a design method, yet also is an action-research method.

While some of these design-led research fields started with collaborations across disciplines (e.g. ethnographers and computer scientists founded CSCW in the early 1990s), today doctoral students typically learn to integrate design practices and research methods in the conduct of their scholarly endeavours. Integrating design practice with research method thus becomes less a matter of coordinating two teams within an organizational bureaucracy, and something more akin to how an individual ethnographer strives for reflexivity concerning her dual roles as participant and as observer. This seems to contrast with the research described in this letter, in which the evaluations were conducted by researchers who did not consider themselves as 'doers' of the design work (if I'm mistaken about this point, I'd welcome being corrected!).

In light of the example set by human-centered design researchers in CSCW and adjacent fields, my more precise question is this: how might the challenges you encountered be reconsidered as having to do with the all too common gap between researchers and practitioners? In particular I'm thinking of these challenges:

- The frustrating communications that resulted from a prolonged design process.
- Research fatigue among participants, and the importance of jointly collecting data for design and for research so that people need not be interviewed twice.
- The need for evaluators to be present, as participant-observers in design work in order to understand the design-energy and design decisions.
- Limited uptake of findings generated by process evaluators.
- Implementers/designers needing to set aside time to participate in a process evaluation.
- The importance of ensuring that the process evaluation is phased and iterative in a manner that matches the design work.

In my experience, these are common problems/risks wherever researchers try to plug into (but aren't leading) a process led by practitioners, regardless of the field of practice. Often this is no fault of the practitioners - understanding and accommodating a scientific agenda is often beyond their remit. These gaps can be alleviated to some degree when there is one integrated design team that has the skills, methodological background, and remit to conduct a design process, produce evidence about user experiences to iteratively inform intervention development, and use this evidence to evaluate and write about their process (see e.g., [4,5,6,7](#)).

To be fair, the approach outlined in this letter seems at present to be the norm rather than the exception in global health journals: enlisting a team of researchers, who are not designers, to evaluate design practitioners, who are not researchers. The reason I find this letter so valuable is that it surfaces a set of challenges that may be inherent in this approach, and the authors should be commended for their frankness.

The authors will need to discern among themselves whether, based on their experience, they would encourage future evaluators of design-led health initiatives to include researchers from disciplines with more established traditions of integrating design research and practice. If this does resonate with your experience, I would offer two sources that you might reference that speak to a way forward for the design for global health community. The ISO standard on human-centered design ⁸ and more recent work on design for global health ⁹ emphasize that human-centered design is a highly multidisciplinary approach to research and practice - it is centrally concerned with the practical task of synthesizing diverse perspectives. This point might be used to

argue that involving design researchers alongside the range of other research disciplines outlined in the A360 approach is a way forward for streamlining adoption of your five recommendations.

As a final point, I would acknowledge that while the tradition of design research with which I am most familiar can well serve the role of process evaluation as it plays out in the global health field, outcomes evaluations and cost-effectiveness studies remain more the domain of public health researchers and health economists. Such studies can be conducted relatively straightforwardly if they focus on the interventions that resulted from a design process. While good enough for most policy purposes, that approach doesn't address the question of what proportion/aspect of good outcomes we might attribute specifically to the application of human-centered design. Given the complexity of the design process I don't know the answer here, and while the letter raises the question, the authors' five recommendations don't seem to address it either. One could follow the approach taken in evaluating the impact of applying behavior change theories in the design of communication interventions - meta-analyses of many studies with and without HCD. Such an approach would likely surface a further round of questions, unaddressed in this letter, about the extent to which the impacts of design work can/should be attributed to the design practitioner's degree of expertise, or whether a project was structured to enable the design practice to thrive, versus something integral to human-centered design as a discipline.

Minor Comments:

It might be helpful if you clarified your use of the terms "implementer" and "designer" throughout the paper. Are you referring to implementers of the design work (i.e., designers), or to other implementers who were operationalizing the interventions that the designers designed, or both?

The reference below to 'design energy' is intriguing; I think readers would appreciate slightly more explanation of what this refers to and in what sense the term is dealt with in the cited reference.

"During intervention development the process evaluation team faced challenges as the fast-paced, highly iterative HCD process meant that the decision making process, the 'design energy', often went undocumented⁴."

In the statement below, I wasn't able to follow how starting with a blank slate may have blocked the the design team from drawing on public health evidence. Design researchers and ethnographers typically strive to approach observational fieldwork in a manner that minimizes the influence of preconceived notions, but this does not prevent them from drawing on relevant theories or evidence during the process of analysis. Was disregarding public health evidence explicitly described as typical of the design process (a point I would disagree with in general, but that may well have been true to your experience), or might it have just been another researcher to practitioner disconnect?

"Designers add value to public health as they take a different approach when addressing an issue, however, in the case of A360, a preference to work from a blank slate free of the constraints of public health evidence may have impacted on an ability to respond to the PE findings."

Typos

"However, the iterative and flexible nature of HCD potentially poses some unique challenges for evaluation^{1,3,7}.~~[remove period]~~ and examples of evaluations of HCD public health interventions have so far been limited^{7,8}."

"In the 'inspiration phase' the designers immerse themselves with girls and **her their** influencers to understand the issues."

"The OE study design was finalised before the interventions themselves so there remained a risk that the evaluation was not optimal." I think this refers to Outcome Evaluation study design; it would be helpful if you could clarify on first use of the acronym.

"distinct imple- mentation phases as it was rolled **out** across Nigeria"

"because A360 was creating a new trans-disciplinary approach"

"more frequent cost data draws were required during the scale-up period to capture how those changes might affect the "production process" and associated costs."

References

1. Sein MK, Henfridsson O, Purao S, Rossi M, et al.: Action Design Research. *Management Information Systems Quarterly*. 2011; **35** (1): 37-56
2. Zimmerman J, Forlizzi J, Evenson S: Research Through Design as a Method for Interaction Design Research in HCI. *CHI*. 2007. 493-502 [Publisher Full Text](#)
3. Spinuzzi C: The methodology of participatory design. *Technical Communication*. 2005; **52** (2): 163-174
4. Perrier T, Dell N, DeRenzi B, Anderson R, et al.: Engaging Pregnant Women in Kenya with a Hybrid Computer-Human SMS Communication System. 2015. 1429-1438 [Publisher Full Text](#)
5. Kolko B, Hope A, Brunette W, Saville K, et al.: Adapting Collaborative Radiological Practice to Low- Resource Environments. 2012. 1-10
6. Kumar N, Lal P, Anderson R, Perrier T, et al.: Projecting health. 2015. 1-10 [Publisher Full Text](#)
7. Valderrama CE, Marzbanrad F, Stroux L, Martinez B, et al.: Improving the quality of point of care diagnostics with real-time machine learning in low literacy LMIC settings. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies 2018*. 2018. 2
8. Human-centered design for interactive systems. *International Organization for Standardization*. 2010. 1-40
9. Holeman I, Kane D: Human-Centered Design for Global Health Equity. *Information Technology for Development*. Forthcoming. [Publisher Full Text](#)

Is the rationale for the Open Letter provided in sufficient detail?

Yes

Does the article adequately reference differing views and opinions?

Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?

Partly

Is the Open Letter written in accessible language?

Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Human-centered design, digital health, community health.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 12 Sep 2019

Aoife Doyle, London School of Hygiene & Tropical Medicine, London, UK

Overview

This letter offers reflections on the authors' experience evaluating an HCD-led global health program called A360. The letter is timely in that it raises questions that many global health programs have grappled with of late. I particularly appreciate the authors' frankness as they describe disconnects between design/implementation work and research. On the whole I am very enthusiastic about this letter moving forward in the publication process. I would like to offer a few reflections of my own in the hope of helping the author's improve their letter and of helping our research community identify better solutions to these deeply important challenges.

Major comments

This letter seems to focus on the relationship between health research and design practice, and my primary question has to do with how the authors' experiences and conclusions might have differed if their project had emphasized linking health research and design research. To be clear, I'm not referring to the way that practitioners refer to doing 'design research' or 'user research' to investigate user experiences and contexts. By design research, I'm referring to the program of empirical study that has been carried out by human-centered design researchers in the fields of design, information systems, computer supported cooperative work (CSCW), human computer interaction, and design anthropology (and beyond) in the last few decades. These fields draw heavily on ethnographic, participatory, and action research methods that are familiar to many process evaluators in global health (no doubt including the authors of this letter). Design researchers in these fields also often draw on domain-specific craft skills that enable visual expression and prototyping, intervention development and implementation, and iteration based on 'hands-on' engagement with end users. That is to say, in fields with a more established tradition of design research, the researchers often are designers. For methods of contemporary importance, see e.g., action design research ¹ and research through design ². For a paper that is older but perhaps more relevant to my comment here, see Spinuzzi's ³ argument that participatory design is a design method, yet also is an action-research method.

While some of these design-led research fields started with collaborations across disciplines (e.g. ethnographers and computer scientists founded CSCW in the early 1990s), today

doctoral students typically learn to integrate design practices and research methods in the conduct of their scholarly endeavours. Integrating design practice with research method thus becomes less a matter of coordinating two teams within an organizational bureaucracy, and something more akin to how an individual ethnographer strives for reflexivity concerning her dual roles as participant and as observer. This seems to contrast with the research described in this letter, in which the evaluations were conducted by researchers who did not consider themselves as 'doers' of the design work (if I'm mistaken about this point, I'd welcome being corrected!).

In light of the example set by human-centered design researchers in CSCW and adjacent fields, my more precise question is this: how might the challenges you encountered be reconsidered as having to do with the all too common gap between researchers and practitioners? In particular I'm thinking of these challenges:

- The frustrating communications that resulted from a prolonged design process.
- Research fatigue among participants, and the importance of jointly collecting data for design and for research so that people need not be interviewed twice.
- The need for evaluators to be present, as participant-observers in design work in order to understand the design-energy and design decisions.
- Limited uptake of findings generated by process evaluators.
- Implementers/designers needing to set aside time to participate in a process evaluation.
- The importance of ensuring that the process evaluation is phased and iterative in a manner that matches the design work.

In my experience, these are common problems/risks wherever researchers try to plug into (but aren't leading) a process led by practitioners, regardless of the field of practice. Often this is no fault of the practitioners - understanding and accommodating a scientific agenda is often beyond their remit. These gaps can be alleviated to some degree when there is one integrated design team that has the skills, methodological background, and remit to conduct a design process, produce evidence about user experiences to iteratively inform intervention development, and use this evidence to evaluate and write about their process (see e.g., [4,5,6,7](#)).

Response: Thanks for these observations. As is common practice in the field of public health, we are engaged by the donors as external evaluators of this programme and so we deliberately had no role in the design process. It is true that many of the challenges might not have been faced if we had been integrated in the implementing/design team but this would have compromised our ability to remain independent.

In line with the references that you cite here, the implementers are conducting their own monitoring and evaluation to examine process level indicators e.g. the number of girls reached with their services. The external evaluation, conducted by researchers who specialise in evaluation, uses more complex evaluation methodologies to evaluate the impact of the intervention on proximate and also more distal outcomes e.g. population level use of modern contraceptives. The use of design research in the field of public health is relatively new and as external evaluators gain experience of design research they may need to consider alternative methodologies or ways of working with implementers. In

response to suggestions from the other reviewers we have edits the recommendations to suggest improved communication between the external evaluators and implementers.

To be fair, the approach outlined in this letter seems at present to be the norm rather than the exception in global health journals: enlisting a team of researchers, who are not designers, to evaluate design practitioners, who are not researchers. The reason I find this letter so valuable is that it surfaces a set of challenges that may be inherent in this approach, and the authors should be commended for their frankness.

The authors will need to discern among themselves whether, based on their experience, they would encourage future evaluators of design-led health initiatives to include researchers from disciplines with more established traditions of integrating design research and practice. If this does resonate with your experience, I would offer two sources that you might reference that speak to a way forward for the design for global health community. The ISO standard on human-centered design ⁸ and more recent work on design for global health ⁹ emphasize that human-centered design is a highly multidisciplinary approach to research and practice - it is centrally concerned with the practical task of synthesizing diverse perspectives. This point might be used to argue that involving design researchers alongside the range of other research disciplines outlined in the A360 approach is a way forward for streamlining adoption of your five recommendations.

As a final point, I would acknowledge that while the tradition of design research with which I am most familiar can well serve the role of process evaluation as it plays out in the global health field, outcomes evaluations and cost-effectiveness studies remain more the domain of public health researchers and health economists. Such studies can be conducted relatively straightforwardly if they focus on the interventions that resulted from a design process. While good enough for most policy purposes, that approach doesn't address the question of what proportion/aspect of good outcomes we might attribute specifically to the application of human-centered design. Given the complexity of the design process I don't know the answer here, and while the letter raises the question, the authors' five recommendations don't seem to address it either. One could follow the approach taken in evaluating the impact of applying behavior change theories in the design of communication interventions - meta-analyses of many studies with and without HCD. Such an approach would likely surface a further round of questions, unaddressed in this letter, about the extent to which the impacts of design work can/should be attributed to the design practitioner's degree of expertise, or whether a project was structured to enable the design practice to thrive, versus something integral to human-centered design as a discipline.

Response: Having read some of the references that you mention above, we have observed that evaluation in the context of design research has focused primarily on process evaluation and there is little or no discussion of how to evaluate the impact of the interventions. This would seem to support our recommendation no.5 that impact and cost-effectiveness studies would be more efficiently designed when the intervention is defined. We agree that traditional outcome evaluation studies will not necessarily 'address the question of what proportion/aspect of good outcomes we might attribute specifically to the application of human-centered design'. We have not addressed this question specifically in our recommendations as our primary objective was to evaluate the impact of

Adolescents 360 and we did not set out to evaluate HCD per se. You raise some interesting questions around the importance of understanding the relative roles of human-centered design as a discipline, the design practitioner, and the context within the design took place. As research expands in this field we hope that these questions will be explored.

Minor Comments:

It might be helpful if you clarified your use of the terms “implementer” and “designer” throughout the paper. Are you referring to implementers of the design work (i.e., designers), or to other implementers who were operationalizing the interventions that the designers designed, or both?

Response: We have updated the text to make this clearer as follows:

‘The A360 initiative is led by Population Services International (PSI) in collaboration with the Society for Family Health (SFH) in Nigeria, design partner IDEO.org, and the Centre for the Developing Adolescent at the University of California, collectively referred to here as the ‘implementers’. The implementers were all involved in the design process. PSI and SFH are responsible for the larger scale operationalising of the designed interventions.’

The reference below to ‘design energy’ is intriguing; I think readers would appreciate slightly more explanation of what this refers to and in what sense the term is dealt with in the cited reference.

“During intervention development the process evaluation team faced challenges as the fast-paced, highly iterative HCD process meant that the decision making process, the ‘design energy’, often went undocumented⁴.”

Response: ‘Design energy’ is a term that was used by IDEO and refers to how decisions were made at key points in the design process. We had included the definition in the recommendations section but we have now added earlier in the letter where we first mention the term.

In the cited reference (Tolley EB: Traditional Socio-behavioral research and human-centered design: Similarities, Unique Contributions and Synergies.2017), the author does not use the term ‘design energy’ but states the following which we feel represents the same idea:

‘The process of arriving at insights from fieldwork is more of a creative than scientific process, with less emphasis on creating an audit trail that can be replicated by others.’ (page 14)

And

‘Both observations and interpretations or insights from data collection activities may be simultaneously catalogued on individual sticky notes and then grouped and further re-grouped on table or wall space to look for key patterns that inform the research objectives. The process of determining what gets onto a sticky note – and what is overlooked, whether the source of the observation or insight is attributed to a particular participant or not and whether some end-user or influencer voices are (over) represented while others have no

representation – is generally not described.’ (page 14)

In the statement below, I wasn’t able to follow how starting with a blank slate may have blocked the the design team from drawing on public health evidence. Design researchers and ethnographers typically strive to approach observational fieldwork in a manner that minimizes the influence of preconceived notions, but this does not prevent them from drawing on relevant theories or evidence during the process of analysis. Was disregarding public health evidence explicitly described as typical of the design process (a point I would disagree with in general, but that may well have been true to your experience), or might it have just been another researcher to practitioner disconnect?

“Designers add value to public health as they take a different approach when addressing an issue, however, in the case of A360, a preference to work from a blank slate free of the constraints of public health evidence may have impacted on an ability to respond to the PE findings.”

Response: We agree that the decision to work from a blank slate is not necessarily reflective of design processes and may have been specific to how IDEO chose to approach the development of A360 solutions. The HCD process, as used during the development of Adolescent 360, did not explicitly involve any review of the public health evidence for or against interventions that have been previously explored to improve adolescent uptake of modern contraceptives. Public health was one of the ‘lenses’ in the Adolescents 360 approach but our evaluation has shown that the HCD lens overshadowed the other lenses and that the public health evidence base was at times disregarded by those leading the HCD lens (see Figure 7 https://itad.com/wp-content/uploads/2019/06/A360-MTR-Report_FINAL_27th-June-2019.pdf) On reflection we have removed the highlighted sentence.

Typos

“However, the iterative and flexible nature of HCD potentially poses some unique challenges for evaluation^{1,3,7}.**[remove period]** and examples of evaluations of HCD public health interventions have so far been limited^{7,8}.”

“In the ‘inspiration phase’ the designers immerse themselves with girls and **her their** influencers to understand the issues.”

“The OE study design was finalised before the interventions themselves so there remained a risk that the evaluation was not optimal.” I think this refers to Outcome Evaluation study design; it would be helpful if you could clarify on first use of the acronym.

“distinct imple- mentation phases as it was rolled **out** across Nigeria”

“because A360 was creating a new trans-disciplinary **ary** approach”

“more frequent cost data draws were required during the scale-up period to capture how those changes might affect the “production process” and associated costs.”

Response: Thanks. We have corrected all these typos.

Competing Interests: No competing interests were disclosed.
