How useful are systematic reviews for evaluating gender norm change interventions?

September 2015

Knowledge to action: Researching gender norms that affect adolescent girls

UKaid
from the British people

Acknowledgements

This Research and Practice Note is part of the Knowledge to Action Resource Series 2015, produced as part of a 4-year DFID-funded programme - 'Transforming the Lives of Adolescent Girls' - involving fieldwork in Ethiopia, Uganda, Nepal and Viet Nam.

This Note was written by Ella Page with Rachel Marcus, and copy edited by Kathryn O'Neill. Thanks to Caroline Harper for comments on a previous version, Sophie Brodbeck for research assistance and Catriona Foley and Gabbi Gray for the formatting of this Note.

The Knowledge to Action Resource Series was funded by DFID.

Series editors: Rachel Marcus and Caroline Harper.

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ

Tel. +44 (0) 20 7922 0300
Fax. +44 (0) 20 7922 0399
E-mail: info@odi.org.uk

www.odi.org
www.odi.org/facebook
www.odi.org/twitter

Readers are encouraged to reproduce material from ODI Reports for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI.

© Overseas Development Institute 2015. This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 3.0).

ISSN: 2052-7209

Cover photo: Family in Viet Nam © Dao Hong Le, 2013

Key points

- This Research and Practice Note draws on our experience of conducting a systematic review on the impact of communication interventions on gender norms that affect adolescent girls.
- Whether it is viable to conduct a systematic review on the effectiveness of interventions to change gender norms depends on how you define norms. If, by norms, you mean a set of widely shared beliefs and common practices, then a systematic review of interventions is viable. But if you use a stricter definition of norms (related to perceptions of standards of behaviour), a systematic review may not be viable, as few evaluations or impact studies measure norms in this way.
- Not all topics will yield sufficient evidence to be worth conducting a review, so choose your topic and questions carefully. For example, communications initiatives proved to be a rich field, where a range of evaluation methodologies have been applied. By contrast, topics such as legal reform and the role of technology in changing gender norms do not lend themselves so readily to a systematic review of interventions.
- Be aware that there may be a systemic bias towards positive results and that you will need to search for literature outside peer-reviewed journals. If we had looked at peer-reviewed journal articles only (and not included grey literature), we would have had 17 instead of 66 studies to include.
- Qualitative and mixed methods research is valuable in looking at how gender norms change. We adapted our quality assessment tools to include qualitative and mixed methods studies but excluded some research methodologies that did not involve comparisons and found it difficult to assess quality of participatory research.
- As with other types of data, there are numerous challenges in using systematic reviews to help us understand how changing gender norms affect adolescent girls. Comparisons can be very difficult because surveys addressing similar issues often use very different indicators, and much of the data does not distinguish between adolescent girls and adult women.

Setting the scene

Systematic reviews have a long history in medical research, synthesising evidence from randomised controlled trials (RCTs) and answering questions about 'what works'. But they are increasingly favoured in international development research because they can bring together the highest quality evidence on a given issue and assess the effectiveness of different interventions, as well as helping to explain why certain types of programme are effective (Mallett et al., 2012; Snilsveit, 2012; White and Waddington, 2012). There have been some short syntheses of how different interventions have attempted to transform social norms in arenas relevant to adolescent girls (e.g. Ball Cooper and Fletcher, 2013) and of communications interventions designed to contribute to changes in gender norms (e.g. Haider, 2011; UNICEF, 2012), but as of 2013 there had been no systematic assessment of this evidence.

So, to fill this gap, we conducted a systematic review to determine what works in terms of how communications interventions can change discriminatory norms that affect adolescent girls in low- and middle-income countries (Marcus and Page, 2014a). This Research and Practice Note draws on our experience of conducting this review and offers guidance on how to get the best out of a systematic review of other related interventions around adolescent girls and gender norm change.

Why do a systematic review?

Systematic reviews aim to guide policy-makers by:

- Answering focused questions (particularly about the effectiveness of interventions) drawing on a comprehensive assessment of past experience
- Gathering all available evidence rather than a cherry-picked sample (Langer and Stewart, 2014)
- Making use of high-quality evidence only, which meets explicit quality criteria.

There are certain principles you should follow when conducting a systematic review (see Box 1 on page 5):

A systematic review is a major undertaking and requires a significant amount of researcher time and resources. In our case, it took at least a year to go through the various stages that we describe in this Note: developing the protocol, crafting the research question, conducting the search, carrying out quality assessment and analysis, and reconsidering evidence where reflection on findings and iterative searching demanded that we do so.

Figure 1: When to conduct a systematic review, and when not to

Depending on the time available, your audience, your research question, and existing research, conducting a **rapid evidence assessment**, evidence **gap map** or other kind of literature review (Hagen-Zanker and Mallett, 2013) may be a better option. Even with limited time and resources, adopting some of the principles of a systematic review (see Box 1) should ensure that you achieve as rigorous a review as possible. The flowchart in Figure 1 describes when it might be appropriate to conduct a systematic review and when it might be best to choose a different type of literature review.

Lessons from our review experience

Craft your review question carefully

Our review question drew on emerging findings about the importance of communications in changing gender norms that affect the lives of adolescent girls in Ethiopia, Nepal, Uganda and Viet Nam. We had already assessed gaps in the existing evidence base, and did a preliminary scan to see whether there was likely to be enough literature to justify a systematic review (recognising that we would eliminate many studies because they were not sufficiently rigorous or relevant).

Our research question was: What types of communications activities, in what contexts, have had the greatest traction in challenging discriminatory gender norms affecting adolescent girls in low and middle-income countries? The review was also partly able to address some sub-questions around processes of change and how different interventions were designed..

Study design

In this section we set out the various stages involved in conducting a systematic review. We followed a process based on standard systematic review principles of transparency, comprehensive searching, and appraising the methodological quality of different studies. We undertook the literature search stage of the process first, followed by study appraisal and exclusion in order to narrow our sources down to the most relevant and highest quality studies. We then conducted the synthesis and analysis stages (see Figure 2 below).

Box 1: Key principles to guide your systematic review

- Have a clear and transparent search plan, and set out your inclusion and exclusion criteria
- You should search academic and grey literature, check references and citations, seek expert recommendations, and hand-search relevant websites
- At least two researchers should decide which studies are included and excluded from the review, and carry out the quality assessment process
- References should be entered into and managed through a database
- Studies should be quality assessed
- Analyse results from all included studies

TIP

Developing your research question

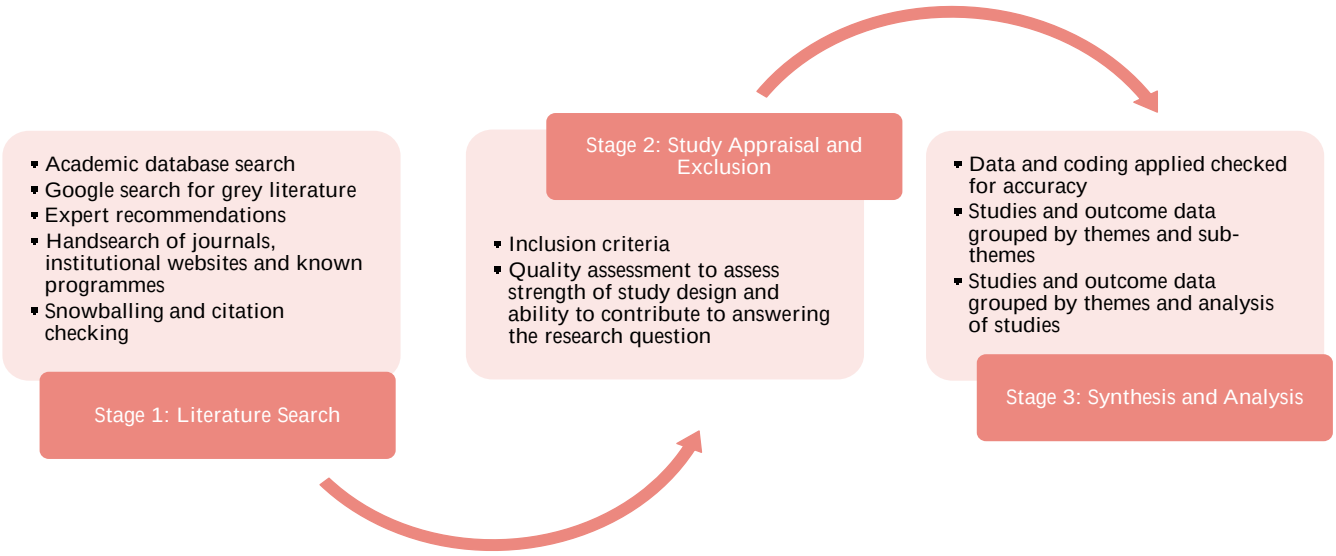
- Spend a week or two scoping the literature to ensure you have a good idea of what type of evidence is available, and whether it lends itself to a systematic review.
- Do you simply want to find out what works? Or do you want to explore how and why different interventions work? The latter will require you to gather additional evidence. Do you have the resources for that?
- Ensure your research question is focused enough to give you a clear idea of the type of programme and type of outcomes you are looking for.
- The available evidence might mean you cannot answer all your chosen sub-questions.

Stage 1: A literature search

We used a comprehensive and iterative search process to find studies from as broad a range of sources as possible, and bring together grey literature (literature that is not published commercially) and academic literature from a range of disciplines and locations. We searched for keywords in academic databases and in the grey literature, in key journals, and on websites of more than 150 organisations. Our initial search located 504 studies, mostly through Google (182), Google Scholar (95) or from hand-searching websites (117). We also sought recommendations from experts and emailed more than 40 organisations to request copies of evaluations and received 25 studies. We also checked citations and references for the studies we found. We only located six studies through academic databases. This may have been because the most useful studies had already been found in Google Scholar. It may also reflect inconsistent index and abstract writing in academic databases for the social sciences – something that has been recognised as problematic in other international development reviews (Hagen-Zanker et al., 2012; Langer and Stewart, 2014).

We tailored our search keywords to the site being searched and built in a period of experimentation so that we could be sure that our search returned the best results possible. Ideally, a literature search would continue until 'saturation point', where no new studies are returned. But time and resource constraints may make this impractical. A well-structured search and comprehensive list of search terms should ensure that you find all the relevant literature.

Figure 2: Our approach to doing a systematic review



The studies we found were initially screened by one researcher who applied inclusion criteria based on Population, Interventions, Comparator, Outcomes and Study Design (PICOS) criteria and the parameters of the review (see Table 1). All the studies included at this stage were uploaded and managed in an EPPI-Reviewer database.

Deciding on relevant outcomes

The outcomes we looked at in our review covered changes in a range of norms, attitudes and behaviour that affect adolescent girls' lives, from early marriage to son preference. We also examined evidence of changes in general attitudes towards gender equality – something that emerged as a distinct area of evidence in the studies we found.

Rather than prejudge which indicators would be relevant, we identified different kinds of indicators: the area of a girl's life that is affected by a given norm, and outcomes that would tell us whether norms were changing (possible outcome indicators) (see Table 2 for examples). We included some other measures that emerged during the review process. Although this approach means it can take longer to carry out the coding and analysis processes, it does increase the depth and breadth of the analysis.

Our focus on communications interventions was, in practice, quite broad. It included activities explicitly intended to change attitudes or practices through spreading information, persuasion or modelling new norms, attitudes and behaviour. We looked at seven types of communication programme and also included programmes that combined communications interventions with other elements such as cash transfers, vocational training or provision of health services. Had we narrowed down our focus, with tighter definitions of the areas of communications we were most interested in, this would perhaps have given the research

Table 1: Population, interventions, comparator outcomes and study design criteria (PICOS)

Population	Does the study take place in at least one low- and middle-income (LMIC) country as defined by the World Bank? Include interventions with any sector of the population or community.
Intervention	Does the study discuss communication activity, TV, radio and online activities, and peer and community activities including interpersonal communication and participatory processes, etc.?
Comparisons	Is there a comparison between two or more comparable units of analysis with and without intervention? This could be achieved via control or comparison group, a longitudinal design or other reasonable method.
Outcomes	Does the study assess changes in gender norms, attitudes or behaviour? Are outcomes related to adolescent girls, or are adolescent girls included among respondents?
Language	Is the study in English, French or Spanish? If so, include.
Publication Type	Exclude if the study is a Master's thesis.

Table 2: Areas of girls' lives we looked at, and possible outcome indicators

Area of girls' lives	Possible outcome indicators	
	Practices and behaviour	Attitudes, perceptions and beliefs
Marriage, family relationships and reproductive issues	• Change in age at marriage • Changes in whether girls have a say over who they marry and when • Changes in the extent to which married girls work outside the household • Changes in ability to/practice of divorce • Changes in ability to negotiate in sexual relationships e.g. when to have sex, use of contraception • Changes in extent of girls' input to decision-making over fertility, e.g. age of childbearing, timing and spacing of children • Change in practice and likelihood of engaging in transactional sex	• Change in perceptions of ideal age at marriage and child bearing • Change in perceptions of girls' right to have a say in who they marry and when • Change in attitude towards married girls working outside the home • Changes in acceptability of divorce • Changes in perceptions and of bride price and dowry • Change in attitude towards girls' decision-making power • Changed perceptions concerning transactional sex
'Son preference' or valuing boys more than girls	• Changes in sex ratio at birth • Change in practices favouring sons – e.g. access to food and health care • Changes in girls' and boys' overall time use and time spent on gendered activities (e.g. cooking, childcare, fetching fuel or water, looking after cattle, fixing house, using technology, etc.) • Girls' increased choice over how they spend their time • Changes in asset ownership	• Change in attitudes to sex-selective abortion • Change in desirability of having a boy/ girl child • Change in attitude and perceptions re resource allocation among children within a household • Changed perceptions towards gendered household activities • Perceptions concerning girls' and boys' ownership and inheritance of assets/ property

team more time to find other programme documents, giving us a deeper understanding of how individual programmes worked, and why.

Stage 2: Study appraisal and exclusion

During our review, we designed the quality assessment process to give equal weight to qualitative, quantitative and mixed methods research. We did this by creating a quality assessment scale, which combined adaptations of the Mixed Methods Assessment Tool (MMAT) and the Maryland Scale for Scientific Methods (MSSM). Rather than aim to find all the flaws in a piece of work, we just tried to identify any errors that were big enough to affect how the results of a study should be interpreted.

We used the MSSM tool (which ranks studies from 1-5 according to research design, focusing on the construction of comparable groups) to ensure that all studies we included had been designed with a valid comparative element, either over time or through the use of intervention and control groups. (The MSSM gives the lowest score to studies that present descriptive statistics with no control group, while it gives the highest score to well-designed RCTs with randomised assignment and comparableintervention and control groups.)

Based on our previous experience of using the MMAT, we modified it to reflect criteria we have found to be critical for identifying the reliability of different studies. Thus we drew on the Critical Skills Appraisal Programme (CSAP) tool for our assessment of qualitative research, to ensure that all research designs are assessed on sampling, research instruments, methods of data collection, and clarity of the analysis process. We introduced a question addressing the technical aspects of RCT design, and simplified and condensed the questions for mixed methods research. Our revised assessment tool also probes whether studies (of whatever design) examined how the cultural, social, political and economic contexts may have affected results. (The original MMAT only probes this issue for qualitative study designs.)

Using our study appraisal tool each study received a percentage score reflecting how well it did in the relevant sections of the tool. This meant we could easily identify high-quality studies (of any methodological tradition) across the sample, so could draw more heavily on the higher scoring studies in our analysis. At this stage, studies which scored less than 50% on our scale were considered to be of poor quality and were therefore excluded. (This cut-off point reflected the need to set the bar high enough so that the studies were methodologically

TIP

Designing and conducting your review

- Write a comprehensive protocol
- Design a research question that is as narrow as possible, while still giving full answers to important questions
- Test keywords and keyword combinations in different search locations. Include words that will narrow down your search results to the type of study you are looking for, like 'impact', 'evaluation', 'research', 'experiment', 'outcome', 'review'
- Use references and citations to broaden your search
- Contact relevant organisations and use search engines to search within websites to find additional material
- Consult literature reviews, programme overviews and other systematic reviews to identify relevant studies

credible but not so high that innovative, participatory designs were automatically excluded.) Decisions were made independently by two researchers who discussed and resolved any differences in opinion.

Despite having adapted the MSSM and MMAT tools to suit our needs, we found some studies that the reviewers felt were good quality but did not fit well within the scale, and so received an unmerited low score (including, for instance, a qualitative study that relied on triangulation rather than comparison between groups or over time). A surprisingly high number of studies did not fully describe their methodology (we excluded 13% of the studies we found for this reason) – a lesson for those writing up research findings.

Stage 3: Synthesis and analysis

In the final stage of conducting your systematic review you will need to summarise and synthesise the findings in order to answer the research question (a systematic review should be replicable so you need to be transparent about how you have organised this process). There are many different ways to synthesise; the most appropriate will depend on the types of evidence you have located (see Walker et al. 2012 for a detailed discussion of different synthesis methods). For our review, we used a narrative synthesis approach, which involved carrying out an in-depth textual analysis of the studies and their similarities, differences and outcomes.

To do this, we identified themes within the studies, used these themes to group the studies, and then carefully worked through each to build a narrative that discussed the outcomes, identified processes of change, and located features of successful programmes. We also looked out for new issues and themes that emerged as we went through the studies. This iterative process meant that we revisited studies at least five times, but found that this was necessary, not only to add codes for new issues but to really get a clear sense of the study and ensure we were being true to it. We also placed more emphasis on studies that had scored particularly well in the quality assessment phase as a way of focusing our analysis.

We found the database invaluable at this point to keep the data manageable. We could easily search for specific reports, but could also insert narrative and descriptive text to record the most relevant features of the studies for analysis, as well storing the details of the outcome data. You will need continued access to the database as you develop your analysis, and will also need to ensure that you have the time to do the material justice with careful analysis. The level of detail you need to include about each programme will vary depending on your target audience; case studies, graphs and tables derived from the evidence can be useful to keep the review from becoming too dense.

TIP

Outcomes

- Allow relevant indicators of outcomes to emerge as part of the review process as well as those you identify up front. But try to narrow down the types of programme you are interested in.
- Where available, a range of programme documents (beyond evaluations and impact studies) may provide additional insights into how the programme achieved certain outcomes.
- Where outcomes are measured in similar ways, a statistical meta-analysis of the data might be possible. But this is often not possible in international development reviews as outcomes are often diverse.

Challenges and constraints of using systematic review evidence to assess change in gender norms that affect adolescent girls

Our review found a diverse evidence base, with a range of study designs and types of programmes that produced a rich analysis across a wide range of discriminatory gender norms that affect adolescent girls. This enabled us to look at programmes that targeted girls and their families – and assessed the impact of their activities with those girls and families – to build a picture of the various factors that contribute to change. In this section, we describe some of the challenges involved in building our evidence base and highlight some areas that require specific attention when planning a systematic review of this kind.

The tendency for quantitative research designs to dominate: Most of the studies included in our review used a quantitative design, with just 16 using mixed methods and 7 using a qualitative design (see Figure 3). Overall, 50% of studies were based on quasi-experimental designs as assessed by the MSSM and just over half (52%) scored 60% or above on our study appraisal tool. Insights from quantitative research were immensely valuable as they showed the numbers of people whose attitudes or practices had changed and, sometimes, the degree of change.

The studies we looked at used a variety of analysis techniques – from simple comparisons of indicators before and after the intervention, to more complex techniques such as propensity score matching and difference-in-difference analysis (which enables researchers to assess how much of any change can be attributed to the programme). For example, Acharya et al. (2009), assessing the impact of the Better Life Options programme to empower adolescent girls in India, used difference-in-difference techniques to establish that 47% of change in participants' attitudes and beliefs was attributable to the intervention.

TIP

Study appraisal

- Put together a team that includes methodological experts from a range of disciplines
- Choose a study appraisal tool that fits with the study research aims
- Change the appraisal tool, if necessary, to deal with unexpected types of evidence
- It is important to track studies you've excluded and why you excluded them (if you amend the appraisal tool or inclusion criteria, you can pick them up more easily).

TIP

Analysis

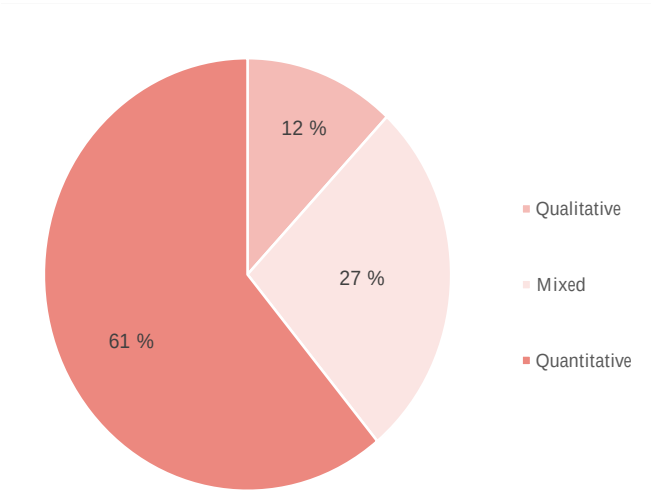
- Choose the most appropriate method for the kind of data you have; narrative methods will be more suited to reviews where study methodologies vary and/or outcome indicators are diverse.
- Keep track of your coding in a database. Being able to search, count and build reports from your data easily is invaluable.
- You may wish to give more weight to studies that have scored particularly high in the quality assessment process.
- Strike a balance between the bigger picture and relevant detail. Include case studies and graphics in the analysis to keep the text interesting.

The proportion of quantitative research was highest in studies dealing with violence and sexual relationships, reflecting a tradition of using Knowledge, Attitudes and Practices (KAP) surveys in research on sexual and reproductive health programming. In addition, some studies based their research tools on the **Gender Equitable Men (GEM) Scale** to assess changes in men's attitudes and practices around gender equality.

Difficulty in assessing qualitative and mixed method studies: Our requirement for an element of comparison may have excluded participatory studies, like Greiner et al.'s (2007) study of Al Amal ('Sails of Hope') (see Box 2) or studies which used the 'most significant change' methodology (Chandurkar et al., n.d.), which are not designed to measure impact in this way.

In many of the studies included in our review, qualitative components were able to suggest why and how interventions were effective – for example, because they enabled people to identify with a fictional character, or provided role models (of good as well as bad behaviour). Qualitative studies can also explore dissenting voices and the reasons people give for not changing what they do (or reverting to a practice they had moved away from). For example, a series of evaluations carried out of the Tostan (Dignity for All) programme in Senegal show that even in villages that have undergone education programmes and made public declarations against FGM/C, there are still dissenting voices. As one elderly woman (a non-participant of the programme) interviewed by Yoder (2008: 54) put it, *'About circumcision, the point is that I no longer have a daughter of the right age. But if I did, I would put up*

Figure 3: Study design



a fight. It would be difficult to have to give it up, I'm not trying to hide it.'

Mixed methods research studies are particularly valuable in assessing changes in norms, attitudes and behaviour because they can tell us not just what has changed but also the underlying reasons why change has occurred (Taylor and Pereznieto, 2014). Of the mixed methods studies included in our review, few achieved the highest scores using our quality appraisal process (reflecting issues with both the tool and the studies themselves). Only five had combined qualitative and quantitative research components to good effect. For example, Shahnaz and Karim's (2008) evaluation of BRAC Employment and Livelihood for Adolescents (ELA) centres in Bangladesh combined both methods to explore (with participants and non-participants) what has changed in girls lives', and how these changes have affected their relationships and decision-making power within the household.

Lack of peer-reviewed evidence: The vast majority of the studies included in our review were found among grey literature (74%), with only 17 studies found in peer-reviewed journals. This low proportion of academic studies may reflect word limits of journal articles, which means there is limited space to describe the methodology in detail.

There was a fairly even split between internal impact or evaluation studies (34) and studies conducted either with or by independent researchers (32 studies). The relatively high proportion of external evaluations may reflect increased donor requirements and funding for external evaluations. For example, the four studies that examine the Behaviour Change Communication Programme (BCCP) in southern Africa were funded by the UK Department for International Development (DFID), while the United States Agency for International Development (USAID) funded a series of evaluations of the Tostan programme (Diop and Askew, 2009; Diop et al., 2004, 2008).

No clear insights into long-term impact: We found a distinct lack of studies assessing change in gender norms that affect adolescent girls over time. Just over half (34) of the studies we looked at took place within a year of the intervention ending or while it was still ongoing; only 12 of the 66 studies took place a year or more after, and 21 were unclear about when the evaluation had taken place. For example, Jewkes et al. (2008) found that two years after the end of the Stepping Stones programme in South Africa, there was no difference between participants and the control group in terms of their involvement in rape or attempted rape, whereas after one year, participants were found to have been significantly less engaged in sexual violence than the control group.

Lack of data disaggregated by age, life stage or socioeconomic group: Although much of the data in the studies we included were well disaggregated by gender, there was less data that had been disaggregated by age. This made it difficult to distinguish between the experiences of adult women and adolescent girls. Although many of the programmes included in the review were small in scale and took place in relatively homogenous communities, we found surprisingly little analysis of how the choice of target group (or a person's socioeconomic status, for example) may have affected outcomes; nor was there much analysis of whether particular programme designs are more successful in reaching (or serving to exclude) certain social groups. Only six evaluations examined how socioeconomic differences affected access to

Box 2: Participatory evaluation: Al Amal ('Sails of Hope')

Al Amal was an 'edutainment' (part entertainment, part education) radio soap opera in Sudan, which aimed to reduce the prevalence of FGM/C. To assess the impact of the programme Greiner et al. (2007) used sketches and photos created by participants to illustrate the storylines they remembered and identified with. Men as well as women emphasised that the storylines around FGM/C resonated most with them, as there were strong parallels with their own lives; they said they had learned about FGM/C through the programme, or that it had reinforced their views that the practice should be stopped.

However, because the evaluation did not include any comparisons, it was excluded from our review.

TIP

Mixed methods research

- If research is presented as mixed methods but relies primarily on one method, assess it on the primary methodology so as not to discount good studies unnecessarily.
- Multidisciplinary teams can be useful to assess interdisciplinary research or research designs and analytical techniques that the review team are less familiar with.

communications programmes, and seven disaggregated differences in outcomes by socioeconomic group.

It was also difficult to identify the geographical areas in which programmes had most impact, even among programmes with a wide reach, such as radio and TV soaps. Hutchinson et al. (2012), investigating the Southern African HIV and AIDS Dissemination Service (SAfAIDS) and OneLove programmes in Swaziland, was one of the few studies to partially disaggregate findings geographically, focusing on programme impacts in border areas, where HIV prevalence is often high.

Lack of detail on programme activities: Very few studies presented the content of programme messages in any detail. We found little evidence about how far certain messages resonated with the target audiences, probably because such issues are discussed in reports of formative research rather than in impact evaluation studies. This lack of detail on message content and how audiences receive messages means we focused more on the type of communication activity. Resource constraints limited the team's ability to follow up with programme designers, implementers and evaluators to gain more information about each programme and any additional documentation that could contribute to a deeper understanding and analysis of how programmes functioned.

Relative impacts of separate programme components: Twenty-five of our 66 studies included more than one element, and 18 included non-communications components. But it proved difficult to disentangle the relative impacts of each component because studies were not designed for this purpose. Only two programmes attempted to analyse and compare the impacts of communications and non-communications components.

Lack of information on cost-effectiveness: Only two studies included any discussion of cost-effectiveness, so we were unable to provide any insight into the cost-effectiveness of different types of communications programmes in our review. Communications activities, particularly radio broadcasts, are sometimes perceived as relatively low cost (compared with livelihoods or scholarship programmes, for example) but we were unable to give any insights into this.

Different amounts of evidence across themes: Our review covered a range of thematic areas in which discriminatory gender norms affect adolescent girls. We found most evidence around changes in gender norms in programmes that addressed early marriage, school attendance and reproductive health, while there was a moderate amount of evidence in studies that looked at FGM/C and violence against women and girls. We found very little evidence around girls' leadership and participation or girls as peer mentors, despite the increasing popularity of programming in this area. This may reflect a programming gap, or it may be that the impact of this kind of programme has not often been measured using comparative studies.

Resources

ODI Programme outputs

Marcus, R. and Page, E. (2014) ***Changing Discriminatory Norms affecting Adolescent Girls through Communications Activities. A Review of the Evidence.*** This is the systematic review we have discussed throughout this Research and Practice Note. It gives full details of the methodology we used, and from the 61 programmes we looked at, finds strong evidence that communications interventions are an effective way to challenge discriminatory gender norms. A summary paper is available **here**.

Page, E. and Marcus, R. (2013) ***Review Protocol: Communication Activities and Discriminatory Gender Norms Affecting Adolescent Girls in Low and Middle Income Countries.*** This document describes how we developed the processes involved in our review and gives details of our search strategy, quality assessment tools and approach to analysis.

Other key literature

Cameron, D., Mishra, A. and Brown, A. (2015) 'The growth of impact evaluation for international development: how much have we learned?' ***Journal of Development Effectiveness.*** This study draws on the 3ie evaluation database to look at how published international development evaluations have changed over time. It finds that the majority of studies are on health, education, social protection and agriculture.

Hagen-Zanker, J. and Mallett, R. (2013) ***How to do a Rigorous, Evidence-Focused Literature Review in International Development.*** This practical guide takes the reader through the stages of a comprehensive literature review and includes reflections on the process of systematic reviews and analysis techniques.

Journal of Development Effectiveness: Special Issue on Systematic Reviews (2012) This set of articles includes lessons learnt and practical guides to conducting a systematic review based on the experience of a range of researchers in international development.

Libraries of systematic reviews

3ie Systematic Review database
Cochrane Library of systematic reviews in health care
Campbell Library of systematic reviews in social development
SUPPORT summaries of systematic reviews in maternal and child health
3ie Systematic Reviews
Prospero registry of reviews in health and social care
EPPI Centre List of Systematic Reviews

Other works cited

Acharya, R., Kalyanwala, S., Jejeebhoy, S.J. and Nathani, V. (2009) *Broadening Girls' Horizons: Effects of a Life Skills Education Programme in Rural Uttar Pradesh*. New York: Population Council.

Ball Cooper, L. and Fletcher, E.K. (2012) *Reducing Societal Discrimination against Adolescent Girls. Using Social Norms to Promote Behaviour Change*. London: Girl Hub.

Chandurkar, D., Negi, A. and Mukherji, S. (n.d.) *Implementing the Most Significant Change technique for monitoring and evaluation of social and behaviour change programme; process, outcomes and lessons learnt.* Discussion Paper 8. Sambodhi Research & Communication Pvt. Ltd.

Diop, N.J. and Askew, I. (2009) 'The effectiveness of a community-based education program on abandoning female genital mutilation/cutting in Senegal', *Studies in Family Planning 40(4): 307-318.*

Diop, N., Moreau, A. and Benga, H. (2008) *Evaluation of the Long-Term Impact of the Tostan Programme on the Abandonment of FGM/C and Early Marriage: Results from a Qualitative Study in Senegal*. Washington DC: Population Council.

Diop, N.J., Faye, M.M., Moreau, A., Cabral, J., Benga, H., Cisse, F., Mane, B., Baumgarten, I. and Melching, M. (2004) *The Tostan Program: Evaluation of a community-based education program in Senegal*. Washington DC: Population Council, GTZ and Tostan.

Greiner, K., Singhal, A. and Hurlburt, S. (2007) 'With an antenna we can stop the practice of female genital cutting": A participatory assessment of Ashreat Al Amal, an entertainment-education radio soap opera in Sudan'. *Investigación Y Desarrollo 15(2): 226-259*.

Haider, H. (2011) *Helpdesk Research Report: Communication Initiatives to Change Attitudes and Behaviours.* Birmingham: GSDRC.

Hutchinson, P., Silvestre, E., Anglewicz, P., Cole, E. and Meekers, D. (2012) *External Evaluation of the Southern African Regional Social and Behavior Change Communication Program, as Implemented in Swaziland*. New Orleans: Department of Global Health Systems and Development, Tulane University School of Public Health and Tropical Medicine.

Jewkes, R., Nduna, M., Levin, J., Jama, N., Bunkle, K., Puren, A. and Duvvury, N. (2008) 'Impact of Stepping Stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: cluster randomised controlled trial', *British Medical Journal* 337(a506).

Langer, L. and Stewart, R. (2014) 'What have we learned from the application of systematic review methodology in international development? A thematic overview', *Journal of Development Effectiveness* 6(3): 236-248.

Mallett, R., Hagen-Zanker, J., Slater, R. and Duvendack, M. (2012) 'The benefits and challenges of using systematic reviews in international development research', *Journal of Development Effectiveness* 4(3): 445-455.

Shahnaz, R. and Karim, R. (2008) *Providing Microfinance and Social Space to Empower Adolescent Girls: An Evaluation of BRAC's ELA Centres*. Dhaka: BRAC.

Skuse, A., Rodger, D., Power, G., Friguglietti, D. and Brimacombe, T. (2013) Communication for Development Interventions in Fragile States: A Systematic Review. *The JBI Database of Systematic Reviews and Implementation Reports* 11.3 (2013): 1-190.

Snilsveit, B., Oliver, S. and Vojtkova, M. (2012) 'Narrative approaches to systematic review and synthesis of evidence for international development policy and practice', *Journal of Development Effectiveness* 4(3): 409-429.

Stewart, R. (2014) 'Changing the world one systematic review at a time: a new development methodology for making a difference', *Development Southern Africa* 31(4): 581–590.

Stewart, R., van Rooyen, C. and de Wet, T. (2012) 'Purity or pragmatism? Reflecting on the use of systematic review methodology in development', Journal of Development Effectiveness 6(1): 58-68.

Taylor, G. and Pereznieto, P. (2014) *Review of Evaluation Approaches and Methods used by Interventions on Women and Girls' Economic Empowerment*. London: Overseas Development Institute.

UNICEF (2012) Report of the 12th UN Round Table on Communication for Development: The Role of C4D in Empowering Adolescent Girls, 14-17 November 2011. New Delhi: UNICEF (United Nations Children's Fund).

Walker, D., Bergh, G., Page, E. and Duvendack, M. (2013) *Adapting Systematic Reviews for Social Research in International Development: A Case Study on Child Protection*. London: Overseas Development Institute.

White, H. and Waddington, H. (2012) 'Why do we care about evidence synthesis? An introduction to the Special Issue on Systematic Reviews', *Journal of Development Effectiveness 4(3): 351-358*.

Yoder, P.S. with Diaye, S. and Diop, N. (2008) *Long-Term Evaluation of the Tostan Programme in Senegal: Kolda, Thies and Fatick Regions*. New York: UNICEF.

ODI is the UK's leading independent think tank on international development and humanitarian issues.

Readers are encouraged to reproduce material from ODI Reports for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. For online use, we ask readers to link to the original resource on the ODI website. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI.
© Overseas Development Institute 2015. This work is licensed under a Creative Commons Attribution-NonCommercial Licence (CC BY-NC 3.0).
ISSN: 2052-7209

All ODI Reports are available from www.odi.org

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
Tel +44 (0)20 7922 0300
Fax +44 (0)20 7922 0399

www.odi.org