# Measuring changes in social norms: learning from Voices for Change

## Summary

### Overview of approach

Voices for Change (V4C) uses a bespoke Attitudes, Practices and Social Norms (APSN) survey as its main tool to measure changes in social norms related to three ‘target behaviours’ in the four states where V4C is operational. Conducted four times over the five-year life of the programme, the APSN allows us to measure:

1. The extent to which the three target behaviours are underpinned by social norms;
2. Population-wide changes in young people’s attitudes and practices towards violence against girls and women, the involvement of women in decision-making, and the acceptance of women’s leadership.

Using a ‘difference-in-differences’ design, we are able to understand the direct effect of V4C interventions by comparing changes over time for young people who have been exposed to V4C interventions versus those who have not.

The survey design has been informed by the six-cell matrix developed by Mackie et al. (Figure 1). In the survey, for each behaviour under examination (violence against girls and women, non-involvement of women in decision-making and non-acceptance of women’s leadership) there are a set of four to six questions, covering cells A to F.

### Figure 1 Six-cell matrix for the measurement of social norm change

<table>
<thead>
<tr>
<th>What self belief is about</th>
<th>Self</th>
<th>Other - 1st Order</th>
<th>Other - 2nd Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normative Expectations</td>
<td>D: What I think I should do</td>
<td>E: What I think others should do</td>
<td>F: What others think I should do</td>
</tr>
</tbody>
</table>
What can we learn from the Attitudes, Practices and Social Norms survey?

The APSN survey offers many insights into how social norms are changing, the contribution V4C is making to these changes, and what is driving the changes, including:

- **Determining whether target behaviours are the result of social norms rather than more individually determined practices.** This will affect implementation approaches and needs to be established early in the programme.

- **Understanding how many people in V4C’s four focal states hold a specific attitude or practice and how these numbers change over time.** This is important to understand as it gives a sense of whether there is a ‘rising tide’ in support of the desired attitudes and behaviours.

- **Understanding the extent to which V4C interventions are contributing to changes in attitudes and practices for the three target behaviours.** This is critical to demonstrate the effectiveness of the programme.

- **Understanding the programme elements that are most effective and the changes they are bringing about.** This helps programme management to invest in what is working and ensures value for money.

- **Understanding who is benefiting from the programme and the effect of the programme on them.** In addition to seeing how programme effects vary by gender and state, this also enables programme management to understand how we are reaching populations with different demographics.

**Lessons from the Attitudes, Practices and Social Norms survey**

Over the life of V4C, we learnt many valuable lessons about how to measure social norm change using the APSN, including:

- **Determining as early as possible whether behaviours targeted by the programme are underpinned by social norms:** It is often assumed that common behaviours are determined by social norms (expectations of others’ approval/disapproval), but this is not always the case. If others’ opinions do not influence behaviour, and instead the behaviour is influenced by personal-level factors – such as how one handles anger or stress – the intervention will need to be structured differently from interventions focused on true social norms where others’ approval matters.

- **Being specific and selective in identifying the target behaviours to be explored through the survey:** The social norms survey approach entails at least four to six questions for each target behaviour, so the length of the survey extends rapidly. To keep the survey a reasonable length, one needs to be strategic in deciding which behaviours to include, such that they are specific enough to elicit meaningful results, yet broad enough to speak to the general behaviour category.

- **Careful question design which is critical when assessing social norms:** The survey team needs to think carefully about what precisely is the norm being explored. Mixing concepts within one four-to-six-question social norm set will muddy analysis and results.

- **Finding ways to measure the effect of large-scale interventions in a statistically valid way wherever possible:** This is challenging because interventions designed to work at scale are often expected to achieve maximum reach quickly. Phasing roll-out can be a way of generating a control group without compromising reach, but this is not always programmatically possible. Alternatively, as in the case of V4C, responses from people who have been exposed to programme interventions can be compared to those who have not, using a ‘difference-in-differences’ measurement design.

- **Adopting a mixed methods approach to get the most insights out of the survey:** Mixed methods allow researchers to use the best tools to answer different questions: quantitative data to understand population-level realities and trends, and qualitative data to dig deeper and understand causes and mechanisms – why certain behaviours are occurring, why people hold certain values, why a person might change their behaviour (or not).

- **Investing in rigorous training of survey teams:** Social norms measurement is a nuanced undertaking and it is essential that survey teams understand the logic behind the questionnaire so that they can effectively convey the questions and, when necessary, explain them.
Measuring changes in social norms: learning from Voices for Change

This paper presents the experience from Voices for Change (V4C) in measuring changes in social norms. It starts by explaining the programme’s approach to changing social norms approach and then goes on to describe the programme’s bespoke Attitudes, Practices and Social Norms (APSN) survey, our main tool for measuring social norms change. Following that, the paper identifies the learning that can be generated through the APSN and the valuable lessons we have learnt about social norms measurement in the course of the programme.

1 Our approach to strengthening the enabling environment for young women’s empowerment

V4C is a five-year programme which seeks to strengthen the enabling environment for young women’s empowerment in Nigeria. It does this by changing people’s attitudes and practices towards violence against girls and women, acceptance of women’s role in decision-making, and support for women standing for leadership positions. The programme targets 3.25 million young women and men, aged 16-25, in four focal states: Enugu and Lagos in the south of the country, and Kaduna and Kano in the north. Young people are specifically targeted as they are seen to be the most receptive to change. However, it is recognised that they are only likely to have the necessary social influence to instigate society-wide change during adulthood. The programme therefore expects that initial change will be observed within a few years of programme implementation, whilst the establishment of new behavioural norms in the three focal areas occurs over a generational time frame.

V4C’s approach to implementation and results measurement has been informed by social norms theory (Figure 2). Social norms theory recognises that individuals’ behaviours are often shaped by behavioural expectations within a social group. People adopt certain behaviours because they believe that most people in their reference group (the people who are important to them) behave in this way and that people in their reference group expect them to behave in this way. To change such socially rooted behaviours, one must influence and change a reference group, rather than isolated individuals. At the same time, it is recognised that these group-based behaviours are affected by wider societal influences coming from the family, community and formal institutions such as legislation and political structures (Figure 3: Social ecological model). Theory therefore suggests that the most effective means of promoting behaviour change and establishing new social norms is to work across the domains in a coordinated way.

Social norms are shared beliefs about what is typical and appropriate behaviour in a valued reference group. They can be defined as a rule of behaviour that people in a group conform to because they believe: (a) most other people in the group do conform to it; and (b) most other people in the group believe they ought to conform to it.


V4C’s implementation approach has therefore sought to create change at three levels (Figure 3):

- Amongst individual young men and women, as well as opinion leaders such as religious and traditional leaders and media personalities;
- Amongst society more broadly using mass media, including social media, to generate public debate about established norms of ‘being a man’ or ‘being a woman’ in Nigeria and celebrating different, more equitable norms;
- Amongst formal institutions, including strengthening the rights of women in legislation, the priority awarded to women in government plans and budgets and the role women play in political platforms.
Implementation has sought to create synergies between interventions at these three levels; for example, connecting women’s political leaders with young women involved in our transformative safe spaces programme, and having known religious leaders speak out in support of new norms on television and radio (Figure 4).

Our annual Attitudes, Practices and Social Norms (APSN) survey is V4C’s main tool for measuring two key programme elements: 1) The extent to which target behaviours are in fact social norms (critical for programme design), and 2) Measuring population-wide changes in young people’s attitudes and practices towards violence against girls and women, involvement of women in decision-making and acceptance of women’s leadership. Conducted four times over the five-year life of the programme, the APSN allows us to track attitudinal and behavioural change amongst our target group, understand the effects of V4C’s interventions and inform programme implementation strategy.

Quantitative panel survey exploring young people’s beliefs about typical behaviours in reference group and appropriateness of those behaviours;
• Qualitative focus groups and interviews with young people, adults and key influencers to explore how and why target behaviours exist;
• Representative sample of 4,798 young people aged 16–25 years old in four Nigerian states;
• Four rounds conducted over the five-year life of the programme;
• Average re-contact rate of over 95 per cent;
• Consists of approximately 250 questions which take one hour to administer;
• Costs about £200,000 to implement
The APSN is a mixed methods survey comprising a quantitative panel survey of young people and qualitative focus groups with young people, adults and key influencers. To measure population-wide change, the panel survey uses a representative sample of 4,798 young people aged 16–25 across the four focal states. The state-wide scale at which V4C operates limits opportunities for having clean treatment and control groups. Instead, using a difference-in-differences design, we are able to understand the direct effect of V4C interventions by comparing changes over time for young people who have been exposed to V4C interventions versus those who have not.

V4C’s social norms implementation approach has explicitly informed the design of the APSN survey. As the behaviours the programme seeks to change are understood to be socially based, to track change, we must not only measure what young people themselves think and do, but we must also assess young people’s expectations of others: their perceptions of how others (in their reference group) think and behave (first-order expectations), as well as their perceptions of how their reference group expects young people themselves to think and behave (second-order expectations). To gain insight into the behaviours people are contemplating and the scope for further shifts in behaviour, we can explore how young people think they and their reference group should behave, as well as what they think others expect of them (normative expectations). Furthermore, both the qualitative and quantitative components of the APSN seek to identify positive and negative sanctions that may affect the rate behaviour changes even as attitudes are shifting in the desired direction.

The six-cell matrix developed by Mackie et al. (Figure 1) neatly summarises the empirical and normative changes V4C is seeking to affect (personal behaviour and attitude, as well as first- and second-order expectations). As such, the design of the APSN survey has closely followed the logic of social norms theory, and it is one of the first large-scale surveys to put this theory into practice. In the survey, for each behaviour under examination (violence against girls and women, non-involvement of women in decision-making, and non-acceptance of women’s leadership) there are a set of four to six questions, covering cells A to F. Figure 6 provides example questions drawn from the APSN survey. The letter against each question points to the cell which the question relates to. So, taking the questions on violence against girls and women as an example, Question A interrogates ‘what I do/what is done to me’ (empirical self), Question B interrogates ‘what others do’ (empirical first order), and so on.

Whilst the six-cell matrix provides consistency in approach, in practice, there are four cells which form the core of the enquiry: A, B, D/E and F. Practically speaking, one question for each of the six cells may not always be necessary, and may, in some cases, not always be logical. For example, one would expect responses for cells D and E to be the same – that is, approval of a behaviour should be fairly consistent, regardless of whether the actor is the respondent her/himself or another person) – and can be covered through one question. Questions relating to cell C are often difficult to communicate and may only make sense if the behaviour is a private one, happening behind closed doors, as in the case of domestic violence.
Figure 6 ASPN survey questions, structured around the six-cell matrix

**VAWG**

A: How often in the last month did a man hit or slap you?  
OR How often in the last month did you hit or slap a woman you know?  
B: In other families around here, how often does a man hit or slap a woman in a month?  
C: How much would [people who matter to the respondent] think that a man in your family hits or slaps a woman?  
D+E: Sometimes a husband is annoyed or angered by the things his wife does. Would you approve or disapprove if a husband hit or beat his wife in the following situations? (Goes out without telling him; neglects the children; argues with him; refuses to have sex; doesn’t cook food properly)  
F: How much would [people who matter to the respondent] approve or disapprove if a man in your family hit or slapped a woman?  

For each of these questions, the respondent could answer ‘often’, ‘sometimes’, ‘rarely’ or ‘never’

**Women’s role in decision-making**

A: What’s the reality now? How much is the woman’s opinion considered in your family?  
B: In general, how much is a woman’s opinion considered in families around here when they make decisions?  
D: In your view, how much should a woman’s opinion ideally be considered in your family?  
E: In general, how much should a woman’s opinion be considered in other families around here when they make decisions?  
F: In your view, how much would [people whose opinion matters to the respondent] think a women’s opinion should be considered in your family?  

For each of these questions, the respondent could answer ‘a lot’, ‘some’, ‘a little’ or ‘none’.

**Women’s leadership**

A: Would you ever like to be selected for leadership of an organisation (Community Development Association (CDA), school, professional/trade association, etc.)?  
[no/probably not/probably yes/yes]  
B: Around here, how often are women selected for leadership of an organisation (CDA, school, professional/trade association, etc.)? [never/rarely/sometimes/often]  
E: Would you approve or disapprove if a woman around here was selected for leadership of political organisation (CDA, school, professional/trade association, etc.)? [strongly approve/moderately approve/neither approve nor disapprove /moderately disapprove/strongly disapprove]  
F: In your opinion, how many people around here approve of women being selected for the leadership of a local organisation? (CDA, school, professional/trade association, etc.)? [very few or none/less than half/about half/more than half/almost everyone]  

For each of these questions, the respondent could answer ‘a lot’, ‘some’, ‘a little’ or ‘none’.

Source: Voices for Change
3 What can we learn from the Attitudes, Practices and Social Norms survey?

The survey methods and social norms approach used by V4C has a number of different applications. We describe the key applications below and provide V4C-oriented examples of how this methodology deepens understanding in important programme areas.

Baseline social norms assessment
Firstly, the APSN survey enables programme implementers to use baseline data (at year 0 or year 1 of the initiative) to identify the extent to which target behaviours are social norms versus more individually determined practices. Violence against women and girls (VAWG) illustrates this point. According to the baseline study, more than 20 per cent of young women experienced violence in their household in the month preceding our survey. Yet, public disapproval of VAWG is already strong (even nearly universal), which suggests that VAWG does not follow the conventional pattern of social norms. 3

The private nature of VAWG likely contributes to the disconnect between public attitudes and private practice. Whilst addressing social norms is important on its own, it may not be sufficient to end violence against women and may need to be complemented by other initiatives which tackle other drivers of VAWG. As discussed next, this suggests a need for good formative analysis at the start of programmes to ascertain the extent to which behaviours are driven by norms and to develop appropriate VAWG interventions.

Conversely, women aspiring to lead exhibit characteristics indicative of a social norm. It is a public behaviour with defined expectations of social sanctions. The study finds that both young men and young women desire to be more active in leadership than they currently are. On a four-point scale, both young men and women’s average interest in taking leadership positions is a 3.2, but they report that actual participation of women in leadership is significantly lower (1.8 on a four-point scale). Support of women in leadership is amplified by V4C exposure – young people familiar with V4C are about 30 per cent more likely to have improved their opinion about female leaders over the last two years.

It appears likely that the expected disapproval from older adults plays an important role in reducing the number of women who stand for leadership positions. Qualitative data highlight a number of common perceptions about women’s ability to lead and make decisions under pressure which justifies the current low level of women in leadership positions. Given the desire amongst young women to be more active in leadership, it appears that this behaviour has the potential for reaching a tipping point. If key influencers such as parents and elders can be persuaded of the merits of having female leaders, a substantial number of young women say that they are ready and interested to fill these roles.

The survey also shows implementers where the population stands on average on a continuum of change, which helps to inform later programming. For example, in the case of V4C and its VAWG work, we discovered that already 65–95 per cent of young people in the target states disapproved of VAWG. Furthermore, at least 80–90 per cent of young people already expect that others would disapprove of VAWG if it happened in their own households. Originally, V4C had planned to focus on an information campaign highlighting why VAWG was undesirable; by discovering the prevalence of anti-VAWG attitudes (and especially expectations of others’ disapproval), our programming was able to adapt and focus on subsequent steps in the continuum of change, such as identifying and addressing ongoing causes of VAWG in the face of widespread disapproval. These results informed the V4C campaign strategy of NO EXCUSE – emphasising that no matter the situation, VAWG is never acceptable.

3.1 Measuring population-level changes in target behaviours

A population-representative survey like the APSN enables the estimation of how many people hold a specific attitude, practice or social norm – and how these numbers change over time. Our survey used Nigerian census data and a household listing technique with enumeration areas to draw a pool of survey respondents that is representative of all Nigerian young people (our target population) in the four states targeted by V4C. This enables us to construct survey weights that tell us how our sample is representative of the full population of young people in target Nigerian states. As a result, we can extrapolate the full reach of our programme and calculate population-wide changes in targeted attitudes and behaviours.

For example, using weighted survey data, we estimate that in the first three years of programme implementation, V4C initiatives had reached over 1.1 million young people in Enugu, Kaduna and Kano states (56 per cent of 16–25 year olds in Enugu and Kaduna; 40 per cent in Kano). We also estimate that V4C has reached an additional 300,000 young people in Lagos State, with a total estimated reach of 1.4 million 16–25 year olds in the four states by 2016. These statistics are based on respondents’ recognition of different types of V4C media and programming, including radio shows, TV, billboards, logos and online spaces.

As for population-wide change, we observed a significant increase in the young people practising the three V4C targeted behaviours. The number of young people who practised all three positive behaviours increased steadily each year by about six per cent. Much of this change was driven by Kaduna State, where 190,000 more young people practised three target behaviours in 2016 (compared to the programme’s first year of implementation).
In Kano, another 55,000 more young people practised two or three positive behaviours in 2016, compared to 2014. Overall, this means that the percentage of young people in our target states that practise two or three positive norms has increased from 60 per cent to 80 per cent in the last three years.

3.2 Identifying the programme’s causal impact in shifting attitudes and behaviours

By contacting the same individuals each year, it is possible to track how key attitudes and practices change over time. Asking about familiarity with the programme of interest then allows an analysis of whether these changes correspond to differences in programme exposure (a ‘difference-in-differences’ design). Questions assessing familiarity with programming should be as specific as possible, designed to jog respondents’ memories. Where relevant, we linked our questions to images of logos, screenshots of online content, photos of billboards, and sound clips of radio programming and speciality pop songs, so that responses would be more accurate and reflect actual rates of programme exposure.

V4C results show that young people familiar with V4C programming by the third year of implementation have significantly more positive views on gender compared to young people not familiar with programme initiatives. These young people (representing 56 per cent of young people in Enugu and Kaduna states) are on average:

- 26 per cent more likely to think women and men should enjoy equal opportunities and respect;
- 26 per cent more willing to support female leaders;
- 15 per cent more willing to speak up against violence against women;
- 11 per cent more likely to say their overall views on gender equality have improved in the last two years.

Furthermore, our results show that V4C programming is linked with higher levels of contemplation, encouraging young people to think more about beliefs that underpin gender stereotypes. We find that the V4C media (TV, radio) corresponds with bigger changes in young people’s general sense of empowerment – especially amongst women – compared to those not exposed to the programme.

3.3 Pathways for improved effectiveness and understanding mechanisms of change

The previous two sections describe effects that can be measured with standard, well-sampled surveys assessing prevalence of target attitudes and behaviours in a given population. APSN’s advantage is that it allows more nuanced understanding of what programmes are likely to be most effective and what kinds of change are caused by programming. Beyond personal attitude and practice, APSN enables an assessment of how expectations of others’ attitudes and behaviours are also changing – key in situations where social approval or disapproval may play a role in influencing personal decisions.

Two years into V4C, we observed that the VAWG behaviour area was showing mixed results amongst the general population, and that there were differences in the types of changes observed for individual attitudes/behaviours and the expectations of others. We took a closer look at the data to better understand these trends. Using regression analysis to control for baseline differences in individuals (such as their gender, state and starting VAWG attitudes), we found much more detailed evidence of 1) the ways in which V4C was working and 2) the ways in which future programming could leverage these changes to further increase impact.
For example, V4C media appears to be increasing awareness of VAWG rates and an acceptability amongst women to identify and report VAWG. V4C media also appears to be communicating to viewers that VAWG is still more pervasive than they previously thought, and that more people still condone the practice than they thought (leading some men to believe VAWG is more normative in some ways). Intensive V4C exposure (through workshops, discussion groups and exercises) has stronger and more consistently positive effects, decreasing approval of VAWG amongst both men and women, as well as perceived social approval of VAWG.

The results also suggest that Kaduna and Enugu states were at different points on the continuum of change at the time of V4C programme initiation. Kaduna residents on average appear to have had less initial knowledge about reasons why VAWG would be unacceptable, and thus we see programming increase a) the acceptability of reporting (and/or unacceptability of the behaviour) and b) awareness of the real prevalence of VAWG in the community. Baseline results showed that Enugu had higher rates of disapproval of VAWG in 2014, so the fact that we see less change in approval amongst Enugu residents two years after is not surprising.

### 3.4 Social norm ‘scores’

With at least four questions per behaviour area, analysis of change along specific dimensions can become unwieldy. One useful way to track overall change in social norms at the individual level is to create a social norm ‘score’. The responses for questions A, B, D and F in the social norms matrix are each calibrated to a four-point scale (to make sure each question gets equal weight in the index score). Then the responses are added together to estimate how far along the continuum of change a respondent is. 0 would indicate the person never practises nor approves of the behaviour, does not expect others to practise the behaviour, and does not expect others to approve of the behaviour. A 16 would indicate the person always practises and approves of the behaviour, knows others practise the target behaviour, and expects others to approve of the behaviour. A score in between indicates a mix of positive and negative attitudes/behaviours.

For example, when we control for other baseline characteristics, we find that exposure to V4C TV and radio programmes corresponds with a significant (positive) increase in young people’s social norm score related to women taking leadership positions. V4C programming correlates with a 0.2-point bigger improvement in their social norm score between 2014 and 2016, compared to people with no V4C exposure. See the following section for further examples of how to use a social norm score.

### 3.5 Which sub-populations are affected most?

Finally, by asking a robust set of demographic questions, we can determine what subgroups are responding positively to core programming. These criteria also make it possible to control for respondent-level variations, better isolating programme effects in an initiative where it is not feasible to implement clean treatment and control groups (like V4C).

For example, we explored whether high- and low-welfare respondents showed similar familiarity with and response to V4C programming. We replicated the Grameen Foundation’s Progress out of Poverty Index to assess which respondents had higher and lower welfare. We find that young people with higher welfare are more likely to recognise V4C programming (including radio and TV). This may not be surprising, given the role that financial resources play in being able to access these communication channels.

At the same time, we find that whilst the poor may not be exposed to V4C quite as often, those who do have V4C exposure respond to programming in the same ways and at similar rates as higher welfare groups. For example, as discussed above, V4C media exposure correlates with a 0.2-point increase in a respondent’s social norm score for women in leadership roles. This effect is consistent for high- and low-welfare groups.

In addition, we find that regardless of programme exposure, low-welfare respondents appear to have had bigger changes in their social norm scores for both women standing for leadership and women’s roles in decision-making, as compared to high-welfare groups. This suggests that bigger society-wide effects of positive norm change are being experienced most at the lower end of the income spectrum.

### 4 Lessons from the Attitudes, Practices and Social Norms survey

Voices for Change’s APSN survey is one of the first globally to implement a population-representative, theory-based study to track social norms and demonstrate programme impact in achieving norms change. Because of the novel nature and scale of the study, we had to innovate in a space where there were limited examples to follow. We learnt many valuable lessons during this process, including through decisions that in hindsight we would have made differently. Key insights are presented here so that future studies can benefit from our experience – both our successes and the hurdles and pitfalls we encountered along the way.
It is important to determine as early as possible whether behaviours targeted by the programme are social norms. It is often assumed that common behaviours are social norms (meaning that their persistence is caused in part by expectations of social approval or disapproval), but this is not always the case. Instead, some behaviours may be determined by personal interest, knowledge or the availability of resources. For example, women may not be standing for leadership positions due to a lack of financial resources and support, rather than widespread disapproval of the practice. Other behaviours may be associated with widespread public (dis)approval, such as violence against women, but because the practice occurs behind closed doors, there is no way for the fear of public sanctioning to influence private behaviour.

If the behaviour is influenced by personal-level factors, the intervention will need to be structured differently from interventions focused on true social norms (where others’ approval matters). Also, public social norms campaigns to change attitudes and behaviour may work differently for behaviours that are private due to a disconnect between action and negative (or positive) public sanctioning. Qualitative enquiry can help determine whether a behaviour is shaped by social norms and is therefore susceptible to change through a social norms approach.

In the programme’s early stages, it is important to correctly specify target behaviours. The social norms survey approach entails at least four questions per target behaviour, so there is a trade-off between richness of information on any given behaviour and the number of behaviours one can ask about in a survey of reasonable length. Therefore, one needs to be strategic in deciding which behaviours to include, such that they are specific enough to elicit meaningful results, yet broad enough to speak to the general behaviour category. For example, V4C chose to look at physical violence as an example of power dynamics between men and women, in part because it was harder to ask focused questions about other types of verbal or emotional control.

However, these latter categories may be easier to change (and have higher starting rates of prevalence), so in retrospect it may have been fruitful to look at these other behaviours instead. Conversations with other team members about the focus of outreach and programming can help target questions appropriately. Returning to the example, ultimately the choice of whether to focus questions on physical violence or emotional control may hinge on what kinds of power dynamics are emphasised in programme outreach.

During this process, programme implementers should think about the full arc of the project. Will the project’s methods, audience, and/or emphasis evolve over the course of the programme? If so, it may be warranted to include questions on a broader range of behaviours. The first survey establishes the baseline attitudes and behaviours against which any change will be measured. If the behaviour is not included in the baseline survey, it is much harder to assess change in this target area over time.
Survey language always matters, but careful question design is especially important when assessing social norms. The survey team should think carefully about what precisely is the norm that should be studied? In the case of women in leadership, for example, is it the opportunity to stand for leadership that matters, the decision to stand, or the actual selection of women to the leadership after they decide to stand for the position? Each of these would require a slightly different set of questions, and mixing the concepts within one four-question social norm set will muddy analysis and results. Similarly, questions in cell B (what others do) are deceptively tricky to write well. This is because it is easy to combine a question about frequency and prevalence into one question. For example, take the question ‘On average, how often do men around here hit or slap a woman?’ This conflates the idea of frequency (how often on average do men in other households hit or slap a woman?) and prevalence (in how many households around here are women beaten with some degree of regularity?). Similarly, for cell C, care should be taken to distinguish between strength of approval on average, and the percentage of people who at least approve of the behaviour to some extent.

Social desirability bias can impact results, and creativity is needed to circumvent this bias, especially when behaviours occur in private. In panel studies, researchers will want to consider whether any changes in behaviour are real (e.g., due to programme impact or broader socio-cultural trends) or merely a consequence of people feeling more (or less) comfortable with the survey exercise over time. If they trust the enumerator more, people may report more accurate responses, making it look like behaviour is changing when in actuality the only change has been trust in the survey environment. If people are afraid of being judged by the enumerator, or that their privacy is compromised, there will likely be underreporting of sensitive behaviours.

Whilst V4C attempted to minimise social desirability bias, however possible, (guaranteeing privacy for respondents, building an atmosphere of trust and non-judgement), it is difficult to identify the extent to which social desirability bias still affects results. As a general rule, questions that ask about ‘what others think/do around here’ are more likely to elicit accurate results, compared to personal questions of ‘What do you think/do?’. That said, when behaviours are private (such as VAWG in the household), respondents may not have an accurate perception of what happens in other households. Where possible, it is useful to look for outside data (such as statistics from non-profit or government sources of reported VAWG) to see whether trends measured in the survey match trends measured elsewhere.

This situation is further complicated when desired changes in the population can actually affect one’s willingness to report a behaviour or attitude. For example, as women feel more self-confident, they may be more willing to report VAWG. Therefore, over time, it may appear that VAWG rates are increasing when in fact women are feeling more comfortable reporting the violence that has been occurring all along. V4C addressed this in the latter rounds of its survey by building in additional questions to test various causal pathways for changes in VAWG rates including actual increase in violence, increased exposure to VAWG messaging (and acceptability to discuss it), increased self-confidence, and increased comfort in the interview process. For private, sensitive behaviours like VAWG, it can be very challenging to overcome social desirability bias completely, but well-planned surveys and outside data can help minimise this bias within and across survey waves.

In terms of research design, there is a tension between implementing an initiative that maximises programmatic reach, and an implementation that allows for a statistically valid measurement of programme effectiveness. As the reach of the programme increases, it is increasingly hard to identify an equivalent control group – and ideally, evaluators would want an equivalent control group so they can compare whether the programme area shows more positive change compared to any changes observed in the control area.

This issue is especially challenging in social norms work when media such as the internet, radio and TV are used to spread programme messages. Media reaches large groups of people with ambiguous or non-existent geographic boundaries, and people who are not in media catchment areas are likely to be systematically different (poorer, less educated) in ways that could affect the measurement. Yet, with social norms-oriented initiatives, it is important to reach large sections of the population, given that perceptions of (and actual) social approval/disapproval affect behaviour.

One way of dealing with this dilemma is to phase in programming in different regions or states over one or more years. This way, change during the first year can be compared between exposed and non-exposed groups, and the large overall number of people reached by the initiative is still reached by programme mid- or endline. V4C needed to roll out its programming simultaneously in all states, making this ‘staggered entry’ design unfeasible. Instead, the programme used two other approaches to obtain a reasonably plausible counterfactual in its analysis: first, after Year 1, programming was scaled back in some local government areas (LGAs). Although the target LGAs were not chosen randomly, it still

Voices for Change Legacy Paper August 2017
Measuring Changes In Social Norms: Learning from Voices for Change
allows for a ‘difference-in-differences’ approach where we compare change over time in target and non-target LGAs. If all LGAs are expected to respond similarly to the programme initiatives, and some receive programming whilst others do not, we can still assess whether social norms improved in the target LGAs as compared to change over time in non-target LGAs. This approach requires asking the same survey questions to the same people at least twice.

In addition, we asked respondents to self-report familiarity with different types of V4C programming. We ask this at the end of the survey, so that any reminders about these programmes do not influence responses to questions about key attitudes and behaviours. If people do not correctly remember V4C exposure, we would expect this to bias against us finding significant programmatic effects (as it would make the data ‘noisier’); yet, we do find strong effects correlated with V4C exposure in all three key behaviour areas.

Whether identifying social norms at baseline, identifying effective mechanisms for change, or tracking this change over time, a mixed methods approach works well (a mixture of quantitative and qualitative methods). V4C’s approach to identifying norms and measuring social norm change is primarily quantitative, with qualitative data being used to help interpret and understand the quantitative data. Mixed methods allow researchers to use the best tools to answer different questions; quantitative data to understand population-level realities and trends, and qualitative data to dig deeper and understand causes and mechanisms – why certain behaviours are occurring, why people hold certain values, why a person might change their behaviour (or not).

A quantitative-led approach is essential for a programme like V4C that operates on a markedly large scale, targeting over 3.25 million young people as well as millions of secondary targets amongst young people and older adults. Focus groups, interviews and other qualitative work can help unpack a group’s attitudes or behaviours, but with such a large population, a large quantitative study is the best way to ensure that the people sampled are in fact representative of the population as a whole.

Representative data lend confidence amongst donors (some of whom are sceptical of the effectiveness of social norms change) making significant investments that the programme is resulting in measurable, population-wide change. The data also facilitate comparison, for example between states and genders, helping implementers to understand where further attention is needed and, in combination with qualitative data, to hone implementation strategies.

When conducting social norms-oriented surveys, it pays to invest in rigorous training of the survey team. Although quantitative data collection is often more straightforward than qualitative data collection, in a social norms survey, data collectors are exploring complicated sets of questions. Tone, emphasis and precise wording all matter to convey questions about the respondent’s expectations of others. It is critical that the survey team understands the logic behind the questionnaire – a half-day training in social norms theory has been valuable in building a deeper understanding of the survey objectives amongst enumerators, resulting in more nuanced feedback during the survey pilot and higher comprehension rates amongst respondents. If necessary (and acceptable to programme researchers), the enumerator should also know how to explain the question using additional words.

5 Further resources


On a scale of 0–100, researchers estimated that the Progress out of Poverty Index (PPI) is a measure of poverty, based on a questionnaire with ten questions about a household’s characteristics and asset ownership, which are scored to compute the likelihood that the household is living below the poverty line. For more information, go to http://www.progressoutofpoverty.org/.

In tipping point theory, change occurs when a certain proportion of influential individuals take up a new behaviour, with others following their lead.

Control variables include gender, age, state, economic status and baseline opinions about women in leadership.

The Progress out of Poverty Index (PPI) is a measure of poverty, and asset ownership, which are scored to compute the likelihood that the household is living below the poverty line. For more information, go to http://www.progressoutofpoverty.org/.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

One way of confirming whether a behaviour is a social norm is to determine whether what people do is the same as what others think people should do (does A or B = F)? A social norm is a behaviour that is informed by what others do (B) and what others think I should do (F). So, if a behaviour is a social norm, then the measurement of B or A, if a behaviour is private and hard to observe in the general population is roughly equal to the measurement of F. In the case of VAWG, then, we would look at whether the rate of VAWG in one’s household was consistent with whether people expect others to approve of VAWG. In this case, we see that occurrence of VAWG is much higher than perceived approval of VAWG, suggesting that VAWG may not follow a classic social norm dynamic.

In tipping point theory, change occurs when a certain proportion of influential individuals take up a new behaviour, with others following their lead.

Control variables include gender, age, state, economic status and baseline opinions about women in leadership.

The Progress out of Poverty Index (PPI) is a measure of poverty, based on a questionnaire with ten questions about a household’s characteristics and asset ownership, which are scored to compute the likelihood that the household is living below the poverty line. For more information, go to http://www.progressoutofpoverty.org/.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

1. The final survey round was conducted in summer 2017, and endline results will be available in September 2017.
3. Approximately 370,000 young people in Enugu; 976,000 in Kaduna; 1,250,000 in Kano; and 650,000 in Lagos.
5. One way of confirming whether a behaviour is a social norm is to determine whether what people do is the same as what others think people should do (does A or B = F). A social norm is a behaviour that is informed by what others do (B) and what others think I should do (F). So, if a behaviour is a social norm, then the measurement of B or A, if a behaviour is private and hard to observe in the general population is roughly equal to the measurement of F. In the case of VAWG, then, we would look at whether the rate of VAWG in one’s household was consistent with whether people expect others to approve of VAWG. In this case, we see that occurrence of VAWG is much higher than perceived approval of VAWG, suggesting that VAWG may not follow a classic social norm dynamic.
6. In tipping point theory, change occurs when a certain proportion of influential individuals take up a new behaviour, with others following their lead.
7. Control variables include gender, age, state, economic status and baseline opinions about women in leadership.
8. The Progress out of Poverty Index (PPI) is a measure of poverty, based on a questionnaire with ten questions about a household’s characteristics and asset ownership, which are scored to compute the likelihood that the household is living below the poverty line. For more information, go to http://www.progressoutofpoverty.org/.
9. On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.
10. On a scale of 0–100, researchers estimated that someone with a poverty index score of 44 would have a 50 per cent chance of living in poverty. Therefore, we use a poverty index of 44 as the dividing line between low- and high-welfare groups.

About V4C

Voices for Change (V4C) is a £29 million programme funded by UK Aid, working to strengthen the enabling environment for gender equality in Nigeria. The programme targets young women and men aged 16–25 years old. It operates in four states in Nigeria: Enugu, Kaduna, Kano and Lagos and for some activities, at federal level. V4C is a unique example of a programme applying social norms theory at scale and is addressing the structural barriers to gender equality; in particular, discriminatory and harmful attitudes, behaviours and social norms. The three normative areas that V4C seeks to change are women’s voice and leadership, women’s role in decision-making and violence against women and girls.

V4C recognises that for young women to be better supported, change needs to happen at scale – not only at the individual level but also within wider society.

- At the individual level, V4C works with adolescent women and girls to provide them with the skills, knowledge and confidence to challenge, together with men, boys, religious leaders, traditional leaders and networks of men and women, discriminatory social norms, and create change in their colleges, homes, workplaces and communities.
- At the community level, V4C works with men and boys, religious and traditional leaders, and networks of women and girls to create a critical mass of support for gender equality, accelerating change and shifting negative norms.
- At the social-structural level, V4C works to change discriminatory laws, create better policies, and direct assets towards women and girls, sending a message about changed social norms through political and legal structures.

Conceived as the pilot stage of a twenty-year vision, V4C began implementation in October 2013 and will end in September 2017.

Credits

This paper is one in a series of V4C legacy papers that have been written to share practical guidance and learning on the different approaches and strategies used by the programme. For more information, visit www.v4c-nigeria.com.

This summary has been written by Elaine Denny and Claire Hughes with contributions from Emeka Nwankwo. Readers are encouraged to quote and reproduce material from this summary in their own publications. In return, V4C requests due acknowledgement and quotes to be referenced as above.