

RESEARCH ARTICLE

Open Access



# Implementation of the Good School Toolkit in Uganda: a quantitative process evaluation of a successful violence prevention program

Louise Knight<sup>1\*</sup>, Elizabeth Allen<sup>1</sup>, Angel Mirembe<sup>2</sup>, Janet Nakuti<sup>2</sup>, Sophie Namy<sup>2</sup>, Jennifer C. Child<sup>1</sup>, Joanna Sturgess<sup>1</sup>, Nambusi Kyegombe<sup>1</sup>, Eddy J. Walakira<sup>3</sup>, Diana Elbourne<sup>1</sup>, Dipak Naker<sup>2</sup> and Karen M. Devries<sup>1</sup>

## Abstract

**Background:** The Good School Toolkit, a complex behavioural intervention designed by Raising Voices a Ugandan NGO, reduced past week physical violence from school staff to primary students by an average of 42% in a recent randomised controlled trial. This process evaluation quantitatively examines what was implemented across the twenty-one intervention schools, variations in school prevalence of violence after the intervention, factors that influence exposure to the intervention and factors associated with students' experience of physical violence from staff at study endline.

**Methods:** Implementation measures were captured prospectively in the twenty-one intervention schools over four school terms from 2012 to 2014 and Toolkit exposure captured in the student ( $n = 1921$ ) and staff ( $n = 286$ ) endline cross-sectional surveys in 2014. Implementation measures and the prevalence of violence are summarised across schools and are assessed for correlation using Spearman's Rank Correlation Coefficient. Regression models are used to explore individual factors associated with Toolkit exposure and with physical violence at endline.

**Results:** School prevalence of past week physical violence from staff against students ranged from 7% to 65% across schools at endline. Schools with higher mean levels of teacher Toolkit exposure had larger decreases in violence during the study. Students in schools categorised as implementing a 'low' number of program school-led activities reported less exposure to the Toolkit. Higher student Toolkit exposure was associated with decreased odds of experiencing physical violence from staff (OR: 0.76, 95%CI: 0.67-0.86,  $p$ -value < 0.001). Girls, students reporting poorer mental health and students in a lower grade were less exposed to the toolkit. After the intervention, and when adjusting for individual Toolkit exposure, some students remained at increased risk of experiencing violence from staff, including, girls, students reporting poorer mental health, students who experienced other violence and those reporting difficulty with self-care.

**Conclusions:** Our results suggest that increasing students and teachers exposure to the Good School Toolkit within schools has the potential to bring about further reductions in violence. Effectiveness of the Toolkit may be increased by further targeting and supporting teachers' engagement with girls and students with mental health difficulties.

**Trial registration:** The trial is registered at [clinicaltrials.gov](http://clinicaltrials.gov), NCT01678846, August 24th 2012.

**Keywords:** Process evaluation, Violence, Schools, Children, Corporal punishment, Uganda

\* Correspondence: [louise.knight@lshtm.ac.uk](mailto:louise.knight@lshtm.ac.uk)

<sup>1</sup>London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Full list of author information is available at the end of the article



## Background

In Uganda, physical punishment in schools has been banned since 1997, and became illegal in May 2016. Despite this, physical punishment persists as normal practice in primary schools. In one study conducted in 2012, over half of school children reported experiencing physical violence from staff in the last week and 8% sought treatment for injury from a healthcare provider [1]. This high level of violence in schools is not unique to Uganda. Recent national prevalence studies have shown that 40% of 13-17 year olds report being punched, kicked or whipped by a teacher in the last week in Kenya and in Tanzania 50% reported experiencing physical violence from a teacher when they were under 18 year of age [2, 3].

The Good School Toolkit developed by Raising Voices, a Uganda-based Non Government Organisation (NGO), is one of the very few rigorously evaluated interventions designed to reduce physical violence from school staff to students. The Toolkit is a violence prevention behavioural intervention that aims to change school operational culture. We recently conducted a trial to assess effectiveness as part of the Good Schools Study. The trial results showed a 42% reduction in relative risk of students experiencing physical violence in the last week from staff (corresponding to an odds ratio: 0.40, 95% CI: 0.26 to 0.64,  $p < 0.001$ ) [4]. However, even with this highly effective intervention, 31% of students in the intervention schools had experienced physical violence from staff in the last week after intervention delivery. This may be due to variation in delivery of the intervention by Raising Voices, school-led Toolkit implementation by school staff, adoption of the intervention by schools, or due to the characteristics of schools or composition of students within schools.

Exploring reasons for variation in intervention impact in complex interventions such as the Good School Toolkit is important to inform intervention development, adaptations, program monitoring, cost-effective implementation and scale-up [5]. Qualitative and quantitative methods for evaluation bring complementary insights into what and how interventions are delivered and received. Quantitative evaluation can not only describe what was implemented but also explore dose response and how delivery and reach may vary across contexts or participant characteristics [6]. Quantitative evaluation can therefore add valuable insight into how implementation is associated with effectiveness, and can highlight inequalities to inform intervention future development. However, quantitative process evaluations measuring implementation are relatively rare, with a limited number of health related behaviour

change interventions reporting on how implementation is related to outcomes [7–10].

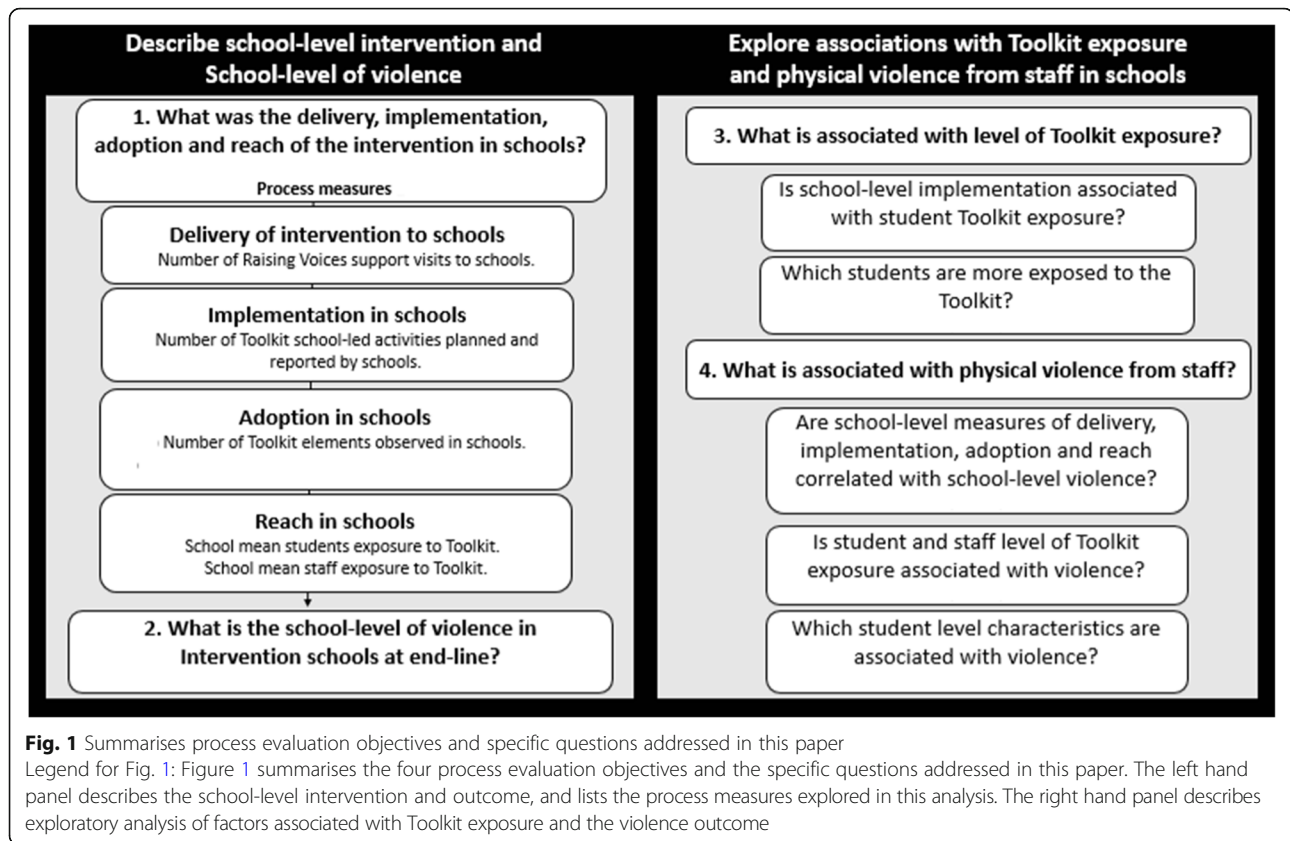
Here we present a quantitative process evaluation of the Good School Toolkit intervention in Uganda that focuses on the Toolkit implementation and explaining variations in effectiveness across schools. In associated papers we present the study protocol [11], main study results [4], qualitative findings on pathways of change [12] and an economic evaluation of the Good School Toolkit [13].

**Objectives** The specific process evaluation objectives are to: 1) describe measures of Toolkit delivery, implementation, adoption and reach in schools, 2) describe the prevalence of physical violence from staff across the intervention schools at baseline and at endline, 3) explore factors associated with student's exposure to the Toolkit, and 4) explore factors (student and school-level) associated with physical violence from staff at endline. Figure 1 summarises the process evaluation objectives, Toolkit implementation process measures and sub-questions addressed in this paper.

## Methods

### Overview of main trial

The Good Schools Study consists of a cluster randomised controlled trial, a qualitative evaluation, an economic evaluation, and the process evaluation presented here. A cross-sectional baseline survey was conducted in June 2012, and endline survey during June 2014 in 42 primary schools in Luwero District, Uganda. Luwero is a large district with urban trading centres and rural sub-districts. Using school enrolment lists, 151 eligible primary schools were identified and grouped into strata according to the sex ratio of their students. Forty-two primary schools were randomly selected proportional to the stratum size, and all agreed to participate. Stratified randomisation was carried out after the baseline survey, with 21 schools receiving the Good School Toolkit and 21 forming a waitlisted control group. All 21 intervention schools completed the Toolkit intervention, which took place over 18 months (corresponding to four school terms) between September 2012 and May 2014. At endline, 92% of the sampled students were interviewed and 91% of all staff interviewed. The study was approved by the London School of Hygiene and Tropical Medicine Ethics Committee (6183) and the Uganda National Council for Science and Technology (SS2520). Our protocol and main trial results, that include details on ethics and consent procedures for all participants in the Good School Study, as well as child protection referral procedures, are published elsewhere [4, 11, 14].



## Intervention

The Good School Toolkit is publicly available at [www.raisingvoices.org](http://www.raisingvoices.org). It consists of a six-step process that involves implementation of a series of about 60 activities described in manuals and supporting leaflets and posters. Activities are coordinated by two lead teacher ‘protagonists’ and two student representatives in each school; some activities involve outreach to parents and the surrounding community. Schools receive one-on-one support visits and phone calls from Raising Voices staff. The intervention uses the Transtheoretical behaviour change model [15] and involves the application of behaviour change techniques shown to be effective in other fields and in other violence prevention interventions, including setting a goal, making an action plan, and providing social support [16]. A description of the Good School Toolkit intervention and a summary of the Toolkit six-step process are presented in Additional file 1.

## Process evaluation design

Drawing broadly on the Grant et al. 2013 process evaluation framework [17] we describe the overall implementation of the Toolkit in terms of four implementation components: Raising Voices delivery of the intervention to schools, school-led implementation of Toolkit activities,

adoption of Toolkit elements in schools and the reach of the intervention to students and teachers at school. A pre-specified set of process data were prospectively captured in the 21 intervention schools during the 18-month implementation of the Toolkit and at the study end-line survey in 2014. All the measures were developed and pilot tested in the field before use and are described with data quality summarised in Additional file 2.

**Delivery of intervention to schools by Raising Voices** was measured using data routinely collected by Raising Voices Programme Officers as part of the intervention. Each of four officers directly supported five or six schools, and a Program Manager provided oversight and audit visits to the schools. All interactions with the schools—including technical support visits, group trainings and telephone calls—were systematically documented by each program officer termly. For analysis, the ‘delivery’ variable used was the number of Raising Voices technical support visits per school.

**School-led implementation of Toolkit activities** was measured using termly ‘action plans’ routinely completed by schools as part of the intervention, and standardised activity monitoring forms distributed to intervention schools collected only for the Good Schools Study. Schools were asked to complete action plans detailing

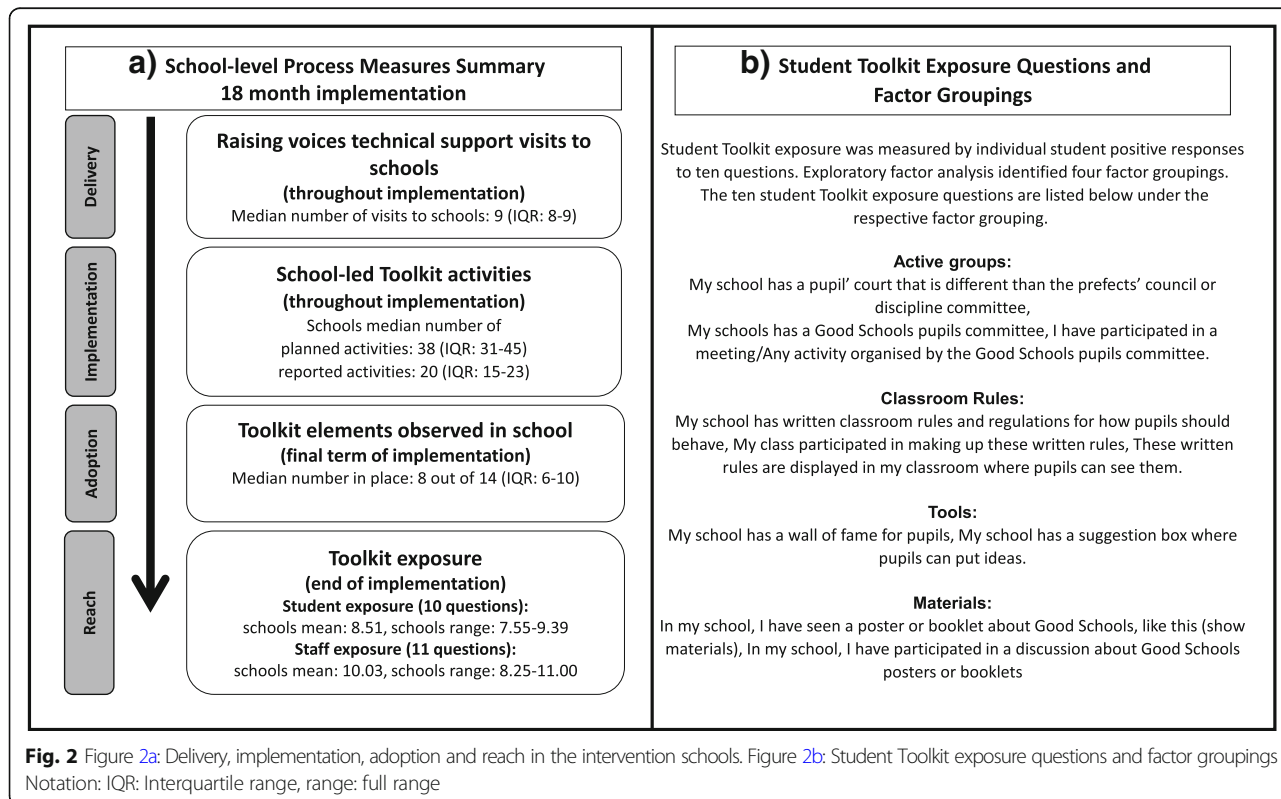
all Toolkit activities they would conduct at the start of each term, and then to record the details of each activity they actually completed on a monitoring form (one monitoring form per activity). The number of planned and completed school-led activities reported were measured throughout implementation and used as the implementation process measures.

**Adoption of Toolkit elements by schools** was tracked by an independent ‘Study Process Monitor’ who was hired specifically to collect data on implementation of the intervention for the Good Schools Study. Once every term, she asked the Head Teacher or teacher protagonist a standard set of questions about Toolkit structures present in each school, a sub-set of which were verified by direct observation (e.g., the presence of a Good School mural). The number of Toolkit elements present in school were measured using data from the last term of implementation and used as the adoption process measure. This multi-item measure was constructed based on fourteen observations (count 0-14, Cronbach alpha 0.72 that are listed in Additional file 2).

**Toolkit reach to students and teachers in schools** was measured as school level aggregate of individual Toolkit exposure, using data from the 2014 endline surveys, conducted as part of the Good Schools Study. The surveys included a set of 10 student questions and 11 school staff questions on exposure to the Toolkit that captured awareness and participation in various Toolkit

activities and processes. Student exposure questions are listed in Fig. 2b and include, “My schools has a Good Schools pupils committee.” Staff questions include “My school has a suggestion box where pupils can put ideas” and “My school has a Good Schools staff committee”. A full list of staff questions is available in Additional file 2. Binary yes or no responses to exposure questions were constructed as a count 0-10 and 0-11 for students and staff respectively, with a higher score representing more exposure. The staff exposure score had high internal reliability as measured by Cronbach’s alpha 0.72. Staff surveyed were 87% teaching staff and 13% administrators, cooks and other staff. The student exposure score had a low Cronbach’s alpha of 0.56 and therefore exploratory factor analysis was performed on student exposure responses. We used tetrachoric correction to account for binary variables, we retained factors with eigenvalues above 1 and Promax rotation was performed [18, 19]. The resulting predicted estimations were generated for each of factor groupings and summed for the total student Toolkit exposure score. The exploratory factor analyses identified four factor groupings that represented exposure to Good School Toolkit: 1) active groups, 2) classroom rules, 3) tools and 4) materials (items described in Fig. 2b).

**Other factors** were measured using items from the baseline and endline surveys, described elsewhere [1, 4]. Violence exposure (in students) and use (in staff) was



measured using an adapted version of the International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool-Child Institutional (ICAST) [20]. Mental health difficulties were measured using the Strengths and Difficulties Questionnaire (SDQ) [21]. Both these measures have been widely used and have established data on aspects of validity and reliability in different contexts [20–25]. All questions were adapted and piloted in the local population prior to use in the survey.

**Data management**

The Study Process Monitor was responsible for all process data and tracked termly submission of data from schools and Raising Voices program team, as well as collation and storage of paper copies. Raising Voices intervention delivery data were entered on to standardised excel entry sheets by programme staff, copies of the school Good School term action plans were collected from schools by Program Officers and number of activities listed on the action plans captured on to an excel tracking form by Study Process Monitor and all other process data were collected on paper and double entered on to Epidata. The Study Manager oversaw data entry and the study London based Data Manager ran all data comparison and double entry cleaning reports. The study endline survey data, including Toolkit exposure, were captured on programmed tablet computers with algorithms designed to eliminate erroneous skips. All further data management and analysis was performed in Stata/IC 13.1.

**Analysis**

The sample size calculation for the main trial is describe in the main trial paper [4]. All process evaluation analysis is exploratory and therefore no sample size calculation relating to process outcomes are specified for these analyses. All analysis was conducted using Stata/IC 13.1. Endline survey data from the twenty-one intervention schools for both students ( $n = 1921$ ) and staff ( $n = 286$ ) were used for analysis. Missing data from baseline and endline surveys were very low (less than 1% for all measures), however there were more missing data for routine implementation measures collected by schools. Data quality issues, including missing data for process measures, are fully described in Table 2.1 in Additional file 2.

**Implementation process and endline violence in schools'**

To describe the overall implementation of the Toolkit in schools, we calculated school level mean or median values of each measure, along with standard deviations or inter-quartile ranges. To describe the variation in levels of violence in the intervention schools, we calculated the mean percentage of students who experienced

any physical violence from staff in the last week, at baseline and endline, along with 95% Confidence intervals.

**Factors associated with students' toolkit exposure**

To explore whether student's endline Toolkit exposure was associated with attending a school with "low", "medium", or "high" Toolkit implementation, two unadjusted linear regression models are fitted, accounting for school clustering by fitting school as a random effect (Table 1). The two models explore if students had a higher mean Toolkit exposure score (outcome) if they attended a school that reported more (a) planned and (b) completed Toolkit activities, over the implementation period.

To explore which student characteristics were associated with student's Toolkit exposure a linear regression model was fitted, adjusting for school clustering. Choice of factors

**Table 1** Are students attending schools that implement a higher number of school-led Toolkit activities more exposed to the Toolkit?

	Unadjusted mean increase in student individual Toolkit exposure (95% CI)
Model a) Students in schools reporting low, medium or high number of <i>planned</i> school-led Toolkit activities, over four terms of implementation:	
Number of students in regression model	$n = 1921$
Students in schools reporting:	
Low number Toolkit activities planned	ref
Medium number Toolkit activities planned	0.14 (0.05-0.22) < 0.003
High number Toolkit activities planned	0.27 (0.15-0.38) < 0.001
Model b) Students in schools reporting low, medium or high number of <i>completed</i> school-led activities, over four implementation terms:	
Number of students in regression model	$n = 1921$
Students in schools reporting:	
Low number of completed Toolkit activities	ref
Medium number of completed Toolkit activities	0.30 (0.20-0.39) < 0.001
High number of completed Toolkit activities	0.23 (0.11-0.34) < 0.001

Notation: Linear regression models showing crude associations between students categorised as attending schools reporting low, medium or high levels of Toolkit implementation measured by number of a) planned Toolkit activities and b) completed Toolkit activities, and association with students self-reported exposure to the Toolkit intervention. School level clustering was accounted for by modelling school as a random effect

Legend: Students in schools that reported low, medium or high *planned* school-led Toolkit activities prospectively on their termly action plans (total number planned for the four school terms over the 18 month implementation period), where: low = less than or equal to 30 planned activities, medium = 31 to 45 planned activities and high = 46 or more planned activities. Students in schools that reported a low, medium or high number of *completed* school-led Toolkit activities by the end of the four terms over 18 month implementation period, where: low = less than or equal to 14 completed activities, medium = 15 to 22 completed activities, high = 23 or more completed activities

explored was informed by the conceptual framework for this analysis shown in Additional file 3. Factors that were found to be significantly associated with Toolkit exposure ( $p$ -value < 0.05) were retained and included in a multivariable model. The reporting of any functional difficulty was identified a priori as a potential important influencing factor and was included in the multivariable analysis despite no significant crude association. Pre-hypothesised interactions between explanatory variables (sex and number of meals eaten, and sex and mental health) were investigated by including interaction terms in the models. Due to evidence of a non-normal distribution of the student's Toolkit exposure measure, non-parametric bootstrapping with 2000 repetitions was used to estimate bias corrected confidence intervals (Table 2).

**Factors associated with physical violence at endline**

Spearman's Rank correlation coefficient was used to examine if schools' level of Toolkit implementation was correlated with the prevalence of violence in schools. Schools were ranked separately by total number of support visits by Raising Voices program staff, number of school-led Toolkit activities planned and reported as completed, number of Toolkit elements observed in place in school and by school mean of aggregate student,

staff and teacher Toolkit exposure. Correlation of each measure with prevalence of physical violence in schools at endline; and change in violence between baseline and endline was explored (Table 3).

To explore whether student's Toolkit exposure to different components of the intervention, measured by the continuous factor scores for each factor grouping 'active groups', 'classroom rules', 'tools' and 'materials', are associated with student reports of physical violence from staff in the last week logistic regression models were fitted, accounting for school clustering (Table 4). Similarly, logistic regression models were fitted, adjusted for school clustering, to explore associations between staff Toolkit exposure and reported use of violence against students in the last week and last term (Table 5). Lastly, to explore which student characteristics were associated with students' self-reported physical violence in the last week from staff, logistic regression models were fitted adjusting for school. Choice of factors explored was based on associations found to be important at baseline and drawing on the conceptual framework for this analysis (Additional file 3). All factors explored for bivariate association are listed in Additional file 4. Factors that were found to be significantly associated with self-reported physical violence ( $p$  < 0.05) were retained and included in

**Table 2** Student factors associated with level of student Toolkit exposure

	Unadjusted mean difference in student Toolkit exposure (95% CI) <i>P</i> -value	Adjusted mean difference in student Toolkit exposure (95% CI) <i>P</i> -value
Number of students in model	<i>n</i> = 1921	<i>n</i> = 1914 (a)
Student in current school for full Toolkit implementation period	0.25 (0.14-0.35) < 0.001	0.24 (0.14-0.34) < 0.001
School grade 2014:		
Primary 5	ref	ref
Primary 6	0.05 (-0.03 to 0.14) 0.211	0.04 (-0.05 to 0.12) 0.431
Primary 7	0.30 (0.20 to 0.39) < 0.001	0.23 (0.14 to 0.32) < 0.001
Poorer mental health	-0.42 (-0.56 to -0.27) < 0.001	-0.36 (-0.51 to -0.20) < 0.001
Female students	-0.18 (-0.26 to -0.10) < 0.001	-0.18 (-0.25 to -0.10) < 0.001
Three or more meals eaten yesterday	0.10 (0.03 to 0.18) 0.009	0.07 (-0.01 to 0.15) 0.084
Absent from school one or more days in last week	-0.12 (-0.20 to -0.02) 0.022	-0.09 (-0.18 to -0.01) 0.064
Any self-reported functional difficulty	0.03 (-0.05 to 0.12) 0.455	0.09 (0.00 to 0.18) 0.046
Any other violence experienced in the last year(b)	-0.10 (-0.18 to -0.02) 0.014	-0.05 (-0.13 to 0.03) 0.206

Notation: CI Confidence Interval. (a) Missing data: six students' responses to being absent from school in the last week and one student response to number of meals eaten yesterday. (b) Any other violence experienced in the last year besides physical violence from school staff

Legend: Linear regression models show individual student factors crude associations with Toolkit exposure and a fully adjusted model that includes all other factors as co-variates, both models account for school clustering by fitting school as a fixed effect

**Table 3** Are school-level process measures correlated with school prevalence of physical violence against students' from staff at endline and change in prevalence over the Toolkit implementation period

		School endline physical violence prevalence (a)	Change in school physical violence prevalence (b)
School-level process measures	School-level mean (sd), full range	Spearman's Rank Correlation Coefficient (rho) ( <i>p</i> -value)	Spearman's Rank correlation Coefficient (rho) ( <i>p</i> -value)
Number of schools	21	21	21
Delivery of intervention			
Total number of Raising Voices support visits to school.	8.76 (1.5), 6-12	0.06 (0.797)	0.18 (0.443)
School-led implementation			
Total number of planned school-led activities.	36.95 (9.09), 18-52	0.06 (0.791)	0.09 (0.694)
Total number completed school-led activities reported.	19.33 (6.44), 6-36	0.23 (0.308)	0.12 (0.615)
School adoption			
Toolkit structural elements observed in place, last term of school implementation (count: 0-14).	8.19 (2.25), 4-12	0.30 (0.181)	-0.50 (0.021)
Toolkit reach			
School mean of student Toolkit exposure (count: 0-10)	8.51 (0.45) 7.55-9.39	-0.31 (0.172)	0.35 (0.125)
School mean of staff Toolkit exposure (count: 0-11).	10.03 (0.70) 8.25-11	0.18 (0.442)	-0.05 (0.816)
School mean of teacher Toolkit exposure (count: 0-11).	10.42 (0.48) 9.29-11	-0.23 (0.313)	0.48 (0.029)

A full description of each school-level process measure is provided in Additional file 2

(a) school prevalence of violence at endline: percentage of students reporting physical violence from staff in the last week post intervention – a negative correlation indicates schools with a higher process measure level correlates with schools that have a smaller proportion of students experiencing violence in school

(b) change in school prevalence of violence over implementation period: calculated as endline minus baseline percentage of students reporting physical violence from staff in the last week – a positive correlation indicates schools with a higher implementation process measure correlates to a larger decrease in violence between baseline and end-line

**Table 4** Student Toolkit exposure measures associations with physical violence from staff experienced by students in the last week

Student Toolkit exposure measures	Physical violence from staff in last week (students self-reports)
	Unadjusted Odds Ratio (95% CI) <i>p</i> -value
Total students in models	<i>n</i> = 1921
Exposure to Toolkit (total factor score)	0.76 (0.67-0.86) < 0.001
Toolkit exposure factor groupings:	
Active Groups	0.60 (0.47-0.78) < 0.001
Classroom Rules	0.59 (0.42-0.82) < 0.001
Tools	0.80 (0.56-1.14) 0.223
Materials	0.62 (0.45-0.85) 0.003

Notation: CI confidence interval. Legend: Logistic regression model of association between student Toolkit exposure and physical violence from staff, self-reported by students. School clustering has been accounted for by modelling school as a fixed effect. Question items in each factor grouping are presented in Fig. 2b

a multivariable model that included a-priori students' exposure to the Toolkit. Pre-hypothesised interactions between sex, mental health, any other violence in the past year and number of meals eaten yesterday, were investigated by including interaction terms in models.

## Results

### Toolkit implementation: Delivery, school-led implementation, adoption and reach

All twenty-one intervention schools completed the six steps of the Toolkit intervention (Additional file 1), although Raising Voices program officers reported that the intensity and quality of each step varied between schools. School implementation measures are summarised in Fig. 2. On average Raising Voices delivery of support visits took place twice per school term and this was similar between schools. The number of planned and reported school-led activities implemented throughout the whole 18 months (four school terms) implementation period varied between schools, ranging from 18 to 52 and 6 to 36 respectively. Little variation in schools reach to

**Table 5** Toolkit exposure and self-reported use of physical violence against students in the last week and last term, presented for all school staff and restricted to teaching staff only

Staff and teachers Toolkit exposure measures	Physical violence use against students In the last week (staff self-reports)	Physical violence use against students In the last term (staff self-reports)
	Unadjusted Odds Ratio (95% CI) <i>p</i> -value	Unadjusted Odds Ratio (95% CI) <i>p</i> -value
Total staff in model	<i>n</i> = 283	<i>n</i> = 283
Staff Toolkit exposure (count 0-11)	1.0 (0.84-1.21) 0.912	0.95 (0.85-1.07) 0.428
Total teachers in model	<i>n</i> = 246	<i>n</i> = 246
Teachers Toolkit exposure (count 0-11)	0.84 (0.66-1.07) 0.148	0.77 (0.61-0.98) 0.031

Legend: mixed-effect logistic regression models of association between all school staff and teacher only Toolkit exposure and self-reported physical violence used against students. School clustering has been accounted for by modelling school as a random effect

students was seen with schools' average student Toolkit exposure at endline ranging between 8 and 9 out of 10 exposure questions. School averages of staff Toolkit exposure varied between 8 and 11 out of 11 questions.

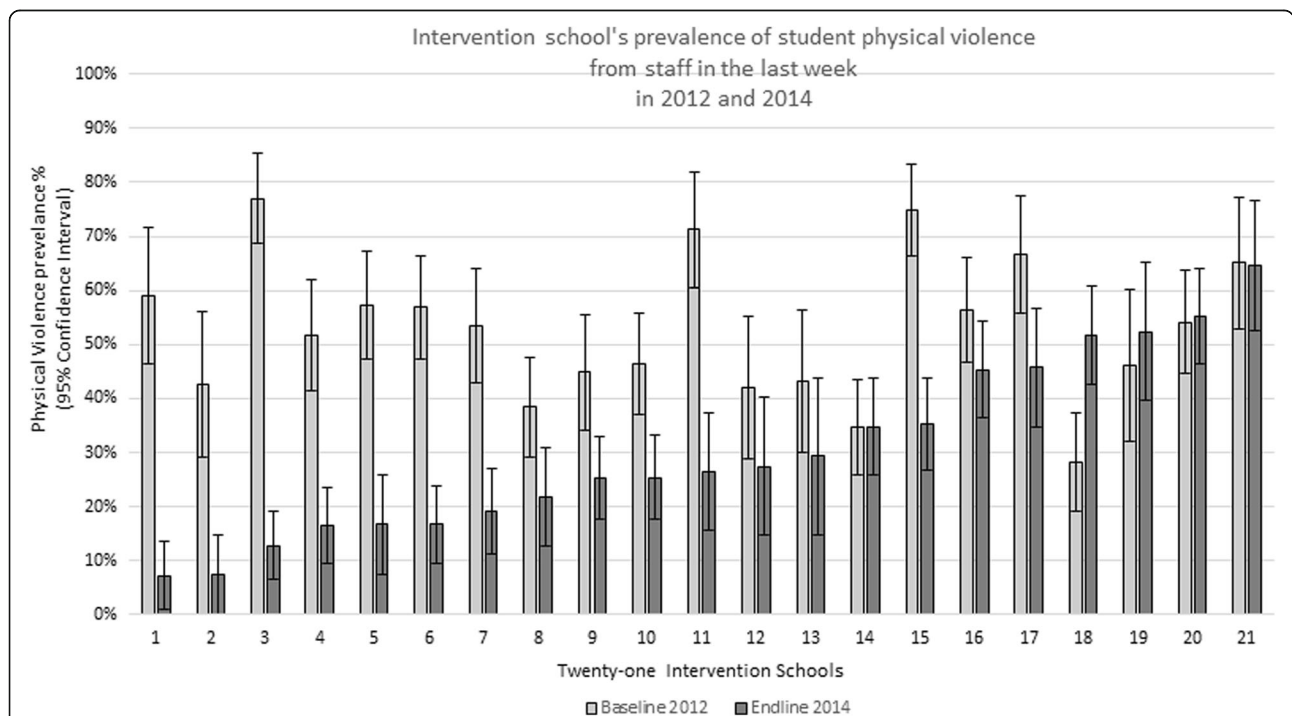
**Variation in the prevalence of physical violence in intervention schools at endline**

At the end of Toolkit implementation, intervention schools had varying proportions of students reporting physical violence from staff in the last week, ranging from 7.25% (95%CI: 0.97%-13.53%) to 64.62% (95%CI: 52.67%-76.56%) (Fig. 3). Although the Toolkit intervention has been proven

to bring about a large average reduction of physical violence in schools, five of the intervention schools remained at similar levels or increased in prevalence of violence at endline compared to baseline, although only one school had a statistically significant increase in violence. In contrast, twelve of the remaining seventeen schools showed a statistically significant decrease in prevalence over the implementation period.

**Factors associated with students' Toolkit exposure**

As expected, student exposure to the Toolkit at endline was associated with attending a school that had reported a



**Fig. 3** Prevalence of physical violence in intervention schools at baseline and endline

Legend: Intervention school-level of physical violence from staff in the last week, reported by students, graph presenting baseline and end-line school means and 95% confidence intervals.



higher number of school-led Toolkit activities implemented over the intervention period. Students in schools with a 'high' or 'medium' number of activities planned had a higher mean exposure score compared to those students in schools with a 'low' number planned. A similar association is seen for students attending schools that report more of the Toolkit activities completed, with students in these schools having a higher mean exposure score compared to those in 'low' implementing schools (Table 1).

When investigating individual student's characteristics and Toolkit exposure, students that had been in the current school for the full implementation period and those students in a higher school grade, had increased odds of exposure to the Toolkit. Girls were much less likely to be exposed than boys. Children experiencing more mental health difficulties were the group of students least likely to be exposed (Table 2).

#### School and individual factors associated with physical violence from staff in schools' at endline

Schools where teachers reported more exposure to the Toolkit had larger decreases in prevalence of school violence between baseline and endline (Table 3). One unexpected observation was that schools with an increased number of Toolkit elements present during the last term of implementation, had a smaller decrease in violence over the implementation period. No other school process measures were significantly correlated with school level

violence at endline or decrease in prevalence of violence over the implementation period.

Table 4 shows that students with increased exposure to the Toolkit have a 24% reduction in odds of experiencing physical violence irrespective of which intervention school they attend. In terms of exposure to specific Toolkit processes, participation in Good School 'active groups', 'classroom rules', and Toolkit 'materials' were each independently associated with reduced odds of experiencing violence, whereas awareness of Toolkit 'tools' did not show a significant association.

Table 5 shows that teachers who were more exposed to the Toolkit reported less use of physical violence against students over the last week and last term, although this was only statistically significant over the past school term period.

Table 6 shows that after accounting for individual level exposure to Toolkit activities, girls, students reporting difficulties with self-care (such as washing or dressing), mental health difficulties and those who experienced other violence within the last 12 months, remained at increased risk of experiencing physical violence from staff after the intervention. Having eaten three or more meals the previous day was associated with lower odds of violence.

## Discussion

### Summary of main findings

The intervention delivery, in terms of technical support visits to schools from Raising Voices program staff, was similar across schools. In contrast, we observed large

**Table 6** Student factors associated with experience of physical violence from staff at endline

Student factors	Unadjusted Physical violence from staff towards students.	Adjusted Physical violence from staff towards students.
	OR (95% CI) <i>p</i> -value	aOR (95% CI) <i>p</i> -value
Number of students in model	<i>n</i> = 1921	<i>n</i> = 1920
Exposure to Toolkit (factor score)	0.76 (0.67-0.86) < 0.001	0.82 (0.72-0.93) 0.001
Female sex	1.38 (1.12-1.70) 0.002	1.31 (1.05-1.62) 0.015
Self-reported functional difficulty with memory and/or concentration	1.38 (1.04-1.84) 0.028	NA (1)
Self-reported functional difficulty with self-care (e.g. washing)	3.68 (1.51-8.95) 0.004	3.36 (1.31-8.60) 0.011
Eaten three or more meals yesterday	0.67 (0.53-0.84) < 0.001	0.74 (0.59-0.94) 0.012
Any other violence experienced in the last year (2)	2.27 (1.82-2.84) < 0.001	2.00 (1.60-2.52) < 0.001
Poorer mental health (3)	3.69 (2.45-5.57) < 0.001	2.51 (1.63-3.85) < 0.001

Notation: aOR adjusted odds ratio, CI confidence interval, NA Not Applicable. (1) Self-reported functional difficulty with memory and/or concentration removed from the final multivariate model due to co-linearity with the mental health measure. (2) Any other violence experience in the last year besides physical violence from school staff. (3) Poorer mental health ranging from 0 (low difficulties) to 2 (high difficulties)

Legend: unadjusted and adjusted logistic regression models are presented, both accounting for school clustering by fitting school as a fixed effect. Adjusted model includes all other student factors as covariates in the model

variation between schools in Toolkit implementation and adoption measured by number of school-led Toolkit activities planned and reported as completed by schools and in the number of Toolkit structural elements observed in place in each school. Despite variation in levels of school-led implementation, at the end of the implementation period we found high levels of Toolkit reach as measured by school mean aggregates of students and staff exposure to the Toolkit, with relatively little variation across schools. Regardless which school they attended, girls, lower grade students and students with mental health difficulties were less exposed to the intervention. In terms of intervention outcomes, although there was a large average reduction in physical violence, the reduction and prevalence at endline varied widely across schools, ranging from 7% to 65% prevalence. Even after the intervention and adjusting for individual Toolkit exposure, girls, students with mental health difficulties, who experienced other violence, had eaten fewer meals or had difficulties with self-care, were at higher risk of violence compared with other students.

### Strengths and limitations

Like all studies, this process evaluation has strengths and limitations. We used a combination of data collected through a rigorous program of research, and routine monitoring data collected by schools themselves as they would do in the absence of formal research. Data on Toolkit activities planned and completed activities reported by schools were incomplete, and in some cases whole term data were not available from schools. Therefore, these measures may not reflect accurately the number of Toolkit activities in all of the schools. This highlights the need for simple tools and process data embedded in the programme implementation. Measures developed for this study had not been fully tested to determine validity and reliability, however all questions were piloted prior to use. Measures development, construction and data quality issues including estimated internal reliability are documented in Additional file 2. Staff and students responded to interviewer-administered questionnaires during baseline and endline surveys. Like all self-reported measures, there may be some social desirability bias in responses. For teachers, we would expect those more exposed to the Toolkit to report using less violence. However, we would expect a bias in the opposite direction for students, where students who are more exposed to the intervention report more violence experience, which could dilute the effect. Strengths of this evaluation include high student and staff response rate and triangulation of wide range of data sources, use of an independent study process monitor and prospective data collection of process measures.

### Why did the intervention work better in some schools than others?

Taken together, our results suggest that increased exposure to the intervention was the main driver of larger intervention effects. This is true for both increased teacher exposure, and increased student exposure, which supports the Toolkit's holistic model of engagement with multiple actors within a school to try to engender school-wide change. These results also highlight the importance of on-going training and activities to ensure newly transferred teachers and students are exposed to, and invested in, the Good School Toolkit intervention.

Counter-intuitively, schools with more Toolkit structural elements observed in place by our Study Process Monitor in the final term of the intervention implementation had smaller decreases in violence. This might be explained by a 'last push' in schools that were slower to implement. This may have resulted in more visible elements in the final term, without full engagement in the program of work required to sustain these elements or investment in the underlying change. None of the other implementation measures captured at the school level were associated with changes in school violence. We are, however, limited by our relatively small sample size of 21 schools.

In terms of exposure to specific Toolkit processes, participation in 'active groups', 'classroom rules', and Toolkit 'materials' were each independently associated with reduced violence, whereas awareness of 'Tools' alone was not. However, qualitatively we found that tools such as the 'wall of fame' were perceived positively by students and staff suggesting these are important in promoting reward and praise in schools [12]. In summary, all Toolkit processes seem to be important to bring about change, a finding that supports the idea that multiple and repeated engagement with Toolkit ideas contributes to intervention effectiveness.

Two schools that represent unexpected outcomes are shown on Fig. 3: school number 21 had the highest prevalence of violence at endline and shows no change over time, and school number 18 had a significant increase in violence at endline compared to baseline. Through our program monitoring, we are aware that both of these schools had changes in staff during the implementation of the Good School Toolkit. Anecdotally, it is possible that new staff may be less invested in the program, or in some cases even reverse policy on corporal punishment or dismantle the school wide intervention—shifts Raising Voices has experienced in other schools following staff turnover. This highlights the importance of strong leadership and ownership for the programme to remain successful and stresses the need for early identification of schools requiring additional on-going support. Raising Voices also reported that motivation of the Good School teacher protagonists was also an important factor

influencing schools sustained implementation. This suggests that identifying and building on protagonists' motivation may also be important for an effective program.

The Good School Toolkit intervention includes activities that foster a supportive school environment, aim to challenge negative social norms, and encourage student participation and confidence. Nevertheless, we saw that girls and students with mental health difficulties were less exposed to the Toolkit, irrespective of which school they attend. Conversely, and in line with the inclusive nature of the intervention, students reporting functional difficulties (for example, with sight or hearing), who were more absent from school, who had experienced other violence and eaten fewer meals, reported the same levels of exposure to the Toolkit as other students in their schools. These findings could indicate that there is something within all schools, reflecting broader societal and gender norms, that is preventing girls and students experiencing mental health difficulties from participating in school Toolkit activities. Also, the difference observed in Toolkit exposure for girls, compared to boys, could help explain the main study finding that the intervention was slightly more effective in reducing violence in boys overall [4].

#### **Which students remain at higher risk of violence after the intervention?**

Thirty-one percent of students in intervention schools still experienced physical violence from staff in the last week at endline, demonstrating that even after a highly effective intervention some children were still more at risk of violence compared to their peers. Even after accounting for level of exposure to the intervention, girls, students reporting difficulty with self-care, students who had eaten fewer meals, reported more mental health difficulties and those who experienced other violence besides physical violence from staff in the last year, remained at higher risk of physical violence from staff after the intervention.

The underlying reasons why girls might participate less in Toolkit activities, and remain at higher risk of physical violence from school staff even if they do participate in Toolkit activities, is an area requiring further investigation. Our findings may reflect the need for a social norm shift and sustained school-wide cultural change to address negative gender norms. In addition, the Toolkit may benefit from additional activities and content intentionally designed to enhance participation of these groups. This is in line with findings from other school-based intervention studies in Sub-Saharan Africa. A review of HIV prevention programs in youth in Sub-Saharan Africa that included twenty studies delivered in schools, or schools and community, concluded that "attention should go to studying implementation difficulties, sex differences in responses to

interventions and determinants of exposure to interventions" [26]. The need for behaviour change at cultural level was highlighted in a sexual reproductive health school-based intervention in Tanzania, where authors emphasised the need to train and monitor teachers to "have supportive relationships with pupils, boost pupil confidence, encourage critical thinking, challenge dominant gender norms, and not engage in physical or sexual abuse." [27]. While this quantitative evaluation supports the idea that the Good School Toolkit can bring about school cultural change around the use of violence, we also show that this is not universal to all schools and that harmful norms around violence use against some students remain irrespective of which school the student attends. Although gender equity is implicit in many of the processes and the design of the Toolkit, there may be value in emphasising and making it more explicit to teacher and students, in doing so highlighting Toolkit activities that specifically support gender equality and address negative gender norms.

Student risk factors for violence might reflect the circumstances that influence their likelihood of being physically punished at school. For example, students who have eaten fewer meals might be hungry and this could trigger punishment for having less attention in class. There is some evidence of behavioural and attention problems among hungry children from studies in the United States [28–30]. Students with mental health difficulties might also have difficulty concentrating and might display behaviours that could be seen as disruptive or challenging to teachers who may not have the tools or techniques to deal with children exhibiting these difficulties [31, 32]. This is particularly true in a context with limited teaching resources and large classroom sizes [12, 33] and suggests that positive discipline alternatives to physical punishment may not be well applied or may not be sufficient to prevent violence in some circumstances. Such circumstances may challenge teachers who are still transitioning to non-violent approaches to maintain discipline in their school. Teacher capacity-building around the extra support and skills required for some students in the classroom learning environment is a potential area to focus Toolkit intervention activities.

Our results draw attention to children that may have complex issues, including being poly-victimised, having difficult home environments and dealing with a variety of mental health difficulties [1, 34–36], highlighting the need for strengthening the intervention around building sustained capacity within the school system that recognises children experiencing overlapping vulnerabilities. The layered supportive environment that the Toolkit fosters might be one of the very few opportunities for marginalised children to be involved in a positive school programme, where they can be supported to form better

relationships and improve communication - skills that can help build confidence and resilience that influence their choices and future life trajectories.

#### Can we develop indicators for programmers to monitor effectiveness?

Collecting data on program implementation, as an indicator of intervention effectiveness, can be useful for future scale-up efforts [8, 37]. The most useful indicators would be school-level, and easily captured during routine implementation monitoring. Unfortunately, none of our implementation measures collected at the school level were associated with the change in school violence over the intervention period or the prevalence measured at endline directly after intervention. These results should be interpreted with caution, as we have low power to quantitatively detect effects across only 21 intervention schools. The lack of association may also be due to the limitations of the monitoring data we collected. Data collected from schools as part of the implementation of the Toolkit had low levels of completeness, and this may have masked a real association between these measures and intervention effectiveness. We also might not have tracked important process indicators. For example, school enforcement of policies or standards promoted by administration are difficult to measure, but may be important predictors of intervention impact. Hence, refinement and reliability testing of in-school assessments, that are in line with the intervention theory of change, could be a useful addition for future program monitoring. This also underscores the importance of capturing qualitative information to understand school context as well as expert programmer's knowledge of school specific issues relating to effective implementation of the Toolkit.

Our individual-level measures of exposure to the Toolkit were associated with intervention effect, as would be expected. School-level aggregate of teacher's exposure was also associated with larger reductions in school violence over the implementation period. This may be a potential indicator that could be used by program implementers—although it would require surveying teachers. In addition, our results indicate that monitoring the number of school-led Toolkit activities planned each term could be a simple way to use routine programme data to identify low implementing schools. However, our results highlight that none of the process measures investigated are good indicators of overall intervention effectiveness, and should therefore not be interpreted in the same way as data on violence outcomes.

#### Conclusion

Even though the intervention is highly effective at reducing violence against children in school, we found that some schools require additional support to bring about

effective and sustained change. It may be possible to increase the effectiveness of the Toolkit by increasing student and teacher exposure. The next layer of Good School programme refinement should attempt to engage with children who were less exposed—in particular, girls, those with poorer mental health and in lower school grade.

#### Additional files

**Additional file 1:** Good School Toolkit description (DOCX 22 kb)

**Additional file 2:** Good Schools Study process evaluation outcome and process measures (DOCX 44 kb)

**Additional file 3:** Summary of all student factors explored in analysis (DOCX 17 kb)

**Additional file 4:** Conceptual frameworks for process evaluation analysis (DOCX 170 kb)

#### Acknowledgements

The Wellington Ssekadde and the Good School implementation team, our team of interviewers and supervisors, Jane Frank Nalubega and the CHA team, our team of interviewers and supervisors, Anna Louise Barr, Heidi Grundlingh, Jennifer Horton, Professor Maria Quigley (independent statistician), our trial steering committee Professor Russell Viner (chair), Dr. Lucy Cluver, and Jo Mulligan.

#### Funding

This work was funded by the MRC/DfID/Wellcome Trust via the Joint Global Health Trials Scheme (to K. Devries), and the Hewlett Foundation and the Oak Foundation (to D. Naker).

#### Availability of data and materials

The datasets generated and/or analysed during the current study will be publicly available on LSHTM repository ten years after end of data collection in 2014.

#### Author's contributions

LK managed the data collection of the endline surveys, led the process evaluation data analysis and drafted the manuscript. EA participated in the design of the study, advised on data analysis, participated in data interpretation and writing of the manuscript. AM and JN participated in data collection and writing of the manuscript. JCC managed the data collection baseline surveys and participated in the interpretation of the data and writing of the manuscript. JS managed data entry, and participated in the writing of the manuscript. SN, NK and EJW participated in data interpretation and writing of the manuscript. DE participated in the design of the study and writing of the manuscript. DN initiated the idea to do the study, participated in the design of the study, obtained funding, provided comments on data interpretation, and participated in drafting of the manuscript. KMD designed the Good Schools Study including the process evaluation, participated in data collection, data interpretation and obtained funding. All authors read and approved the final manuscript.

#### Ethics approval and consent to participate

The study was approved by the London School of Hygiene and Tropical Medicine Ethics Committee (6183) and the Uganda National Council for Science and Technology (SS2520). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Written voluntary informed consent was obtained from all individual participants included in the study. Headteachers provided consent for schools to participate in the study. Parents were notified and could opt children out of participation in survey data collection. Children themselves provided consent for participation. Staff provided consent for participation in survey data collection.

### Competing interests

Dipak Naker developed the Good School Toolkit and is a Co-director of Raising Voices and Sophie Namy is the Raising Voices Learning Coordinator. Angel Mirembe and Janet Nakuti are employed in Raising Voices monitoring and evaluation division, but were managed by LSHTM staff during the study. No other author declared competing interests.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. <sup>2</sup>Raising Voices, Kampala, Uganda. <sup>3</sup>Makerere University, Kampala, Uganda.

Received: 3 March 2017 Accepted: 13 April 2018

Published online: 09 May 2018

### References

- Devries KM, et al. School violence, mental health, and educational performance in Uganda. *Pediatrics*. 2014;133(1):e129–37.
- UNICEF, *Violence against Children in Tanzania: Findings from a National Survey, 2009*. Summary report on the prevalence of sexual, physical and emotional violence, context of sexual violence, and health and Behavioural consequences of violence experienced in childhood. . 2011, UNICEF Tanzania, Centers for Disease Control and Prevention, and Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania.
- UNICEF. Violence against children in Kenya: findings from a 2010 National Survey. Nairobi, Kenya: Division of Violence Prevention: National Center for Injury Prevention and Control; 2012.
- Devries KM, et al. The good school toolkit for reducing physical violence from school staff to primary school students: a cluster-randomised controlled trial in Uganda. *Lancet Glob Health*. 2015;3(7):e378–86.
- Michie S, et al. Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implement Sci*. 2009;4:40.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.
- Wierenga D, et al. What is actually measured in process evaluations for worksite health promotion programs: a systematic review. *BMC Public Health*. 2013;13:1190.
- Hargreaves JR, et al. Measuring implementation strength: lessons from the evaluation of public health strategies in low- and middle-income settings. *Measuring implementation strength*. 2016;31(7):860–7.
- Plummer ML, et al. A process evaluation of a school-based adolescent sexual health intervention in rural Tanzania: the MEMA kwa Vijana programme. *Health Educ Res*. 2007;22(4):500–12.
- Bonell C, Jamal F, Harden A, et al. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. Southampton (UK): NIHR Journals Library; 2013. (Public Health Research, No. 1.1.) Chapter 8, Research question 3: process evaluations. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK262778/>.
- Devries KM, et al. The good schools toolkit to prevent violence against children in Ugandan primary schools: study protocol for cluster-randomised controlled trial. *Trials*. 2013;14:232.
- Kyegombe N, et al. How did the good school toolkit reduce the risk of past week physical violence from teachers to students? Qualitative findings on pathways of change in schools in Luwero, Uganda. *Soc Sci Med*. 2017;180:10–9.
- Greco G, Knight L, Ssekadde W, et al. Economic evaluation of the Good School Toolkit: an intervention for reducing violence in primary schools in Uganda. *BMJ Global Health* 2018;3:e000526.
- Child JC, et al. Responding to abuse: Children's experiences of child protection in a central district, Uganda. *Child Abuse & Neglect*, 2014;38(10):1647–58.
- Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot*. 1997;12(1):38–48.
- Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol*. 2008;27(3):379–87.
- Grant A, et al. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials*. 2013;14:15.
- Brown MB. Algorithm AS 116: the tetrachoric correlation and its asymptotic standard error. *Appl Stat*. 1977;26:343–55.
- Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Res Eval*. 2005;10(7). Available online: <http://pareonline.net/getvn.asp?v=10&n=7>.
- Zolotor AJ, et al. ISPCAN Child Abuse Screening Tool Children's Version (ICAST-C): Instrument development and multi-national pilot testing. *Child Abuse Negl*. 2009;33(11):833–41.
- Goodman R, et al. Using the strengths and difficulties questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br J Psychiatry*. 2000;177:534–9.
- Runyan DK, Dunne MP, Zolotor AJ. Introduction to the development of the ISPCAN child abuse screening tools. *Child Abuse Negl*. 2009;33(11):842–5.
- Runyan DK, et al. The development and piloting of the ISPCAN child abuse screening tool—parent version (ICAST-P). *Child Abuse Negl*. 2009;33(11):826–32.
- de Vries PJ, et al. Measuring adolescent mental health around the globe: psychometric properties of the self-report Strengths and Difficulties Questionnaire in South Africa, and comparison with UK, Australian and Chinese data. *Epidemiol Psychiatr Sci*. 2017;1-12.
- Goodman A, Goodman R. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry*. 2009;48(4):400–3.
- Michielsen K, et al. Effectiveness of HIV prevention for youth in sub-Saharan Africa: systematic review and meta-analysis of randomized and nonrandomized trials. *AIDS*. 2010;24(8):1193–202.
- Wight D, Plummer M, Ross D. The need to promote behaviour change at the cultural level: one factor explaining the limited impact of the MEMA kwa Vijana adolescent sexual health intervention in rural Tanzania. *A process evaluation*. *BMC Public Health*. 2012;12:788.
- Murphy JM, et al. Relationship between hunger and psychosocial functioning in low-income American children. *J Am Acad Child Adolesc Psychiatry*. 1998;37(2):163–70.
- Jyoti DF, Frongillo EA, Jones SJ. Food insecurity affects school children's academic performance, weight gain, and social skills. *J Nutr*. 2005;135(12):2831–9.
- Alaimo K, Olson CM, Frongillo EA Jr. Food insufficiency and American school-aged children's cognitive, academic, and psychosocial development. *Pediatrics*. 2001;108(1):44–53.
- Schulte-Körne G. Mental health problems in a school setting in children and adolescents. *Dtsch Arztebl Int*. 2016;113(11):183–90.
- Nelson JR. Designing schools to meet the needs of students who exhibit disruptive behavior. *J Emot Behav Disord*. 1996;4(3):147–61.
- Elbla AIF. Is punishment (corporal or verbal) an effective means of discipline in schools?: case study of two basic schools in greater Khartoum/Sudan. *Procedia - Social and Behavioral Sciences*. 2012;69(Supplement C):1656–63.
- Clarke K, et al. Patterns and predictors of violence against children in Uganda: a latent class analysis. *BMJ Open*. 2016;6(5):e010443.
- Nalugya-Sserunjogi J, et al. Prevalence and factors associated with depression symptoms among school-going adolescents in Central Uganda. *Child Adolesc Psychiatry Ment Health*. 2016;10:39.
- Devries KM, et al. Witnessing intimate partner violence and child maltreatment in Ugandan children: a cross-sectional survey. *BMJ Open*. 2017;7(2):e013583.
- Panovska-Griffiths J, et al. Optimal allocation of resources in female sex worker targeted HIV prevention interventions: model insights from Avahan in South India. *PLoS One*. 2014;9(10):e107066.