Social constructs, behaviour change, and the uptake of community-based WASH interventions:

Metrics and analytical approaches for measuring collective efficacy

Maryann Greene Delea

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of London

2019

Department of Disease Control

Faculty of Infectious & Tropical Diseases

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by grants from the Bill & Melinda Gates Foundation and The World Bank Group's Strategic Impact Evaluation Fund

DECLARATION

I, Maryann Greene Delea, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that it has been indicated in the thesis.

Maryann Greene Delea

October 2018

ABSTRACT

It has become commonplace in international development to intervene upon communities with interventions that require collective action without first gauging the communities' perceptions regarding their ability and autonomy to engender and maintain change. Programmes and research studies employing community-based interventions often overlook important interpersonal behavioural factors that may well affect uptake and effectiveness. Social constructs such as collective efficacy, social capital, social cohesion, and social norms are important interpersonal behavioural factors and predictors of collective action, which may be needed to realise downstream health and development impacts.

This doctoral thesis examined the conceptualisations of various social constructs and their theoretical underpinnings. Theoretical examinations were used to establish hypotheses regarding the underlying structure of collective efficacy (CE). A factor analytic approach was used to develop four CE measurement scales from data collected in Odisha, India and Amhara, Ethiopia to empirically test these hypotheses. The India CE scale was subsequently used to ascertain whether there were independent associations between CE-factors and the uptake and influence of a community-based water supply and sanitation intervention.

Compared to controls, intervention households had higher CE factor scores for village leadership (β=0.16, 95% CI: 0.08, 0.25) and agency (β=0.08, 95% CI: 0.01, 0.15), and lower scores for social disorder, conditional on education. Prevalence of improved WASH behaviours was associated with CE factors: improved water piped on-premise was associated with village leadership (aPR=2.25, 95% CI: 1.12, 4.53); improved on-site toilets with social disorder (aPR=0.90, 95% CI: 0.84, 0.97); enclosed bathing rooms with social response (aPR=1.12, 95% CI: 1.02, 1.23); and utilisation of improved sanitation facilities by all family members with agency (aPR=1.17, 95% CI: 1.00, 1.37) and village leadership (aPR=3.86, 95% CI: 1.67, 8.97). Agency, social response, and social disorder factors were associated with nutritional outcomes. Implications for enhanced intervention design, targeting, and evaluation are discussed throughout.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank the numerous individuals who have invested in me – whether through a contribution of time, energy, knowledge, love, companionship, or laughter – your support did not go unnoticed, and is much appreciated.

First, I would like to extend my sincere gratitude to my supervisor, Tom Clasen, who provided me the opportunity to pursue my doctorate, and gave me the support and space to investigate something I thought was important. Thank you for guiding me along the way, and being patient with me throughout the process. Thank you for having the confidence in me to complete this journey.

I would also like to thank Matt Freeman for your flexibility and guidance over the last three and a half years. It has been a pleasure working with you to get the *Andilaye* trial off the ground and exploring some unexpected areas of interest. Thank you Corey Nagel for the time you spent discussing my analysis plans and models with me. I appreciate your insights and guidance.

Thank you Gloria Sclar for being a great collaborator, travel companion, and friend. Thank you for teaching me the ins and outs of both rigorous and practical qualitative research. I learnt quite a lot from you along the way. Thank you for your intellectual contributions to the hypothesised collective efficacy framework, suggested refinements to the India and Ethiopia collective efficacy surveys, and reflections on the results and write-up of the two related papers.

Thank you to the co-authors of the papers presented herein, for your important contributions. Thank you to Bethany Caruso for your invaluable thoughts and guidance regarding the formative research, survey design and administration, factor analyses, and Research Paper 2 revisions. Thank you Regine Haardörfer for your insights into psychometric analysis options and suggestions regarding the gendered approach to scale development. Thank you Mulat Woreta and Abebe Gebremariam for your insights into the local context, and Robert Dreibelbis for your suggestions about how to effectively frame Research Paper 2.

Thank you to my colleagues, friends, and family for your unyielding support and encouragement. Thank you Jed Snyder for your contributions coordinating the *Andilaye* trial, making us all look good with your amazing graphical design skills, and carrying the project on through the finish line. Thank you to everyone in room 411 – Helen Buxton, Alex Czerniewska, Jo Esteves-Mills, Lauren D'Mello-Guyett, Aurelie Jeandron, and Robert Dreibelbis; and other members of the Environmental Health Group, particularly Jessica Petz and Sian White – for your friendship, patience, and encouragement. Special thanks to Fiona Majorin and Miles Kirby for welcoming me to London, being great friends over the years, and providing insight and guidance into the last phases of the doctorate.

Thank you to the study teams, particularly Mulat Woreta, Resom Berhe, and Kassahun Zewudie in Ethiopia, and Parimita Routray, Munmun Dasmohapatra, Alfred Mohanty, and Manaswani Rout in India. Thank you to the enumeration teams in both countries for your efforts and energetic approach to this work.

I would also like to thank the communities – too many to name – in which I have worked, particularly those from the two study sites highlighted in this thesis, but also those from the ten other countries in which I have worked over the past 12 years. These communities and their members have taught me – as an outsider, and in some cases, a participant observer – the true meaning of the themes covered in this thesis: community, togetherness, sharing, cooperation, and collective action. It does, indeed, take a village. The experiences I have had working and living in these communities no doubt enriched this intellectual inquiry, enhanced my ability to reflect on the data in a meaningful way, and informed the direction of the studies and the interpretation and the data.

Last, but certainly not least, I am extremely grateful and deeply indebted to my family – my sisters Deirdre and Kristin Delea, and my mother Nancy Delea – for your steadfast support, consistent encouragement, and humour. Without you, this pursuit would not have been possible. Thank you to my dad, John Delea. While our time together on this Earth was far too short, you left an indelible mark on my life, teaching me that education is a priority, hard work pays off, and persistence is key. Most importantly, you impressed upon me the importance of integrity, resilience, and being kind and generous even in the face of adversity. I have felt your presence with me during the most difficult parts of this journey.

TABLE OF CONTENTS

TABLE OF FIGURES (exclusive of figures within research manuscripts/papers)

TABLE OF TABLES (exclusive of tables within research manuscripts/papers)

ACRONYMS

ARHB	Amhara Regional Health Bureau
CE	Collective efficacy
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CI	Confidence interval
CLTSH	Community-led total sanitation and hygiene
CTT	Classical test theory
DIF	Differential item functioning
EFA	Exploratory factor analysis
FMoH	Federal Democratic Republic of Ethiopia's Federal Ministry of Health
HAZ	Height-for-age z-score
HEP	Health Extension Programme
HEW	Health extension worker
IRT	Item reduction theory
MANTRA	Movement for Action Network for Transformation of Rural Areas
MIMIC	Multiple indicators multiple
NFHS	National Family Health Survey
NGO	Non-governmental organisation
NTDs	Neglected tropical diseases
PCA	Principal components analysis
PR	Prevalence ratio
RCT	Randomised, controlled trial
RMSEA	Root mean square error of approximation
RMSR	Root mean square residual
RQ	Research question
STH	Soil-transmitted helminthiasis
TLI	Tucker-Lewis index
WASH	Water, sanitation, and hygiene
WAZ	Weight-for-age z-score
WHZ	Weight-for-height z-score
WLSMV	Weighted least-squares with mean and variance adjustments

PROLOGUE

Chapter 1: Introduction and background

1.1 Thesis introduction

This doctoral thesis examines the conceptualisation and attributes of several social constructs that social, behavioural, and behaviour change theories suggest are antecedents of cooperative behaviour and collective action. While various social constructs are considered, the thesis focuses on collective efficacy (CE), and builds a case for its utility in examinations of behavioural factors influencing the uptake of community-based interventions predicated on collective action.

Theoretical underpinnings of social constructs were examined in order to inform implementation science research that was conducted with the intension of providing recommendations for improving public health and development practice. While much of the work presented herein is specific to water, sanitation, and hygiene (WASH) interventions, related findings can be extended to any community-based programme or intervention that requires cooperative behaviour or collective action. This thesis research was conducted under the auspices of two separate evaluations of WASH interventions in India and Ethiopia. Findings from these studies have important implications for programme and policy, and recommendations regarding intelligent intervention design, targeting, and evaluation are presented.

1.2 Chapter overview

This chapter presents general information related to the origin, rationale, aims, and objectives of this thesis research. The chapter also provides information regarding the candidate's contributions. The chapter ends with an overview of the content of subsequent sections and chapters.

1.3 Origin of doctoral thesis research and problem addressed

In early 2015, my supervisor approached me about a study he was planning in India that would evaluate the effectiveness of a promising intervention that leverages a communal good – a

village-level water distribution system that is piped into each household via three taps – to ensure village-wide saturation of improved, on-site household sanitation facilities coverage and utilisation. This programming approach was unique in that it focused on collective action, and seemed to be fully interrupting the setting in which open defecation typically occurs in India – off-site water sources that allow for post-defecation ritual cleansing and menstrual hygiene management. Given the implementing organisation's village engagement and enrollment approach (outlined in Chapters 3 and 5), he and the programme administrators were interested in trying to assess social cohesion in programme villages, and investigating whether this might influence intervention effectiveness.

1.3.1 What: Selecting an appropriate social construct for the task at hand

I spent the subsequent months exploring various social constructs and building a case for how collective efficacy would be appropriate to investigate given the nature of the construct and how ubiquitous it is in social psychology and behaviour change theory. I felt strongly that it was important to focus the investigation on CE as opposed to other social constructs given the theoretical and empirical evidence base that indicates CE shapes a group's decision to set and pursue common goals [1, 2], which influences performance and collective action. Evidence also suggests that this social construct is predictive of group performance, including the amount of effort a group or collective will spend working toward communal goals and the level of persistence they will expend in doing so when group efforts fail to produce the desired results [2, 3]. Conscious goals are important to consider when assessing factors influencing collective action given they are known to affect action [4-6]. Not incorporating CE in examinations of collective action could therefore result in the oversight of these factors and their influence on targeted behavioural outcomes. While this thesis focuses on CE, it considers concepts related to social capital and social cohesion as related influencers of collective action and cooperative behaviour.

Social cohesion, while a seemingly important factor, may not necessarily translate to collective action or cooperative behaviour, especially if other psychosocial, contextual, and technological factors prevent a collective (e.g., a village, community group) from setting or pursuing communal goals. Similarly, while social capital plays an important role in behavioural control perceptions, examining cognitive and structural aspects of capital without considering aspects of informal social control and social cohesion may result in an incomplete assessment of the behavioural antecedents of collective action. As indicated in relevant literature, focusing on

social cohesion or social capital alone may provide necessary yet insufficient examination of the mechanisms through which cooperative behaviour and collective action are facilitated [7, 8].

1.3.2 How: Measuring latent social constructs: Positivist versus interpretivist perspectives

It is not possible to observe or measure CE directly because it is a latent construct [9]. A positivist approach, however, posits that certain latent aspects of society, such as CE, can be studied by leveraging scientific evidence to reveal the nature of how things work [10]. Such an approach utilises scientific theory and methods to produce quantitative measures of latent constructs in a manner that allows researchers to explore the relationships between those latent constructs and their constituent components (i.e., sub-constructs – e.g., construct domains, dimensions, facets). With the proper study design, a positivist approach also permits the examination of the causal relationships between latent constructs and important behavioural outcomes and health impacts.

Some, particularly interpretivists, may argue that the scientific method cannot generate an accurate portrayal of the social realm. Here, I will neither argue against the benefits, nor highlight the limitations of pursuing other, non-positivist epistemological approaches to examining social constructs [11]. However, I suggest that through the use of widely accepted scientific and analytical methodologies, a positivist approach that incorporates formative qualitative research activities can be employed to generate meaningful information that advances general understanding about CE and its constituent sub-constructs. Such information can be used to further inform programme and policy.

1.3.3 Why: Assessing relationships between social constructs, behavioural outcomes, downstream health & development impacts

To date, the public health community in general, and WASH sector more specifically have not readily incorporated assessments of CE into the design, targeting, and evaluation of interventions [12]. From a theoretical standpoint, and as evidenced in other sectors, CE is an important antecedent of collective action [13]. Therefore, it was hypothesised that overlooking or underestimating CE's role as a factor of collective uptake of community-based interventions may result in the attenuation of behavioural adoption and downstream health and development impacts. The work presented in this thesis tests this and other related hypotheses. Findings presented herein may help address knowledge gaps and provide information that can be used to improve community-based programmes.

1.3.4 So what: Implications for the design, targeting, and evaluation of community-based programmes

The overarching goal of this doctoral research was to produce evidence-based recommendations to strengthen community-based programmes. As a result, the findings reflected in this thesis highlight implications for the design, targeting, and evaluation of community-based programme interventions, particularly those that are predicated on collective action.

1.4 Candidate's contributions to the thesis research and how they advance knowledge of the subject

As is typical with much public health research, this thesis research involved a team of individuals contributing toward its overarching goal. This thesis was supervised by Dr. Thomas Clasen at LSHTM. Dr. Matthew Freeman from Emory University is the Principal Investigator of the Ethiopia trial presented herein. He provided additional supervisory support. Funding for the research was provided by the Bill & Melinda Gates Foundation and The World Bank Group's Strategic Impact Evaluation Fund. My colleague, Gloria Sclar, a Public Health Research Associate at Emory University, provided technical assistance on the qualitative aspects of the formative work in India and Ethiopia, and the design of the intervention developed for the cluster-randomised control trial (RCT) in Ethiopia. She also provided her own contributions to the development of our hypothesised CE framework – this was a joint effort. The genesis of the RCT's intervention was a collaborative effort, with important contributions provided by Siraj Mohammad, Mulat Woreta, Kassahun Zewudie, and Resom Berhe from the Emory Ethiopia Bahir Dar regional office. Efforts provided by our cadre of facilitators, field supervisors, and data collection team were critical to the execution of this work. Given my involvement in the design, management, and writing up of the study proposal, field protocols, baseline and midline reports, intervention design content, and other programme-related documents, some of the text I present herein is reflected in related study documents.

While this work reflects a collaborative effort from many individuals, I led the intellectual inquiry, from determining which social construct should be investigated and managing the technical direction of the formative research to the design of the two collective efficacy studies highlighted herein, and the analysis of the resulting data. I conceived of the research, created the research questions related to each inquiry, developed all field protocols and tools, and trained enumerators. Feedback was sought and received from other team members, and final

versions of protocols and study tools reflect inputs from several individuals. I managed and led the collective efficacy studies.

Aside from designing and managing the collective efficacy and costing sub-studies in India, I did not contribute to the overarching study design or household survey instruments related to the matched cohort study. However, I was heavily involved in the design and management of the cluster-randomised control trial in Ethiopia. The Principal Investigator and I designed the RCT, I developed all field protocols, and worked with the above-mentioned team members to design the intervention being evaluated in the RCT.

This thesis research contributes to existing knowledge in several ways. First, Research Paper 1 provides information and recommendations for the selection of appropriate social constructs, and their incorporation into the design, targeting, and evaluation of community-based interventions. Second, only a few psychometric analyses of CE have been carried out, and to my knowledge, none of these has used data obtained from low literate, rural populations from low- and middle-income countries. In addition to generating four CE scales (three in Ethiopia, one in India), our full CE surveys will be published. Any of these tools can be adapted and deployed for future assessments of CE. Finally, no known studies have thoroughly examined independent associations between empirically-derived CE factors and the uptake and influence of WASH interventions. Discussion of the implications of the results of this thesis work for improved design, targeting, and evaluation of community-based interventions is provided throughout.

I am the first author of the three research papers presented in this thesis. One paper has been published in a peer-reviewed journal; the other two papers have been prepared for submission to peer-reviewed journals. A cover sheet that includes details regarding the role I played in the development and execution of the work proceeds each research paper included herein.

1.5 Thesis purpose, aims, and objectives

The purpose of this doctoral thesis was to further elucidate how, why, and when various social constructs should be assessed to improve the design, targeting, and evaluation of community-based interventions. The thesis aimed to specifically examine collective efficacy, and determine whether evidence suggests it is associated with collective WASH behaviours and improved WASH practices. This body of work encompassed several specific objectives, each with its own set of research questions. To facilitate navigation through this thesis, Table 1.1 summarises each

research objective, along with the related research questions, and indicates where in the thesis these items are addressed.

Table 1.1 Thesis objectives and research questions

Objective	Related research questions (RQs)	Thesis chapters
To examine conceptualisations of various social constructs and their underpinnings in behaviour and behaviour change theory to elucidate when and why various social constructs should be assessed to inform the design, targeting, and evaluation of community-based interventions	**RQ1.** How are various social constructs conceptualised, what are the underlying theoretical foundations of each, and how do they relate to proposed causal mechanisms of behaviour and behaviour change? **RQ2.** Which social constructs should be employed to assess the effect of socially-influenced change mechanisms on the uptake, influence, and sustainability of community-based interventions? **RQ3.** What are the advantages and limitations of various social constructs?	**Chapter 2:** Research Paper 1
To elucidate the underlying structure of CE through the development, refinement, and validation of CE measurement metrics	**RQ1.** Which sub-constructs (e.g., constituent domains, factors/dimensions, and facets) are salient for measuring CE? **RQ2.** Are the psychometric characteristics of the resulting CE measurement models compelling in terms of their ability to demonstrate construct validity, or the degree to which the scale measures what it purports to measure? **RQ3.** Are there important differences in the measurement of CE between men and women; individuals/households with leadership status vs. those without?	**Chapter 4:** Research Paper2 - Ethiopia scales **Chapter 5:** Research Paper 3 - India scale
To investigate the relationship between CE and intervention uptake, influence, and sustainability of improved WASH behaviours	**RQ1-3.** (see RQ1-3 above) **RQ4.** Is there evidence of an association between CE factor scores and uptake of a community-based water supply and sanitation intervention (various indicators along the causal chain – behavioural outcomes, child nutritional status)?	**Chapter 5:** Research Paper 3

1.6 Thesis structure

This thesis contains seven chapters organised into several sections. The document commences with a Prologue that presents an introduction to the thesis and provides background information. The main body of the thesis is divided into three sub-sections, with research findings presented as three separate research papers. Discussion, reflections, and conclusions related to the overarching thesis is provided in the Epilogue. A summary of the specific content of each section is provided below.

PROLOGUE

Front content related to the thesis is presented in the Prologue. This section is comprised of **Chapter 1**, which introduces the thesis and provides background information. The origin of the thesis and its specific aims and objectives are presented herein.

PART I: SOCIALLY-INFLUENCED BEHAVIOURAL ANTECEDENT CONSTRUCTS

The first part of the thesis provides an overview of various socially-influenced latent constructs that represent antecedents of improved behaviours. **Chapter 2** provides a theoretical overview of social constructs as mechanisms of behaviour change. This chapter lays the foundation for the social constructs that are discussed throughout the thesis, and highlights knowledge and evidence gaps related to collective efficacy. This chapter **contains the manuscript for Research Paper 1**, which synthesises information from behaviour and behaviour change theory to summarise attributes related to various social constructs. A case is made for better incorporation of collective efficacy measurement to inform the design, targeting, and evaluation of community-based programmes.

PART II: PSYCHOMETRIC & COGNTIVE VALIDATION OF COLLECTIVE EFFICACY SCALES

The second part of the thesis focuses on the development, refinement, and validation of locally adapted collective efficacy scales. **Chapter 3** provides a summary of the research contexts and study designs of the overarching, anchor studies in which this collective efficacy research was conducted. This chapter also summarises how an iterative approach was used, to leverage learning from the first collective efficacy inquiry carried out in India to inform further refinements for similar work carried out in Ethiopia. **Chapter 4** provides a summary of the five-step scale development, refinement, and validation process that was used to construct all collective efficacy measurement scales, and provides details and justification for the methodological decisions made in the conduct of the psychometric analyses. This chapter **contains Research Paper 2**, which provides a detailed account of collective efficacy scale development, highlighting work conducted in Ethiopia.

PART III: COLLECTIVE EFFICACY & THE UPTAKE & INFLUENCE OF A COMMUNITY-BASED WASH PROGRAMME

The third part of the thesis demonstrates how the collective efficacy scale developed for the rural Odisha, India context was employed to examine evidence of independent associations between collective efficacy-related factors and behavioural and nutritional outcomes. **Chapter 5** provides information related to behavioural and nutritional outcomes observed through the

evaluation of the intervention. This chapter **contains the manuscript for Research Paper 3**, which presents results regarding evidence of associations between collective efficacy factors, and the uptake of a community-based water supply and sanitation intervention, and improved WASH behaviours more generally.

EPILOGUE

The thesis concludes with a reflection of the results emerging from this body of work. **Chapter 6** contains a discussion that explores how these findings corroborate existing evidence, yet provide unique contributions challenging the progression of the social constructs agenda. This discussion focuses heavily on programme implications related to the employment of collective efficacy measurement to inform the design, targeting, and evaluation of community-based interventions, especially those predicated on collective action. This chapter also reviews the limitations of the research. **Chapter 7** contains summarised conclusions and additional recommendations for research and practice.

1.7 Chapter references

1. Bandura, A., *Exercise of Human Agency through Collective Efficacy.* Current Directions in Psychological Science, 2000. **9**(3): p. 75-78.
2. Bandura, A., *Self-efficacy mechanism in human agency.* American Psychologist, 1982. **37**(2): p. 122-147.
3. Durham, C.C., D. Knight, and E.A. Locke, *Effects of Leader Role, Team-Set Goal Difficulty, Efficacy, and Tactics on Team Effectiveness.* Organizational Behavior and Human Decision Processes, 1997. **72**(2): p. 203-231.
4. Ryan, R.M. and E.L. Deci, *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.* American Psychologist, 2000. **55**(1): p. 68-78.
5. Bandura, A. and E.A. Locke, *Negative self-efficacy and goal effects revisited.* J Appl Psychol, 2003. **88**(1): p. 87-99.
6. Locke, E.A. and G.P. Latham, *Building a practically useful theory of goal setting and task motivation: A 35-year odyssey.* American Psychologist, 2002. **57**(9): p. 705-717.
7. Ostrom, E., *Social capital : a fad or a fundamental concept?*, in *Social capital : a multifaceted perspective.*, P. Dasgupta and I. Serageldin, Editors. 2000, World Bank: Washington, D.C. p. 172-214.
8. Bisung, E., S.J. Elliott, C.J. Schuster-Wallace, D.M. Karanja, and A. Bernard, *Social capital, collective action and access to water in rural Kenya.* Social Science & Medicine, 2014. **119**: p. 147-154.
9. Rabe-Hesketh, S. and A. Skrondal, *Classical latent variable models for medical research.* Stat Methods Med Res, 2008. **17**(1): p. 5-32.
10. Berger, P.L. and T. Luckmann, *The social construction of reality : a treatise in the sociology of knowledge*. 1966, London, England: Penguin Books.
11. Achinstein, P. and S.F. Barker, *The Legacy of logical positivism : studies in the philosophy of science*. 1969, Baltimore: Johns Hopkins Press.

12. Glanz, K. and D.B. Bishop, *The Role of Behavioral Science Theory in Development and Implementation of Public Health Interventions.* Annual Review of Public Health, 2010. **31**(1): p. 399-418.
13. Bray, S.R., *Collective Efficacy, Group Goals, and Group Performance of a Muscular Endurance Task.* Small Group Research, 2004. **35**(2): p. 230-238.

PART I: SOCIALLY-INFLUENCED BEHAVIOURAL ANTECEDENT CONSTRUCTS

Chapter 2: Summary of social constructs: Implications for public health and development practice

2.1 Chapter overview

In order to determine which social constructs should be considered and employed for intervention design, targeting, and evaluation, it is important to understand the attributes of each. This chapter examines the conceptualisations of various social constructs, their unique attributes, and their underpinnings in behaviour and behaviour change theory.

2.2 Framing of research questions

To date, the WASH sector continues to overlook some key psychosocial influencers of behavioural change and maintenance by largely examining and addressing individual and household-level behavioural factors. Community-level, or interpersonal factors may be stronger influencers of action at collective levels, the levels at which change may be required to realise health gains [1-3]. Therefore, Research Paper 1 explores the nature of various social constructs through the mapping of each construct's respective conceptualisation, including its domains and dimensions. This was done in order to identify gaps, outline construct attributes, and make recommendations for which constructs are the most appropriate to employ when evaluating the effectiveness of community-based interventions and the performance of development programmes.

2.2.1 Research Paper 1 objective

The objective of this research was to further elucidate how, why, and when various social construct metrics should be employed to evaluate community-based interventions. Related findings will provide information regarding construct characteristics that can guide programme implementers and researchers in selecting appropriate constructs for intervention design and evaluation.

2.2.2 Research questions

The following research questions were investigated:

RQ1. How are various social constructs conceptualised, what are the underlying theoretical foundations of each, and how do they relate to proposed causal mechanisms of behaviour and behaviour change?

RQ2. Which social constructs should be employed to assess the effect of socially-influenced change mechanisms on uptake, effectiveness, and sustainability of community-based interventions?

RQ3. What are the advantages, challenges, and limitations of various social constructs?

2.3 Collective behaviours and practices: Cooperative behaviour and collective action

Throughout this thesis, collective behaviours and practices such as collective action are noted as key behavioural outcomes of interest. These behavioural outcomes reflect interactions between individuals seeking to achieve a communal goal or common-pool resource, and are therefore relevant to an array of public health and development programming. Collective action is necessary for the uptake of different types of community-based interventions, from those promoting disaster risk reduction [4] to mass drug administration, installation and maintenance of communal water points [5], and improved sanitation practices [6].

Coordinated efforts, such as those reflected through cooperative behaviour and collective action, require a group to work to achieve or obtain a communal goal or resource. Therefore, it is important to consider the myriad factors contributing toward these behavioural outcomes, including perceptions regarding the group's perceived ability and autonomy to set, pursue, and achieve the communal goal or obtain and maintain the common-pool resource (i.e., collective efficacy).

2.4 Additional notes regarding social, behavioural, & behaviour change theories

There are several different types of social, behavioural, and behaviour change theories represented in the literature. One dichotomy worth noting is that of explanatory, or behavioural versus change theories. Explanatory or behavioural theories are often linear in nature, and as their name indicates, they serve to explain, through numerous influencing factors or predictors, why behaviours occur [7]. Conversely, change theories tend to be more cyclical, as they portray dynamic processes of change, which may reflect intricate interactions between various influencing factors [7, 8]. It is important to understand the distinction between explanatory and change theories in order to know

which type is appropriate for use, given the task at hand. For example, an explanatory theory is useful and appropriate for helping to synthesise findings from formative work, to identify barriers or facilitators of change. However, these theories would not be appropriate, on their own, to inform the design an intervention that seeks to both identify mechanisms of action *and* address antecedents of behaviour.

Another dichotomy of note is that of stage theories versus predictive theories. The central focus of stage theories is the stages through which an individual or group must move, over time, in order for change to occur. Predictive theories focus on identifying factors and causal pathways, irrespective of time, that influence the likelihood of performance or non-performance of a given behaviour or practice [9]. Critiques of both types of theories have been put forward. Predictive theories are more comprehensive than stage theories with regard to the various factors, determinants, and pathways of change that are considered, and some perceive them as more transferrable to and meaningful for intervention design and evaluation [9]. While these types of theories represent two different schools of thought, integrated theories that consider aspects related to both do exist.

Research Paper 1 does not discriminate against any specific type of theory, as I propose they operate in concert with, rather than independently of each other. Given the theoretical nuances and potential implications for intervention design and evaluation, type of theory was examined rather than restricted.

2.5 Collective efficacy and other social constructs as interpersonal factors influencing collective behaviours

2.5.1 Preamble to Research Paper 1

Research Paper 1 examines pertinent socially-influenced constructs that are predictors of collective action. Conceptualisations of these constructs are summarised, and the theoretical foundations of each are highlighted. Resulting information is synthesised to describe how each construct relates to proposed causal mechanisms of behaviour and behavioural change. Recommendations are set forth for the incorporation of social constructs in public health and development work that employs community-based interventions, as indicated by a synthesis of the constructs' attributes.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Maryann Delea
Principal Supervisor	Dr. Thomas Clasen
Thesis Title	Social constructs, behaviour change, and the uptake of community-based WASH interventions: Metrics and analytical approaches for measuring collective efficacy and other social construct

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Social Science and Medicine
Please list the paper's authors in the intended authorship order:	Maryann G. Delea, Thomas F. Clasen
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I synthesised information regarding the conceptualisations of each social construct presented in this paper to develop a social constructs framework, and provide guidance for the inclusion of social constructs. I wrote the first draft of the paper.

Student Signature: [redacted] Date: 16/10/2018

Supervisor Signature: [redacted] Date: 10-18-2018

2.6 Research Paper 1: Using a social constructs framework to improve the intervention design, targeting, and evaluation of community-based public health and development programmes

Maryann G. Delea [1,2], Thomas F. Clasen [1,2]

[1] Department of Disease Control, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom;

[2] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States of America

Status: Prepared for submission to *Social Science and Medicine*

ABSTRACT

Trials of certain community-based interventions, particularly those related to water, air quality, sanitation, and hygiene (WASH), have failed to demonstrate positive health and development gains. However, most of these studies have not incorporated certain social constructs into the design, targeting, or diagnostic assessments of the interventions under evaluation. When they have been incorporated, the scope of the selected constructs has been limited, and has not always aligned well with the attributes requiring examination. Improving upon the selection and inclusion of social constructs into community-based programmes and interventions will allow for further examination of the specific pathways that either drive or impede behaviour change and sustained adoption of improved behaviours and practices. This article summarises concepts related to several social constructs commonly identified as theoretically-grounded and evidence-based predictors of behaviour and behaviour change, and provides guidance for their inclusion in the design, targeting, and evaluation of community-based behaviour interventions. Constructs such as collective efficacy, social capital, social cohesion, and social norms are examined. An emphasis was placed on the discussion of efficacy-related constructs, as efficacy appraisals are major determinants of goal setting, task choice, and willingness to expend effort and resolve during task or goal pursuit. Synthesised findings are presented in a social constructs framework that highlights each construct's unique attributes and contribution, along with information regarding when and why the constructs

should be considered. Nuances related to motivational concepts of efficacy, behavioural control, locus of control, and agency are explored, and a case is made for better integration of collective efficacy as a behavioural antecedent of collective action to be considered. While the findings of this paper can be used to enhance community-based development programmes, at large, we examine these social constructs and their constituent components through a public health lens, with a specific focus on community-based WASH interventions.

1. Introduction

Impact evaluations of some community-based interventions have yielded lower than expected behavioural outcomes and health and development gains [10, 11]. However, when interpreting these findings, it is important to note the intervention content, implementation approaches, and intervention techniques that were employed, along with how the intervention was targeted and evaluated. Programmes and research studies employing these community-based interventions often overlook important interpersonal behavioural factors that may well affect the uptake and effectiveness of such interventions. Social constructs such as collective efficacy, social capital, social cohesion, and social norms are important factors contributing to collective action [12-14], which is often needed to bring about downstream health and development impacts. Underestimating or overlooking these behavioural factors promotes a “black box” mentality that supports oversimplified conceptualisations of causal chains. Failing to unpack these issues and collect data on critical mechanisms through which sustained behavioural adoption occurs prevents a more thorough examination of the behavioural antecedents of cooperative behaviour and collective action. Such investigations may be needed to help explain why community-based interventions are not bringing about desired, yet plausible health and development impacts.

Within the water, sanitation, and hygiene (WASH) sector specifically, the scope and extent to which social constructs have been incorporated into programmes design and evaluation is somewhat limited. For instance, social capital is often employed exclusively to examine interpersonal factors contributing to collective action [5, 15, 16]. This occurs despite conclusions that have been made and widely disseminated by theorists and practitioners that social capital is necessary, but insufficient for investigating factors driving collective action [15, 17].

Collective efficacy, social cohesion, social capital, and social norms are all latent social constructs hypothesised to influence the uptake, effectiveness, and sustainability of interventions, particularly those contingent upon collective action. Empirical evidence and behavioural theories support the plausibility of these hypotheses. For example, evidence suggests that communities high in social construct measures, such as social capital, witness higher uptake of WASH interventions, and consequently, related health benefits [5, 6, 15, 18]. Evidence also suggests that the influence of collective action is far-reaching, as it enables communities to achieve not only sanitation-related goals, but also attainment of other basic services [19]. However, a more nuanced understanding of 'community' and the role various social constructs play in the success of community-based interventions is needed [20] in order to determine causal mechanisms of improved WASH behaviours and related health impacts.

In the behavioural sciences literature, there are differing views regarding both the conceptualisations of social constructs (i.e., phenomena used to describe and predict them), and how these constructs and constituent sub-constructs are related. There are also differing views regarding when and how to employ social constructs to examine factors related to cooperative behaviour and collective action. Regardless of how social constructs are defined, the idea that they mediate and moderate collective action and the success of community-based interventions is supported by theory and evidence [12, 14, 15], and is broadly acknowledged in practice. However, sound guidance regarding the appropriate employment of these social constructs in the context of community-based programming is deficient. Information that is currently available often focuses on the use of specific constructs, across all contexts, without regard to construct-specific attributes and underlying causal mechanisms suggested by behavioural change theories and empirical testing thereof. Selecting appropriate social constructs that match the specific need of a programme or intervention will allow for further examination of the specific interpersonal factors that either drive or impede behaviour change and sustained adoption of improved behaviours and practices. More nuanced examinations into these specific factors that serve to catalyse and maintain change can help pinpoint the pathways through which change either does or does not occur. Identifying the variety of behavioural factors, interpersonal or otherwise, that prohibit change from occurring may help to explain why current interventions, such as those trialled in the WASH sector have failed to yield sustained behaviour change and biologically plausible health improvements.

Evidence suggests that interventions developed with an explicit theoretical foundation are more effective than those that are not [7, 21]. While progress has been made in recent years to improve the incorporation of theory into intervention design and implementation, there remain opportunities for improvement [21]. One such opportunity is better inclusion of intervention content and implementation strategies that leverage theoretical considerations to address higher-order behavioural factors such as those related to interpersonal, organisational, community, and environmental influences. Interventions that focus exclusively on individual-level behavioural factors fail to fully incorporate construct theories that build on intrapersonal factors (e.g., personality traits, affective states that promote internal motivation, beliefs about personal abilities, coping skills, self-esteem) yet also consider the influence of the individual's social settings (e.g., affective states that promote collective motivations, beliefs about collective abilities, normative nature and pressures of familial ties, referent or social networks) [22].

When designing, targeting, and implementing theoretically-informed interventions, particularly community-based interventions predicated on collective action, it is important to consider various social constructs, especially those pervasive across behaviour change theories. Acknowledging and understanding the inherent characteristics of, and the nuances between various constructs can help ensure interventions incorporate the most appropriate constructs to address the behavioural factors needed to achieve the programme's goal. To date, many community-based interventions have focused on the normative cognition realm, or the range of constructs related to the norms, values, attitudes, and beliefs that create an "affective culture" of trust and solidarity that fosters cooperation and collective action [23]. However, many of these interventions tend to under-utilise constructs related to the instrumental cognition realm, or those constructs that create an "effective culture, with shared confidence in the methods and feasibility of cooperative or collective undertakings" (e.g., efficacy-related constructs) [23]. Interventions intended to yield cooperative behaviour or collective action should consider not only factors related to the normative cognition realm, but also those related to the instrumental cognition realm.

The aim of this paper was to further elucidate when and why various social constructs should be employed for the design, targeting, and evaluation of theoretically-grounded community-based interventions, particularly those predicated on collective action. We sought to provide information related to the specific conceptualisations between various social constructs and the nuances thereof. In order to draw attention to collective efficacy as a neglected social construct, and to

demonstrate its potential utility, we summarise and synthesise the relevant behaviour and behaviour change literature. We highlight the various attributes of each construct, and provide guidance on when and why these constructs should be considered for inclusion in community-based programmes. This guidance is intended to help programme implementers and researchers select the most appropriate constructs to include in the design, targeting and evaluation of their community-based programmes. Finally, we examine and summarise the nature of and nuanced differences between various social constructs to further guide programme implementers and researchers in selecting appropriate constructs.

2. Cooperative behaviour and collective action

Cooperative behaviour and collective action are behavioural outcomes produced by various antecedents, or precursors of behaviour. According to collective action theory, public goods or common resources yield collective benefits, but cooperative behaviour of a group's individuals is influenced by factors such as the size of the group and the tendency for some members to "free-ride" on efforts provided by others [24]. In other words, cooperative behaviour is influenced by behavioural antecedents such as social norms (e.g., empirical and normative expectations regarding cooperation), intra-group normative control (e.g., negative sanctions for violators, positive sanctions for compliers and the strength and monitoring thereof), efficacy-related factors, and the group's dependence on the common good [13, 24, 25].

While experimental testing of this theory has revealed that collective norms and values are modifiers of cooperation toward collective undertakings, evidence suggests that other factors, several of which tap to self- and collective efficacy, influence the rate of contribution toward the common goal as well [17]. Such factors include a common understanding amongst the group (e.g., clear goals regarding the development and management of the common good), past experience or group performance, and presence of leadership [12, 14]. Additional theoretical and empirical evidence suggests that factors related to the collective's socio-structural (e.g., social networks and social support structures) and cognitive features (i.e., perceptions of trust, norms of reciprocity, and values that operate within those structures), and cohesion facilitate group coordination and cooperation for mutual benefit [26, 27].

3. Behavioural antecedents of collective action: Key social constructs and their conceptualisations

3.1 Efficacy

In general, efficacy pertains to perceptions regarding the ability and autonomy to pursue a task or communal goal (respectively for self- and collective efficacy), performance during task/goal pursuit, and the amount of effort and persistence that will be expended [12, 28]. Albert Bandura popularised efficacy through the development of his self-efficacy theory in the late 1970s. Bandura drew on the prevailing psychological theories of his time, such as Rotter's theory of personality [29], to develop this self-efficacy theory [28]. While he first developed and detailed efficacy at an individual level, he went on to extend the concepts of his self-efficacy theory to higher ecological levels in his examinations of collective efficacy [12]. To date, self-efficacy remains more prevalent in behaviour and behaviour change literature and public health and development practice than collective efficacy. While prior efforts have been made to test the generalisability of efficacy theory [30, 31], these efforts have largely focused on explorations of self-efficacy. This may, perhaps, be an artefact of common approaches that tend to address independent, individual-level behavioural factors, yet aim for change on higher, aggregate levels.

Efficacy expectations influence performance expectations across several different dimensions, including magnitude, strength, and generality. In other words, some individuals or group members may perceive that the magnitude of their capability is limited to simple tasks, while others' may perceive that their capacity to perform simple tasks extends to more difficult tasks as well [28]. Strength of efficacy perceptions is important to consider with regard to performance expectations in that weak efficacy perceptions may be easily disrupted, while strong efficacy perceptions may endure, even when challenged [28].

3.1.1 Self-efficacy

Self-efficacy reflects an individual's perceptions regarding his or her ability to pursue a course of action related to a particular problem or task [28]. The construct is multidimensional in that it is influenced by performance accomplishments (i.e., prior mastery and/or failure experiences related to task pursuit), vicarious experiences (i.e., observing, either directly or symbolically, others pursuing tasks and the consequences of their pursuits), verbal persuasion (i.e., suggestions or inducements that influence self-efficacy perceptions), and emotional arousal (i.e., physiological states that may indicate personal competence) [28]. Given past performance influences efficacy perceptions, efficacy may increase along with successive attempts to accomplish a task or pursue a

goal, or at each subsequent stage along the stages of change toward the adoption of a novel behaviour or innovation [32]. Similarly, specific and somewhat difficult goals serve to reinforce positive efficacy perceptions as well as improved performance [28, 33, 34]. Self-efficacy has been well adopted in the examination of individual-level behavioural prediction [35].

Translating or extending concepts of efficacy to a group context may be pertinent in several situations. For instance, collective efficacy may be important in situations in which individuals alone do not have or perceive to have control over social factors or ecological conditions, and instead demonstrate a willingness and perceived ability to cooperate to establish and work toward common goals [12, 35]. While a sizable theoretical and empirical evidence base regarding perceived efficacy amongst groups exists [36-39], applications of the concept of efficacy amongst higher-order collectives, particularly in examining uptake and effectiveness of community-based interventions, remain relatively limited. The theoretical and empirical foundations of self-efficacy are relevant to collective efficacy [12], and result in a sound theoretical and empirical foundation upon which the examination of collective efficacy is grounded.

3.1.2 Collective efficacy

Collective efficacy is a socially-influenced latent construct that draws upon a combination of cognitive and socio-structural factors which facilitate peoples' shared beliefs in their collective power, or ability to come together to execute actions related to a common goal [12, 40]. In addition to shaping a group's decision to set and pursue common goals, collective efficacy also influences the amount of effort and resolve group members will expend in working toward those goals [40]. While findings from psychometric examinations of collective efficacy align to some degree, the number and nature of related factors differ. When considered in combination, psychometric examinations suggest that factors tapping to social cohesion, informal social control, and collective behavioural control are important sub-constructs of collective efficacy [41-44]. In their own analyses, researchers suggest that these factors include social attachment, social networks and personal agency, community organisation and leadership, associational participation, social response, common vision; trust, activism, informedness, belonging, association; and group competence and task analysis [41, 43, 44].

Judgements regarding collective efficacy develop when individuals within a collective consider personal, group, and situational factors [40, 45]. It is these perceptions, or judgements regarding

collective efficacy, that influence whether communal goals are set, and how a group performs as it pursues the communal goal [37, 38, 42, 46]. Evidence suggests that perceptions regarding efficacy are better predictors of subsequent behaviour than prior performance or goal attainment [40].

Efficacy constructs, both self- and collective efficacy, have been applied across several sectors, including sports, education, health, and therapy [28, 34]. These constructs are highly prevalent in the theoretical literature, and are cited as key constructs in the Theory of Planned Behaviour [47, 48], Social Action Theory [30], Goal and Goal Setting Theories [37, 49], and the Social Ecology Model of Behaviour Change [50], amongst others. Perhaps one of the reasons efficacy constructs are so prevalent in theory and behavioural prediction is the motivational nature of related constructs. This renders efficacy appraisals a major determinant of goal setting, trigger of task pursuit (individual) or cooperative behaviour and collective action (collective), and strong predictor of the quantity and duration of effort that is placed on pursuing a goal when obstacles are presented [28, 37-39].

3.1.3 The distinction and relationship between self- and collective efficacy

The main difference between self- and collective efficacy is the entity or level at which efficacy judgments are being made – at the level of the individual or a higher-order group (e.g., self-identified community, endogenously or exogenously organised group, village). Unlike self-efficacy, collective efficacy reflects a socially-influenced construct related to beliefs and expectations (personal normative beliefs, empirical and normative expectations) regarding cooperation and perceptions regarding a collective's ability and autonomy to engender and maintain change, which trigger collective action toward a communal goal [38]. An individual's perceptions regarding his or her own abilities and autonomy to pursue a task and his or her own ability and autonomy to contribute to a common goal may influence perceptions regarding the group's overall performance while pursuing that goal. In other words, important relationships between self- and collective efficacy may exist [12].

3.2 Social norms

Social norms refer to the shared understandings, rules, and expectations regarding how people behave (i.e., practices that are obligatory, permitted, or forbidden) [51, 52]. While there are different schools of thought regarding the terminology and nuances between social norms sub-constructs (e.g., descriptive norms, empirical expectations; injunctive norms, normative expectations), there is consensus amongst theorists that social norms act to maintain group

coherence and behavioural standards [52]. A collective practice is considered a social norm if people engage in the practice as a result of a socially conditioned preference emerging from beliefs regarding what other people do (i.e., empirical expectations) and beliefs regarding what other people think one ought to do (i.e., normative expectations) [53].

Social norms are contingent upon reference groups and notions of identity [54]. Social norms theory propose that biased perceptions may cause an individual to: 1) believe his or her own attitudes and behaviours differ from those in his or her referent group when, in fact, they do not (i.e., pluralistic ignorance), or 2) falsely believe that his or her own attitudes and behaviours are similar to others in the referent group when they are not (i.e., false consensus) [55, 56]. An individual's perceptions regarding what others in the referent group do, and what they think others ought to do is important in that they are documented predictors of behaviour [36]. In order to correct misconceptions, normative feedback or information regarding the actual attitudes and behaviours of individuals in the referent group can be provided [55].

From a theoretical and empirical perspective, the degree to which social norms are maintained, and influence behaviour depends on exogenous and endogenous social control mechanisms, such as collective sanctions and compliance norms [25, 57]. Evidence suggests that there is an important distinction between exogenous and endogenous control mechanisms and their influence on cooperative behaviour [14], as externally imposed sanctions tend to "crowd out" endogenous cooperative behaviour [58]. Social norms also appear prominently in the theoretical literature. They are presented as a key construct in the Theory of Reasoned Action [59], Network Theory of Collective Action [25], Social Change Theory [60], and the Theory of Interpersonal Behaviour [61], amongst others.

3.3 Social capital

Social capital is a multifaceted social construct that refers to the features of social organisation (e.g., social structures, social networks and support) and the norms and values that function within them to facilitate coordination and cooperation for mutual benefit [23, 26, 27]. It is generally accepted that there are two different forms of social capital: socio-structural and cognitive social capital. Socio-structural forms of social capital are objective, and manifest as social networks and organisational groups. Dimensions/factors and facets tapping to this social capital domain include not just groups and networks, but also information and communication structures, community

leadership and organisation, collective representation (of local interests amongst higher level authorities), and associational participation [23]. Cognitive forms of social capital are subjective, and relate to shared norms, values, attitudes, and beliefs. Dimensions/factors and facets tapping to this social capital domain include perceptions regarding social trust, connectedness, and safety; norms of reciprocity and mutual aid; and collective values [23, 62]. Cognitive forms of social capital predispose people to mutually beneficial collective action [23], which operate through socio-structural forms of social capital. Both forms of social capital are informed by social interactions and expectations, and both are mutually reinforcing [23].

Evidence suggests that actual and potential social capital increases the amount or probability of a group's mutually beneficial cooperative behaviour and collective action [23]. In other words, different forms and features of social capital may serve as resources that facilitate collective action [26, 63]. However, examinations of social capital, while perhaps necessary, are exclusively insufficient [14]. In other words, assessing social capital without assessing other social constructs known to be predictors of collective action may provide an incomplete evaluation of the interpersonal antecedents of collective action.

3.3.1 Bonding, bridging, and linking aspects of social capital

There are different types and quality of social capital. Theorists and practitioners have reflected on the concept of homophily (i.e., tendency of individuals to associate and bond with similar others) to distinguish between the quality and types of social capital [64-66]. Bonding social capital reflects relationships amongst social network members with similar social identities (i.e., homophilous), such as relatives, friends, or neighbours [26]. This type of social capital usually reflects the strongest connections, but yield the “least valuable by-product” [67]. Bridging social capital reflects relationships amongst people who are dissimilar, or differ in terms of social identity (i.e., heterophilous). Relationships between community members at large (i.e., non-leaders) and community leaders, local politicians, individuals of different ethnicity, age, socio-economic status [66] would reflect bridging social capital. While bridging social capital represents weaker connections than bonding social capital, it typically generates a more valuable “by-product” [67]. Linking social capital reflects the extent to which individuals build relationships with power structures, and organisations such as local government authorities [66]. Linking social capital represents the weakest type of connection, yet yields the most valuable by-product, as it provides a connection and therefore access to power structures [67]. The by-products of bridging and linking

social capital are meaningful in that they provide opportunities for exposure to novel ideas, values, and perspectives [66].

3.4 Social cohesion

Some conceptualisations of social cohesion propose it is the presence of strong social bonds, sense of connectedness, and the absence of social conflict [63]. Other conceptualisations suggest social cohesion reflects the behaviours and attitudes within a community that represents the propensity of its members to cooperate [68]. Social cohesion manifests as individuals who demonstrate a willingness and ability to work together to address common needs, overcome barriers, and consider the diverse interests amongst the group [69]. Theoretically proposed sub-construct domains include social inclusion, common values and civic culture, social capital, social order and control, and social mobilisation [69, 70]. Proposed dimensions/factors and facets include social solidarity, norms of reciprocity, social attachment (place attachment and place identity), social bonds (sense of belonging and trust), associations that bridge social division, willingness to participate and promote social justice or social control, absence of disparities or other forms of polarisation, absence of social conflict, shared outcomes [63, 70]. Here, social inclusion, as a sub-construct, promotes equity with regard to opportunities, and removes barriers to participation within community structures.

Literature regarding social cohesion, as a stand-alone social construct, exists. However, there is a dearth of literature regarding rigorous measure and psychometric determination of social cohesion, and its influence on behaviour. In general, literature supporting conceptualisations of social cohesion as a stand-alone construct are less common than alternative conceptualisations that conceived of social cohesion as a constituent of other social constructs. More commonly, social cohesion presents in the literature as a sub-construct of social capital or collective efficacy [41, 42, 62, 71]. Despite disparities regarding its conceptualisation, social cohesion is believed to influence the quality, effectiveness, and sustainability of interventions, especially those based on action at the community level [72, 73]. The influence of social cohesion on collective behaviour may be explained, at least partially, by the theory of diffusion of innovations, which suggests that innovative behaviours diffuse much more rapidly in communities that are cohesive and in which members know and trust each other [74]. Social cohesion is also considered in the Network Theory of Collective Action [25], Social Action Theory [30], and the Ecological Model of Health Behaviours [75].

3.4.1 Social capital vs. social cohesion

Many theorists and practitioners view social cohesion as a constituent component of social capital [15, 18, 62, 71, 73]. In general, these conceptualisations and evaluations thereof posit that social cohesion is a cognitive component of social capital that facilitates collective action via structural manifestations of social capital (e.g., formal associations, community groups) [62]. However, some conceptualisations of social cohesion view social capital as a constituent component of cohesion that draws on bonding and bridging social capital [63, 70]. In other words, social cohesion is conceived by some theorist as a as a social construct with *foundations* in local social capital, where the social capital of a particular community "take[s] on a strong sense of local space, albeit with ambiguous and fluid boundaries" [70]. Social capital is therefore proposed as a "partial re-conceptualisation of social cohesion" with social network and community links [70]. These theorists suggest that social cohesion examines concepts (i.e., sub-constructs) that move beyond those of social capital [63]. Regardless of the structure of their conceptualisations, social cohesion and social capital consider "collective, ecological dimensions of society" [63]. It is suggested, however, that these constructs be distinguished from social networks and social support, which operate at the individual level [63].

3.5 Synthesis of social construct conceptualisations

Table 1 summarises attributes of the various social constructs, and provides an overview of each construct's operational definition, related sub-constructs, and when and why the construct should be considered for use to inform the design, targeting, and evaluation of community-based interventions.

3.6 Nuances related to motivational concepts

When it comes to elucidating socially-influenced factors influential in determining whether an individual or group of individuals choose to engage in or adopt various behaviours, the domain of motivational factors that comprise locus of control, behavioural control, efficacy, and agency emerge in several theories. The set of constructs in this motivational domain tap to slightly different conceptual facets, but in essence belong to the same domain. The labels used to describe these factors may also account for differences in how theorists viewed the nature, origins, and intervening processes by which motivational factors affect intentions and behaviour. This may result in some conflation of terms. Below, common factors in this motivational domain are clarified.

Table 1. Social Constructs Framework: Summary of social construct attributes

Construct	Operational definition	Constituent sub-constructs	When to use this construct: Potential uses	Why use this construct: Attributes related to collective action	Sources
Self-efficacy	An individual's perceptions and confidence regarding his/her ability to pursue a course of action related to a particular task or problem	Performance accomplishments, vicarious experience, verbal persuasion, emotional arousal (physiological states)	**To determine and assess:** task pursuit and performance; motivational components of volitions (e.g., intentions as indicators of how hard an individual is willing to try, of how much effort s/he is planning to exert); feed-forward control, as well as feedback **To design, target, and evaluate interventions that:** address individual-level, intrapersonal motivational factors related to task pursuit, enhance one's sense of performance accomplishments & outcome expectations	Construct reveals information regarding individual-level motivation, demonstrated predictor of task pursuit, the amount and duration of effort an individual will expend in the pursuit of a task or personal goal. Better predictor of subsequent behaviour than prior task performance. May increase along with successive attempts to accomplish a task or pursue a goal, or at each subsequent stage along the stages of change toward the adoption of a novel behaviour or innovation. Self-efficacy appraisals may guide the selection of individual-level action strategies. Self-efficacy appraisals regarding an individual's abilities and autonomy to contribute to communal goals are related to collective efficacy perceptions.	Bandura 1977, 1982; Galavotti et al. 1995; Ewart 1991; Ajzen 1991; Bagozzi 1992; Delea et al. 2018
Collective efficacy*	A group's shared beliefs and confidence in its collective power, or ability to come together to execute actions related to a common goal	Collective behavioural control, informal social control, social cohesion; collective performance accomplishments, vicarious experience, verbal persuasion, emotional arousal (physiological states)	**To determine and assess:** collective goal setting, performance as the collective takes action toward communal goal(s); motivational components of volitions (e.g., intentions as indicators of how hard a group/collective of individuals is willing to try, of how much effort they are planning to exert); feed-forward control, as well as feedback **To design, target, and evaluate interventions that:** address collective-level, interpersonal factors related to communal goal-setting and performance during pursuit thereof, motivational components of volitions enhance a collective's sense of its performance accomplishments and outcome expectations	Construct reveals information regarding collective (i.e., group, community)-level motivation. Collective efficacy appraisals shape a group's decision to set and pursue common goals, influences the amount of effort and resolve group members will expend in working toward those goals. Better predictor of subsequent collective behaviour than prior performance or goal attainment. May increase along with successive attempts to pursue a goal, or at each subsequent stage along the stages of change toward the adoption of a novel behaviour or innovation. Collective efficacy appraisals may guide the selection of collective-level action strategies.	Bandura 1982, 2000; Galavotti et al. 1995; Sampson et al. 1997; Goddard et al. 2004; Carroll et al. 2005; Delea et al. 2018
Social norms	Collective or shared understandings, rules and expectations	Empirical expectations (i.e., beliefs regarding the behaviours, practices of others) and normative	**To determine and assess:** empirical expectations and normative expectations, sanctions (i.e., rewards for compliance, punishments for non-compliance) that employ intra-group normative	Construct represents interpersonal behavioural factors related to socially conditioned preferences regarding behaviours and practices of interest. They act to maintain group coherence and	Heckathorn 1990; Crawford and Ostrom 1995; Ostrom 2000;

	regarding everyday life (i.e., how individuals behave, and obligatory, permitted, or forbidden practices)	expectations (i.e., beliefs regarding what people ought to do), social control mechanisms (e.g., endogenous and exogenous sanctions, compliance norms)	control to influence individuals/the group as a whole to adopt mutually beneficial cooperative behaviours **To design, target, and evaluate interventions that:** address interpersonal behavioural factors related to socially conditioned preferences regarding behaviours and practices of interest, and providing normative feedback as an approach for correcting misperceptions that influence behaviour	behavioural standards. An individual's perceptions regarding what others in the referent group do, and what they think others ought to do is important in that they are documented predictors of behaviour. In order to correct misconceptions, normative feedback or information regarding the actual attitudes and behaviours of individuals in the referent group can be provided.	Berkowitz 2004; Bicchieri 2017
Social capital†	Features of social organisation (e.g., social structures, social networks and support) and the norms and values that function therein to facilitate coordination and cooperation for mutual benefit; assets operating to create/enhance bonds, bridges, linkages	**Structural components:** groups & networks (social networks, social support), community leadership & organisation, associational participation, information and communication **Cognitive components:** social trust, reciprocity & mutual aid, connectedness, safety, collective norms & values	**To determine and assess:** actual or potential structural and cognitive components (social, psychological, cultural, cognitive, and institutional assets) that increase the amount or probability of a group's mutually beneficial cooperative behaviour and collective action **To design, target, and evaluate interventions that:** address interpersonal behavioural factors related to social structures and networks, and the trust and norms of reciprocity that work within them to create externalities for the collective	Construct explores bonding, bridging, and linking features of social structures (e.g., networks) and the levels of interpersonal trust and norms of reciprocity which act as resources for individuals, and facilitators of collective action. Necessary but exclusively insufficient (in and of itself) for the development and assessment of collective action.	Bourdieu 1977; Putnam 1983, 2004; Uphoff 1999; Berkman & Kawachi 2000; Ostrom 2000; Forrest & Kearns 2001; Grootaert & van Bastelaer 2002, 2013
Social cohesion‡	Presence of strong social bonds, sense of connectedness, and the absence of social conflict	Social inclusion, common values and civic culture, social mobilisation, social capital, social order and control	**To assess:** Social bonds (sense of belonging and trust), social inclusion, social solidarity, social attachment (place attachment and place identity), norms of reciprocity, associations that bridge social division, willingness to participate and promote social justice or social control, and social control **To design, target, and evaluate interventions that:** Create social bonds and a sense of unity/solidarity, foster an environment that promotes social inclusion, minimise or address disparities/other forms of polarisation and social conflict, and create perceptions and norms regarding shared outcomes	Construct and sub-constructs such as social inclusion and attachment may be important factors to consider given they serve to minimise barriers to participation and promote equal access to opportunities and resources. Communities or collectives high in social cohesion may facilitate the diffusion of innovations, as innovative behaviours diffuse much more rapidly in communities that are cohesive and in which members know and trust each other.	Berkman & Kawachi 2000; Forrest & Kearns 2001; Rogers 1983

Notes. * Some conceptualisations of social capital have included collective efficacy as a sub-construct (Berkman & Kawachi 2000, Perkins & Long 2002). † Social capital and/or aspects thereof (i.e., constituent sub-constructs) have been conceptualised as components of collective efficacy (Carroll et al. 2005, Delea et al. 201) and social cohesion (Berman & Kawachi 2008, Forrest & Kearns 2001).
‡ Social cohesion and/or aspects thereof (i.e., constituent sub-constructs) have been conceptualised as components of social capital and collective efficacy (Sampson et al. 1997, Delea et al. 2018).

Distinctions are summarised, and concepts are presented in terms of their levels of specificity in Table 2.

Beliefs regarding competency and efficacy, performance expectancies (i.e., expectancies of success or failure), and perceived control over outcomes (i.e., behavioural control) should be differentiated. In fact, Bandura indicates that "[p]erceived self-efficacy and beliefs about the locus of causality must be distinguished…" [28].

The concept of locus of control emerged from Rotter's Social Learning Theory [29], and was used to describe both outcome expectancies, and the value placed on an outcome [76]. Locus of control is principally concerned with causal beliefs about action-outcome contingencies [29]. An internal locus of control reflects situations in which individuals believe their actions dictate resulting outcomes, while an external locus of control reflects situations in which outcomes are a product of external forces or factors [77]. Locus of control reflects a higher-level, domain-specific concept. In some theories, locus of control is synonymous to self-efficacy [28]. However, several theories distinguish between locus of control and efficacy, and suggest that efficacy is only reflected when locus of control is perceived to be internal as opposed to external [76, 77].

Behavioural control relates to perceptions regarding one's control over the performance of a behaviour [47]. In other words, behavioural control is "the person's belief as to how easy or difficult performance of the behaviour is likely to be" [78]. Behavioural control reflects intentions, and an intention is only expected to manifest when an individual maintains behaviour control perceptions [49]. Ajzen suggests that self-efficacy is a component of behavioural control [48]. This concept is slightly more specific than locus of control in that it is a behaviour-specific concept.

Bandura set forth a social cognitive model of motivation that focuses on efficacy and agency perceptions [12, 28]. Efficacy is described elsewhere in this paper, and agency reflects the capacity of an individual to act in a given environment [79]. Unlike Expectancy Value Theory [80] and Attribution Theory [81], Social Cognitive Theory [79] suggests that efficacy expectations affect behaviour, which influences outcome expectations and actual downstream

Table 2. Distinctions between common motivational concepts

Label	Type of theory employing label	Definition	Specificity (i.e., level of evaluation)	Author/year
Locus of control				
	Locus of control theories, cognitive-affective theories	A concept used to describe both outcome expectancies, and the value placed on an outcome; primarily concerned with causal beliefs about action-outcome contingencies	Domain-specific	Bandura 1977*, Rotter 1966
Behavioural control				
	Cognitive-affective theories, locus of control theories	Perceived control over the performance of a behaviour	Behaviour-specific	Ajzen 1985†, Ajzen & Madden 1986
Efficacy				
	Social cognitive theories, cognitive-affective theories, interpersonal construct theories	*Self-efficacy:* An individual's confidence in his/her ability to organise and execute a given course of action to solve a problem or accomplish a task	Task-specific	Bandura 1977, Galavotti 1995
		Collective efficacy: A group's perceptions regarding its ability and autonomy to set and pursue a communal goal	Goal-specific	Bandura 1997, Galavotti 1995
Expectancies				
	Cognitive-affective theories, expectancy-value theories	Judgement that task completion or behaviour change will produce a specified outcome	Task/goal-specific	AbuSabha 1997, Bandura 1977, Westby 2005, Schwarzer 2008
Agency				
	Locus of control theories, social cognitive theories	The capacity of an individual to act in a given environment	Context and task/goal-specific	Skinner 1995, 1998; Bandura 2001

Notes. * Bandura proposes self-efficacy is synonymous to locus of control. † Ajzen suggests self-efficacy is a component of behavioural control.

outcomes. Both agency and efficacy are more specific than behavioural control, as they are task/goal-specific.

Outcome expectancies reflect an individual's judgement that task completion or behaviour change will produce a specified outcome [76]. Expectancies influence behaviour, and are shaped by mastery experiences and reinforcement [28]. Theories drawing on outcome expectancies propose that individuals typically take into account several possible outcomes, and weigh them against perceived benefits before deciding whether to pursue a task [76]. In other words, an individual's beliefs regarding expected outcomes and the value of those outcomes has a considerable influence on motivational processes. Positive outcome expectancies and perceptions of self-efficacy work jointly with each other, and influence intention [82].

4. Discussion

This paper sought to examine characteristics of the various social constructs commonly identified as theoretically-grounded and evidence-based predictors of cooperative behaviour and collective action. Appropriately incorporating these constructs into the design, targeting, and evaluation of community-based interventions may serve to elucidate and address persistent shortcomings in behaviour change interventions. Information regarding when and why to use each construct was provided in order to help guide programme implementers and researchers in their consideration and selection of behavioural antecedents. Given the prevalence of efficacy-related social constructs in the behaviour and behaviour change literature, various aspects of these motivational constructs were examined.

Our summary highlighted the unique contributions of each social construct. Findings suggest efficacy-related constructs allow for an examination of perceived capability and autonomy to pursue tasks/goals, and motivational factors related to task or communal goal setting, performance during pursuit thereof, and the amount of effort and persistence that will be exerted in pursuit of the task or goal. Social norms investigate perceptions related to the attitudes and behaviours of others as well as beliefs regarding what people should do. These perceptions influence socially conditioned preferences regarding behaviours and practices. Social capital relates to the socio-structural (e.g., social networks, social support) and cognitive (e.g., trust, norms of reciprocity) features of a group, community, or collective that serve as resources for individuals, and facilitators of collective action. Social cohesion pertains to the strength of social bonds, sense of connectedness, and the absence of social conflict within a

group, community, or collective. Social cohesion and sub-constructs such as social inclusion and attachment may be important factors to consider given they serve to minimise barriers to participation and promote equal access to opportunities and resources.

In order to facilitate further incorporation of efficacy-related constructs into community-based interventions, we outlined distinctions between the domain of motivational factors that comprise locus of control, behavioural control, efficacy, and agency constructs, as they are often conflated in the literature. These findings indicate that there are various levels of specificity with regard to these motivational concepts. As such, when designing and evaluating interventions, it is important to ensure that the level of specificity of the intervention techniques employed to addressed key behavioural factors align with the specificity of the motivational factor itself. For example, a behaviour-specific intervention technique should be employed to address behavioural control-related factors, while a task or goal-specific intervention technique should be employed to address efficacy-related factors. Distinctions should be made between locus of control and efficacy [28].

Findings from our review demonstrate inconsistencies with regard to the conceptualisations and the underlying structures of various social constructs. Rather than striving for consensus regarding the conceptualisations of these constructs, it is important to ensure that interventions appropriately consider relevant sub-constructs that attend to context-specific barriers of behavioural adoption and programmatic goals. Theory can and should be used to inform intervention design, and these theories can be used to determine how to cite related factors as constituents of those constructs. For example, proposed sub-constructs, or factors related to social capital and social cohesion overlap broadly. Factors and related sub-constructs indicated in the theories used to develop and evaluate the intervention at hand should be addressed via the intervention, as indicated by the programme's theory of change.

Examinations of social constructs are often limited to a single construct, across all contexts, without regard to the construct's attributes and whether it is best suited to explain or predict behaviours promoted by the intervention of interest. For instance, despite the fact that efficacy-related constructs are highly prevalent in theoretical literature and the construct relates to important concepts such as a collective's perceptions regarding its abilities to pursue common goals, the construct is rarely employed in community-based intervention evaluation. Instead, social capital is often exclusively employed to examine disparities between the uptake and maintenance of community-based interventions related to the development and maintenance

of communal goods or resources [5, 15, 18]. The exclusive selection of this social construct occurs despite conclusions that social capital may be necessary, but solely insufficient for the development and assessment of collective action [17]. Findings from our social constructs examination demonstrate that investigation of social capital, while useful for assessing attributes related to cognitive and structural capital, fail to consider key motivational factors that trigger communities or collectives to cooperate, coordinate, and take action. Factors such as those related to perceived capabilities and autonomy to pursue change, whether a group will set and pursue communal goals, and the amount of effort and resolve that will be expended while working toward those goals are only explored through examinations of collective efficacy.

When used appropriately and in combination, measures of social constructs and their related constituents can be used to help explain why and how community-based interventions were successful (or not) in bringing about collective action. They can also be used to target communities that are open to and ready for change, and are on-board to employ cooperative solutions. Prior to selecting social constructs for inclusion, programme implementers should clearly specify the programme's goals, objectives, and evidence-based barriers and facilitators of the behaviours being targeted by the programme (i.e., those identified through formative research). Detailing this information will provide information regarding the factors that need to be addressed by the intervention, and through the use of the social constructs framework presented herein, assist the implementer or researcher in identifying the particular construct to which those factors tap. By designing community-based interventions to act on various constructs and constituent factors identified as change mechanisms, purposive action can be taken to enhance these constructs, and may in turn, improve intervention uptake and effectiveness [23].

2.7 Chapter references

1. Fuller, J.A. and J.N. Eisenberg, *Herd Protection from Drinking Water, Sanitation, and Hygiene Interventions.* Am J Trop Med Hyg, 2016. **95**(5): p. 1201-1210.
2. Garn, J.V., S. Boisson, R. Willis, A. Bakhtiari, T. al-Khatib, K. Amer, W. Batcho, P. Courtright, M. Dejene, A. Goepogui, K. Kalua, B. Kebede, C.K. Macleod, K.I.M. Madeleine, M.S.A. Mbofana, C. Mpyet, J. Ndjemba, N. Olobio, A.L. Pavluck, O. Sokana, K. Southisombath, F. Taleo, A.W. Solomon, and M.C. Freeman, *Sanitation and water supply coverage thresholds associated with active trachoma: Modeling cross-sectional data from 13 countries.* PLOS Neglected Tropical Diseases, 2018. **12**(1): p. e0006110.
3. Oswald, W.E., A.E. Stewart, M.R. Kramer, T. Endeshaw, M. Zerihun, B. Melak, E. Sata, D. Gessese, T. Teferi, Z. Tadesse, B. Guadie, J.D. King, P.M. Emerson, E.K. Callahan, D. Flanders, C.L. Moe, and T.F. Clasen, *Active trachoma and community use of sanitation, Ethiopia.* Bulletin of the World Health Organization, 2017. **95**(4): p. 250-260.
4. Story, W.T., H. Tura, J. Rubin, B. Engidawork, A. Ahmed, F. Jundi, T. Iddosa, and T.H. Abrha, *Social capital and disaster preparedness in Oromia, Ethiopia: An evaluation of the "Women Empowered" approach.* Social Science & Medicine, 2018.
5. Person, M., M. Delea, J. Garn, K. Alexander, B. Abaire, and M. Freeman, *Assessing the influence of social capital on water point sustainability in rural Ethiopia.* Journal of Water, Sanitation and Hygiene for Development, 2017. **7**(4): p. 611-622.
6. Dickin, S., E. Bisung, and K. Savadogo, *Sanitation and the commons: The role of collective action in sanitation use.* Geoforum, 2017. **86**: p. 118-126.
7. Davis, R., R. Campbell, Z. Hildon, L. Hobbs, and S. Michie, *Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review.* Health Psychology Review, 2015. **9**(3): p. 323-344.
8. Head, K.J. and S.M. Noar, *Facilitating progress in health behaviour theory development and modification: the reasoned action approach as a case study.* Health Psychology Review, 2014. **8**(1): p. 34-52.
9. Figueroa, M.E. and D.L. Kincaid, *Social, Cultural and Behavioral Correlates of Household Water Treatment and Storage*, in *Health Communication Insights*. 2010, Johns Hopkins Bloomberg School of Public Health, Center for Communication Programs: Baltimore.
10. Clasen, T., S. Boisson, P. Routray, B. Torondel, M. Bell, O. Cumming, J. Ensink, M. Freeman, M. Jenkins, M. Odagiri, S. Ray, A. Sinha, M. Suar, and W.-P. Schmidt, *Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial.* The Lancet Global Health, 2014. **2**(11): p. e645-e653.
11. Stoller, N.E., T. Gebre, B. Ayele, M. Zerihun, Y. Assefa, D. Habte, Z. Zhou, T.C. Porco, J.D. Keenan, J.I. House, B.D. Gaynor, T.M. Lietman, and P.M. Emerson, *Efficacy of latrine promotion on emergence of infection with ocular Chlamydia trachomatis after mass antibiotic treatment: a cluster-randomized trial.* Int Health, 2011. **3**(2): p. 75-84.
12. Bandura, A., *Exercise of Human Agency through Collective Efficacy.* Current Directions in Psychological Science, 2000. **9**(3): p. 75-78.
13. Ostrom, E., *Governing the commons : the evolution of institutions for collective action.* 2015.
14. Ostrom, E., *Collective Action and the Evolution of Social Norms.* The Journal of Economic Perspectives, 2000. **14**(3): p. 137-158.
15. Bisung, E., S.J. Elliott, C.J. Schuster-Wallace, D.M. Karanja, and A. Bernard, *Social capital, collective action and access to water in rural Kenya.* Social Science & Medicine, 2014. **119**: p. 147-154.

16. Dean, A.J., K.S. Fielding, J. Lindsay, F.J. Newton, and H. Ross, *How social capital influences community support for alternative water sources.* Sustainable Cities and Society, 2016. **27**: p. 457-466.
17. Ostrom, E., *Social capital : a fad or a fundamental concept?*, in *Social capital : a multifaceted perspective.*, P. Dasgupta and I. Serageldin, Editors. 2000, World Bank Publications: Washington, D.C. p. 172-214.
18. Cameron, L.A., S. Olivia, and M. Shah, *Initial Conditions Matter: Social Capital and Participatory Development.* IZA Discussion Papers, 2015. **No. 9563**.
19. Prabhakaran, P., K. Kar, L. Mehtab, and S. Chowdhurya, *Impact of community-led total sanitation on women's health in urban slums: A case study from Kalyani Municipality.* IDS Evidence Report, 2016. **No. 194**.
20. Hathi, P., D. Spears, and D. Coffey, *Can collective action strategies motivate behaviour change to reduce open defecation in rural India?* Waterlines, 2016. **35**(2): p. 118-135.
21. Glanz, K. and D.B. Bishop, *The Role of Behavioral Science Theory in Development and Implementation of Public Health Interventions.* Annual Review of Public Health, 2010. **31**(1): p. 399-418.
22. Petraitis, J., B.R. Flay, and T.Q. Miller, *Reviewing theories of adolescent substance use: Organizing pieces in the puzzle.* Psychological Bulletin, 1995. **117**(1): p. 67-86.
23. Uphoff, N., *Understanding social capital: learning from the analysis and experience of participation*, in *Social Capital: a Multifaceted Perspective*, P. Dasgupta and I. Serageldin, Editors. 1999, World Bank Publications: Washington, D.C.
24. Olson, M., *The logic of collective action: public goods and the theory of groups*. 2012, Cambridge, Mass: Harvard University Press.
25. Heckathorn, D.D., *Collective Sanctions and Compliance Norms: A Formal Theory of Group-Mediated Social Control.* American Sociological Review, 1990. **55**(3): p. 366-384.
26. Putnam, R.D., *The Prosperous Community: Social Capital and Public Life.* The American Prospect, 1993(13): p. 35-42.
27. Bourdieu, P., *The forms of capital*. Handbook of Theory and Research for the Sociology of Education, ed. J. Richardson. 1986, New York: Greenwood.
28. Bandura, A., *Self-efficacy: Toward a unifying theory of behavioral change.* Psychological Review, 1977. **84**(2): p. 191-215.
29. Rotter, J.B., *Generalized expectancies for internal versus external control of reinforcement.* Psychological Monographs: General and Applied, 1966. **80**(1): p. 1-28.
30. Ewart, C.K., *Social action theory for a public health psychology.* American Psychologist, 1991. **46**(9): p. 931-946.
31. Neff, J.A. and S.A. MacMaster, *Applying Behavior Change Models to Understand Spiritual Mechanisms Underlying Change in Substance Abuse Treatment.* The American Journal of Drug and Alcohol Abuse, 2005. **31**(4): p. 669-684.
32. Galavotti, C., R.J. Cabral, A. Lansky, D.M. Grimley, G.E. Riley, and J.O. Prochaska, *Validation of measures of condom and other contraceptive use among women at high risk for HIV infection and unitended pregnancy.* Health Psychology, 1995. **14**(6): p. 570.
33. Schunk, D.H., *Goal Setting and Self-Efficacy During Self-Regulated Learning.* Educational Psychologist, 1990. **25**(1): p. 71-86.
34. Eccles, J.S. and A. Wigfield, *Motivational beliefs, values, and goals.* Annual Review of Psychology, 2002. **53**(1): p. 109.
35. Kelder, S.H., D. Hoelscher, and C.L. Perry, *How individuals, environments, and health behaviours interact*, in *Health Behavior: Theory, Research, and Practice*, K. Glanz, B. Rimer, and K. Viswanath, Editors. 2008, Jossey-Bass: San Francisco, CA.

36. Ahern, J., S. Galea, A. Hubbard, and S.L. Syme, *Neighborhood smoking norms modify the relation between collective efficacy and smoking behavior.* Drug and Alcohol Dependence, 2009. **100**(1): p. 138-145.
37. Durham, C.C., D. Knight, and E.A. Locke, *Effects of Leader Role, Team-Set Goal Difficulty, Efficacy, and Tactics on Team Effectiveness.* Organizational Behavior and Human Decision Processes, 1997. **72**(2): p. 203-231.
38. Bray, S.R., *Collective Efficacy, Group Goals, and Group Performance of a Muscular Endurance Task.* Small Group Research, 2004. **35**(2): p. 230-238.
39. Kocaeksi, S. and A. Ezgi Gazioglu, *The Evaluation of Self-efficacy, Collective Efficacy Beliefs in Handball in Terms of Gender.* Procedia - Social and Behavioral Sciences, 2014. **159**: p. 125-127.
40. Bandura, A., *Self-efficacy mechanism in human agency.* American Psychologist, 1982. **37**(2): p. 122-147.
41. Delea, M., G. Sclar, M. Woreta, R. Haardörfer, C. Nagel, B. Caruso, R. Dreibelbis, A. Gobezayehu, T. Clasen, and M. Freeman, *Collective Efficacy: Development and Validation of a Measurement Scale for Use in Public Health and Development Programmes.* International Journal of Environmental Research and Public Health, 2018. **15**(10): p. 2139.
42. Sampson, R.J., S.W. Raudenbush, and F. Earls, *Neighborhoods and violent crime: a multilevel study of collective efficacy.* Science, 1997. **277**(5328): p. 918-24.
43. Carroll, J.M., M.B. Rosson, and J. Zhou, *Collective efficacy as a measure of community*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2005, ACM: Portland, Oregon, USA. p. 1-10.
44. Goddard, R.D., W.K. Hoy, and A.W. Hoy, *Collective Efficacy Beliefs:Theoretical Developments, Empirical Evidence, and Future Directions.* Educational Researcher, 2004. **33**(3): p. 3-13.
45. Hipp, J.R., *Collective efficacy: How is it conceptualized, how is it measured, and does it really matter for understanding perceived neighborhood crime and disorder?* Journal of Criminal Justice, 2016. **46**: p. 32-44.
46. Hodges, L. and A.V. Carron, *Collective efficacy and group performance.* International Journal of Sport Psychology, 1992. **23**(1): p. 48-59.
47. Ajzen, I., *The theory of planned behavior.* Organizational Behavior and Human Decision Processes, 1991. **50**(2): p. 179-211.
48. Ajzen, I., *Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior.* Journal of Applied Social Psychology, 2002. **32**(4): p. 665-683.
49. Bagozzi, R.P., *The Self-Regulation of Attitudes, Intentions, and Behavior.* Social Psychology Quarterly, 1992. **55**(2): p. 178-204.
50. Panter-Brick, C., S.E. Clarke, H. Lomas, M. Pinder, and S.W. Lindsay, *Culturally compelling strategies for behaviour change: A social ecology model and case study in malaria prevention.* Social Science & Medicine, 2006. **62**(11): p. 2810-2825.
51. Lewin, K., *Field Theory in Social Science*. 1951, New York: Harper and Row.
52. Crawford, S.E.S. and E. Ostrom, *A Grammar of Institutions.* The American Political Science Review, 1995. **89**(3): p. 582-600.
53. Bicchieri, C., *The grammar of society : the nature and dynamics of social norms*. 2006, New York: Cambridge University Press.
54. Paluck, E. and H. Shepard, *The Salience of Social Referents: A Field Experiment on Collective Norms and Harassment Behavior in a School Social Network.* Journal of Personality and Social Psychology, 2012. **103**(6): p. 899-915.
55. Berkowitz, A.D., *Social norms approach: Theory, research, and annotated bibliography*. 2004: New York.

56. Shelton, J.N. and J.A. Richeson, *Intergroup contact and pluralistic ignorance.* J Pers Soc Psychol, 2005. **88**(1): p. 91-107.
57. Schwartz, S.H., *Normative Influences on Altruism11This work was supported by NSF Grant SOC 72-05417. I am indebted to L. Berkowitz, R. Dienstbier, H. Schuman, R. Simmons, and R. Tessler for their thoughtful comments on an early draft of this chapter*, in *Advances in Experimental Social Psychology*, L. Berkowitz, Editor. 1977, Academic Press. p. 221-279.
58. Frey, B.S., *How Intrinsic Motivation is Crowded out and in.* Rationality and Society, 1994. **6**(3): p. 334-352.
59. Fishbein, M. and I. Ajzen, *Belief, attitude, intention, and behavior : an introduction to theory and research*. 1980, Reading, Mass.: Addison-Wesley.
60. Thompson, B. and S. Kinne, *Social change theory: applications to community health*, in *Health Promotion at the Community Level: New Advance*, N. Bracht, Editor. 1999, SAGE Publications, Inc.: Thousand Oaks, California.
61. Triandis, H.C., *Interpersonal behavior*. 1977, Monterey, Calif: Brooks/Cole Publishing Co.
62. Grootaert, C. and T. Van Bastelar, *Understanding and Measuring Social Capital A Multidisciplinary Tool for Practitioners.* 2013.
63. Berkman, L.F., I. Kawachi, and G.M. Maria, eds. *Social Epidemiology*. 2 ed. 2014, Oxford University Press.
64. Lin, N., *Social capital : a theory of social structure and action*. 2011, Cambridge, UK; New York: Cambridge University Press.
65. Gittell, R.J. and A. Vidal, *Community organizing : building social capital as a development strategy*. 1999, Thousand Oaks, Calif.: Sage Publ.
66. Szreter, S. and M. Woolcock, *Health by association? Social capital, social theory, and the political economy of public health.* International Journal of Epidemiology, 2004. **33**(4): p. 650-667.
67. Hawkins, R.L. and K. Maurer, *Bonding, Bridging and Linking: How Social Capital Operated in New Orleans following Hurricane Katrina.* The British Journal of Social Work, 2010. **40**(6): p. 1777-1793.
68. Acket, S., M. Borsenberger, P. Cikes, and F. Sarracino, *Measuring and validating social cohesion: a bottom-up approach*. 2011, CEPS/INSTEAD Luxembourg.
69. OECD, *Perspectives on Global Development 2012: Social Cohesion in a Shifting World*. 2012, OECD: Paris, France.
70. Forrest, R. and A. Kearns, *Social Cohesion, Social Capital and the Neighbourhood.* Urban Studies, 2001. **38**(12): p. 2125-2143.
71. Perkins, D.D. and D.A. Long, *Neighborhood sense of community and social capital: A multi-level analysis*, in *Psychological sense of community: Research, applications, and implications*, A. Fisher, C. Sonn, and B. Bishop, Editors. 2002, Plenum: New York. p. 291-318.
72. Cameron, A. and D. Miller *A Practitioner's Guide to Cluster-Robust Inference.* The Journal of Human Resources, 2015. **50**(2): p. 317-372
73. Colletta, N.J. and M.L. Cullen, *Violent conflict and the transformation of social capital : lessons from Cambodia, Rwanda, Guatemala, and Somalia* 2000, The World Bank: Washington, D.C.
74. Rogers, E., *Diffusion of innovations*. Third ed. 1983, New York, New York: Macmillan Publishing Company, Inc.
75. McLeroy, K.R., D. Bibeau, A. Steckler, and K. Glanz, *An ecological perspective on health promotion programs.* Health Educ Q, 1988. **15**(4): p. 351-77.

76. AbuSabha, R. and C. Achterberg, *Review of Self-Efficacy and Locus of Control for Nutrition- and Health-Related Behavior.* Journal of the American Dietetic Association, 1997. **97**(10): p. 1122-1132.
77. Katz, D.L., *Behavior Modification in Primary Care: The Pressure System Model.* Preventive Medicine, 2001. **32**(1): p. 66-72.
78. Ajzen, I. and T.J. Madden, *Prediction of goal-directed behavior-attitudes, intentions, and perceived behavioral-control.* Journal of experimental social psychology, 1986. **22**(5): p. 453-474.
79. Bandura, A., *Social cognitive theory: an agentic perspective.* Annu Rev Psychol, 2001. **52**: p. 1-26.
80. Atkins, L., J. Francis, R. Islam, D. O'Connor, A. Patey, N. Ivers, R. Foy, E.M. Duncan, H. Colquhoun, J.M. Grimshaw, R. Lawton, and S. Michie, *A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems.* Implementation Science, 2017. **12**(1): p. 77.
81. Jones, E.E., *How Do People Perceive the Causes of Behavior? Experiments based on attribution theory offer some insights into how actors and observers differ in viewing the causal structure of their social world.* American Scientist, 1976. **64**(3): p. 300-305.
82. Schwarzer, R., *Modeling Health Behavior Change: How to Predict and Modify the Adoption and Maintenance of Health Behaviors.* Applied Psychology, 2008. **57**(1): p. 1-29.

PART II: PSYCHOMETRIC & COGNITIVE VALIDATION OF COLLECTIVE EFFICACY SCALES

Chapter 3: Research contexts and anchor study designs

3.1 Chapter overview

This doctoral research consisted of examinations of collective efficacy in two separate contexts – rural Odisha, India, and peri-urban and rural Amhara, Ethiopia. Performing this work within the context of two different studies conducted on two different continents allowed for an iterative scale development and refinement approach. This chapter presents information about the contexts in which this research was conducted, including information about the national and sub-national settings, anchor study aims and objectives, and details related to the interventions under evaluation. A brief summary of the timing of the work performed in each context is also provided.

3.2 Employment of data from India and Ethiopia – an iterative scale refinement process

Chapter 2 established a case for the utility of CE in the examination of antecedents of collective practices. There, I highlighted the dearth of evidence and psychometric examinations thereof on CE and the uptake of community-based interventions coming from low- and middle-income countries, where reliance on collective action and cooperative behaviour is common. To address this gap, I led a team effort to design a series of studies intended to develop, validate, and employ scales to assess CE and compare perceptions between men and women. With the team, I hypothesised that CE is an important antecedent of cooperative behaviour and collective action needed to bring about sustained adoption of improved WASH practices at a level required to realise health impacts. As such, two studies, highlighted in this thesis, were designed to examine CE at various points in the programme implementation cycle. These two CE studies were embedded within two anchor studies evaluating the effectiveness of two WASH interventions, one in Odisha, India, and one in Amhara, Ethiopia.

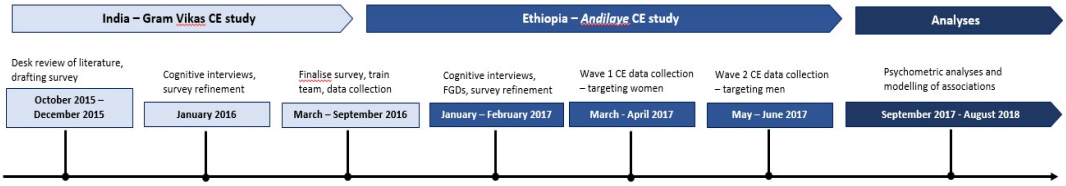
In each of the contexts in which this thesis research was conducted, the anchor WASH study provided a unique opportunity for investigating the role that CE and its constituent factors play

in uptake of improved WASH behaviours and practices. The design of the two anchor studies differed, as did the interventions being evaluated and the contexts in which the studies were being carried out. Each anchor study had its strengths and weaknesses with regard to the aspects of CE that could be explored. Together, these studies allowed for an iterative learning and scale development process while also providing an early opportunity to identify the CE factors that may transcend contexts.

The India study provided an opportunity to develop and administer an early version of the CE survey. Given the scope of the study, and the number of other sub-studies couched within it, our formative research was somewhat limited in time and scope (details provided in Research Paper 3 and its Supplemental Material). We also needed to limit the number of items our survey contained given the number of other survey modules that needed to be deployed during the same study rounds. Finally, the lack of pre-intervention measures of CE and its related factors limited examinations into causal effect and directionality of the relationship between CE factors and indicators of improved WASH practices.

Despite these limitations, the India work provided rich information and lessons that were taken into a subsequent examination of CE in an ex-ante cluster-randomised controlled trial I was helping to design and manage in Ethiopia. Given the nature of the intervention we were designing for that study, I advocated for the incorporation of the collective efficacy survey. Our funding partners were supporting a six-month formative research phase, so we were able to carry out additional formative work that was more extensive in terms of scope and time, which allowed for further refinement of the CE survey instrument. We were also able to leverage lessons learnt from the India study. As such, the Ethiopia work was more comprehensive than, and benefited from, the India work. Below, details are presented on each study context and each anchor study in which the two CE studies comprising this thesis research were situated. Figure 3.1 provides a visual depiction of the timing of the overarching CE inquiry. It is worth noting that while the India CE study was conducted first, the CE scales from the Ethiopia study were developed first. The reason being the Ethiopia dataset represented a more nuanced and comprehensive assessment of CE. Based on additional formative work and emerging literature, 14 survey items were added to the tool previously used in India prior to piloting and administration in Ethiopia. While the India dataset contained data from both men and women, the Ethiopia dataset contained data on household-level male-female dyads. Therefore, as a result of conversations with statisticians and other colleagues with scale development experience, I decided to analyse the more comprehensive dataset first.

Figure 3.1 Thesis research timeline



3.3 Why examine collective efficacy in the context of WASH interventions

On an annual basis, approximately 1.7 billion cases of diarrhoeal disease occur amongst children across the globe [1]. Diarrhoeal disease is the leading cause of malnutrition, and at 1.3 million deaths per year globally, it represents the ninth leading cause of death [2]. However, evidence suggests that the disease burden associated with diarrhoea can be prevented through WASH interventions [3-5]. Estimates from burden of disease studies substantiate the need to improve water and sanitation to prevent diarrhoeal disease burden [6]. As with diarrhoeal disease, poor WASH conditions are also associated with neglected tropical diseases (NTDs), such as soil-transmitted helminthiasis (STH), schistosomiasis, and trachoma, amongst others [7-9].

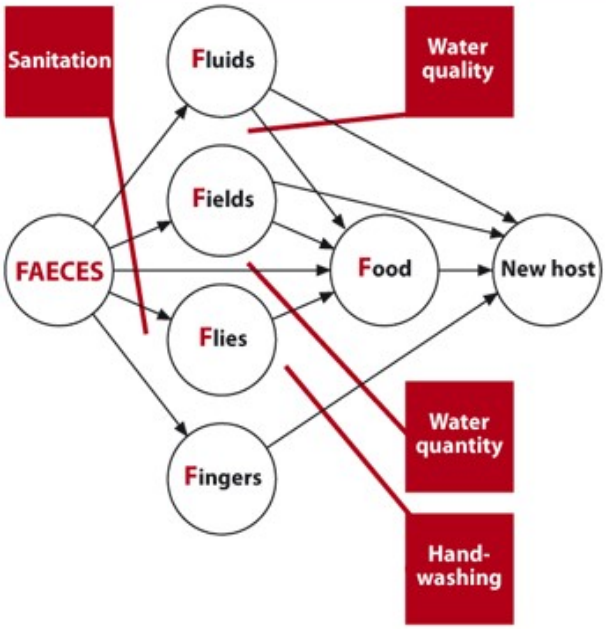
An examination of Wagner and Lanoix's F-diagram [10] reveals the pathways through which many WASH-related diseases are transmitted (Figure 3.2). This diagram depicts important routes of faecal-oral transmission, which occurs when pathogens from faeces or faecal particles from one host pose a risk to a new host (or the same host, as may be the case with auto-inoculation). A lack of adequate sanitation, poor hygiene, and inadequate water quality and quantity perpetuate faecal-oral transmission. These factors are summarised by the five F's in the F-diagram: fingers, flies, fields, fluids, and food [10].

While these factors represent important transmission routes for faecal-oral transmission, they do not fully enumerate all routes important to the transmission of all WASH-related diseases. For example, fomites may also play an important role in indirect person-to-person transmission of WASH-preventive diseases, such as trachoma [11, 12]. Fomites are inanimate objects that can become contaminated and serve as mechanisms of transfer between individuals. Evidence suggests that objects such as bedsheets, clothing, and shared cloths or towels can become contaminated with *C. trachomatis*, the causal agent of trachoma, and serve to transmit the bacteria from one person to another [12].

Addressing shortcomings in environmental conditions and personal practices through WASH interventions may serve to provide barriers to the transmission of WASH-preventive diseases.

The intersecting red lines depicted in Figure 3.2 indicate factors that can interrupt the transmission of pathogens along the various routes. A considerable evidence base indicates that interventions such as point-of use filters with safe water storage; water piped on-premise, particularly with continuous availability; handwashing with soap; and those which improve coverage and utilisation of improved sanitation facilities are associated with lower risk of diarrhoeal morbidity [13]. Further evidence suggests that improved sanitation is not only protective against diarrhoea, but also active trachoma, some STH species, and schistosomiasis [14]. Some, albeit weak, evidence suggests that improved sanitation may also be protective against stunting (i.e., height-for-age z-score) [14].

Figure 3.2 Adaptation of Wagner & Lanoix's F-diagram



Source: https://newint.org/features/2008/08/01/toilets-facts

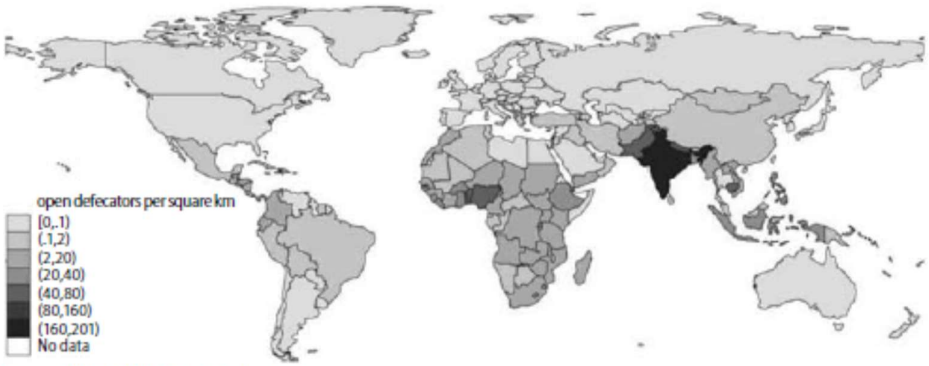
Despite these findings, overall quality of evidence is low, and recent experimental studies have revealed lower than expected health gains related to exposure to WASH interventions [15-17]. However, in response to these emerging findings, questions are being posed regarding: 1) whether it is appropriate to deploy sanitation and hygiene interventions without accompanying water supply improvements [18, 19], and 2) whether a certain level of sanitation utilisation and hygiene practices are required in order to achieve measurable health impact [18]. This second question, in particular, is relevant to this thesis work given evidence suggests that collective action is required for WASH interventions to reach the coverage and utilisation levels likely required to realise health gains through "herd protection" [20-22]. Therefore, this investigation into CE and other social constructs that represent behavioural factors of collective practices is pertinent, particularly given the current climate of WASH research and intelligent intervention design.

3.4 Water, sanitation, and hygiene in India

3.4.1 Intervention context – Odisha, India

Although the country has made recent progress, a large proportion of the world's open defecation occurs in India. Approximately 37% of the 892 million people worldwide who practice open defecation reside in India [23]. While the data were captured over five years ago, Figure 3.3 provides a visual depiction of the number of open defecators per square kilometre, with India clearly standing out as the country with the highest density of open defecators. In order to address this issue, the Prime Minister launched the Swachh Bharat Mission in October 2014, with the goal of ending open defecation in India by 2019.

Figure 3.3 Open defecators per square kilometre (2012)



Source: UNICEF-WHO Joint Monitoring Programme.

The recent National Family Health Survey-4 (NFHS-4) conducted during 2015-2016 provides information on population, health and nutrition. This survey indicates, that in India, on average, 90% of households report using an improved drinking-water source, with only minor disparities between urban and rural settings (91% vs. 89%, respectively) [24]. Sanitation, however, remains more of an issue, with 48% of households in the country reporting use of an improved sanitation facility. Unlike improved drinking-water source utilisation, there is a large disparity between urban and rural settings with regard to utilisation of improved sanitation facilities. Seventy percent of urban households report using an improved sanitation facility while only 37% of rural household report using an improved facility. Although less than half of the households in the country use an improved sanitation facility, at 48%, this utilisation statistic represents a 19% increase from the third round of the NFHS, conducted during 2005-2006 [24]. The national prevalence of reported diarrhoea in the two weeks preceding the survey was 9%. Stunting and underweight remain issues on a national level, with 38% of children aged less than five years

having a height-for-age z-score two standard deviations below the median reference population, and 36% of children less than five having a weight-for age z-score two standard deviations below the median reference population. There are urban-rural disparities for both of these nutritional outcomes, with children in rural areas doing worse by approximately seven percentage points for both outcomes. While wasting was less prevalent than stunting and underweight, one in five children were still two standard deviations below the median reference population for weight-for height z-score. There was only a slight difference in wasting between urban and rural settings [24].

Odisha State in eastern coastal India is largely rural, with only 17% of the households in the state residing in urban areas. The large majority of people in Odisha are Hindu, with 95% of heads of household classifying themselves as such [24]. Odisha is the third most populous state in terms of its tribal population [25], with 23% of households belonging to a scheduled tribe. Another 20% of households in Odisha belong to a scheduled caste, and 35% to another backward caste [24]. Nearly two-thirds of households (65%) in Odisha practice open defecation, with rural households representing a majority (72%) of this open defecation burden. Improved drinking water in Odisha is similar to the national average, with 89% of households using an improved source of drinking water [24]. Reported diarrhoea prevalence in the two weeks preceding the survey was similar to the national average (10%). Stunting, underweight, and wasting amongst children in Odisha were all similar to national averages. This was the site of the evaluation of a water supply and sanitation intervention, implemented by Gram Vikas, a local non-governmental organisation (NGO). This evaluation was the anchor study of the collective efficacy work conducted in India.

3.4.2 Anchor study details: Gram Vikas MANTRA evaluation

During June 2015 – October 2016, a matched cohort study was carried out in Odisha, India. The aim of the study was to evaluate the effectiveness of the Gram Vikas MANTRA intervention, a combined water supply, sanitation, and hygiene intervention, on behavioural outcomes and health impacts [26]. The study had five objectives, one of which was to develop and employ a theoretically-grounded and evidence-based CE scale to determine whether CE was associated with intervention status, uptake of improved WASH facilities coverage and utilisation behaviours, and nutritional outcomes amongst children aged less than five years. Villages maintain some level of innate CE, or perceptions regarding their ability and autonomy to come together to work toward a common goal, regardless of whether they have been intervened upon with a community-based programme. Therefore, measurement of CE perceptions in

control communities was of interest not only for counterfactual purposes, but to ascertain baseline, innate CE perceptions amongst individuals in villages that were not intervened upon.

I led the CE and cost-effectiveness components of the Gram Vikas evaluation. Further details regarding the design and main effects of the matched cohort study have been described elsewhere [26, 27].

The matched cohort study was designed to retrospectively investigate the effects of a water supply and sanitation intervention several years after its completion. The retrospective study design allowed for an examination of the sustainability of the effects of the intervention. This design was selected because of the team's experience with poorly delivered and adopted interventions during its preceding sanitation trial in Odisha, India [17] and similar intervention fidelity and compliance issues observed by other studies [16]. It was decided that moving forward with a retrospective evaluation of a high quality intervention would facilitate the examination of the health impacts of a water supply and sanitation intervention. It is worth noting, however, that like other observational study designs, a matched cohort study cannot rule out unmeasured confounders between the matched villages.

Sample size determinations were made based on data obtained through simulations estimating the log odds of diarrhoeal disease (the primary study outcome) via a multilevel random effects model parameterised with data from a previous study carried out in Puri, Odisha. Details related to those model parameters are published elsewhere [26], and indicated the need to enroll 45 villages per study arm, and 26 households per village. Consequently, 90 villages were enrolled in the matched cohort study.

Villages that Gram Vikas formally engaged for their MANTRA intervention, with intervention start dates between 2003-2006, were eligible for inclusion in the matched cohort's intervention arm. Forty-five villages were randomly selected for study enrollment amongst those meeting eligibility criteria. A multi-step restriction, matching, and exclusion process was used to match 45 control villages to the randomly selected intervention villages, as detailed elsewhere [26]. Households with a child under five years of age were eligible for enrollment, and up to 40 households per village were enrolled. If there were more than 40 eligible households in a given village, 40 eligible households were randomly selected from the village's census list. Various data were collected across four rounds in randomly selected villages in Ganjam and Gajapati, Odisha during June 2015 – October 2016.

Given the nature of Gram Vikas' village engagement and enrollment process, and the emphasis the programme approach places on cooperative behaviour and collective action, it was deemed necessary to examine factors related to these collective behaviours. This was reflected in the specific objective of the matched cohort study pertaining to the development and validation of a CE scale to ascertain whether CE factors were associated with intervention status and the uptake of improved WASH behaviours. This CE work was prioritised as one of five specific study objectives because the Gram Vikas MANTRA intervention is predicated on a community's ability and willingness to take collective action in order adopt the behaviours and practices promoted by the programme.

3.4.2.1 Details regarding MANTRA programme

Gram Vikas spent over a decade working with participant villages to develop, refine, and pilot its Movement for Action Network for Transformation of Rural Areas (MANTRA) programme. The MANTRA programme ensures village-wide saturation of high-quality, on-site sanitation and hygiene facilities by leveraging a common good – a village-level, improved water distribution system that is piped throughout the village, with each household compound receiving three taps. However, in order to ensure the village receives the communal good, each household in the village – regardless of its wealth status, class, and caste/tribe, must install: 1) a high-quality improved household toilet that meets the specifications set forth by Gram Vikas, with one household tap piped into the facility; and 2) an enclosed bathing room, with an additional tap or shower in the facility. According to Gram Vikas' theory of change, water access is the key to ending open defecation in rural India. This is evident in the emphasis the organisation places on eliminating the need to fetch water outside of the home, which changes the scene for defecation, and interrupts the setting in which open defecation traditionally takes place in India – off-site water sources that allow for post-defecation ritual cleansing.

Gram Vikas will not initiate village enrollment into MANTRA before a village demonstrates its unanimous commitment to the programme. Details regarding the MANTRA programme are presented below, and highlight how factors related to collective behaviour are deeply engrained in the programme's design. MANTRA is implemented in three phases, which are carried out over approximately three years. The programme commences with the Motivational Phase, then continues through to Implementation and Completion Phases.

Phase I: Motivational Phase. The Motivational Phase take place over an average of 8-12 months, during which Gram Vikas staff visit the village to establish contact with village leaders,

Panjiat, and ward members to gauge interest in becoming engaged with the programme and gain high-level community buy-in. After official village leaders are contacted, Gram Vikas staff help facilitate the establishment of a Village General Body (VGB), which is comprised of one man and one woman from every household in the village. Then, the VGB must demonstrate its unanimous commitment to participate in the MANTRA programme. This requires the VGB to meet at least once per month, and demonstrate 100% participation in village meetings. Once all households are in attendance during village meetings, the VGB must establish a contract, or agreement between it and Gram Vikas. This contract enumerates the contributions required by each group of actors (e.g., Gram Vikas, village households), and is signed by each household in the village. The VGB must also register for legal status in accordance with the Society Registration Act of 1860, and establish three village committees. An Executive Committee must be established, and be comprised of ten elected village representatives – five men and five women from different castes/tribes and classes, proportional to village caste/tribe/class distribution, to allow for representation across the village. The final member is one representative from Gram Vikas. In addition to the Executive Committee, Village Water and Village Sanitation Committees should also be established.

The Executive Committee must meet weekly, and is responsible for managing the village's Corpus Fund. Gram Vikas requires that 1,000 Rupees per household be deposited into the village's Corpus Fund. However, households pay into this fund as they are able, and household-specific contributions are decided upon by the Village General Body. During this process, each man and woman from each household in the village have one vote to decide the quantity each household must contribute to the village's Corpus Fund. This allows for wealthier households to contribute more on behalf of other poor or marginalised households, and all village members have the opportunity to weigh into the decision. Every month, during the Motivational Phase, a household contribution is collected by the Executive Committee until the Corpus Fund is complete. The fund is then deposited into a fixed deposit bank account in the name of the Village General Body, with the Executive Director of Gram Vikas serving as a signatory, in an official capacity, on the account. The principle on this account cannot be touched, but the interest can be used by the village to expand the village water distribution system and new sanitation and hygiene facilities for new families in the village. Significantly, only if this village has demonstrated its cooperative behaviour and collective action through these milestones will it be graduated to the Implementation Phase. According to Gram Vikas, a substantial proportion of villages first approached about participating in the MANTRA programme do not complete

the Motivational Phase. It is possible that the villages that do successfully complete the Motivational Phase differ, in terms of CE, from the villages that do not.

Phase II. Implementation Phase. The Implementation Phase, which typically takes place over 18-36 months, consists of the construction of improved, on-site toilets and enclosed, on-site bathing rooms in each household compound as well as the construction of the village-level water tank, tower, and distribution system. Gram Vikas promotes a dual pit toilet with concrete sludge diverter, stone or cement ring-lined pit (one metre deep, two feet below ground, and one foot above ground). The user interface is a pour-flush toilet with one tap in the toilet facility. The sanitation and hygiene block (Figure 3.4) is completed with an enclosed bathing room, typically situated next to the toilet in the same block, and contains one additional tap or shower head. The superstructure of these sanitation and hygiene facilities is comprised of seven-foot walls and a six-foot tall door, with a one-foot gap for light and ventilation. Gram Vikas suggests that the sanitation and hygiene block be within 12-15 feet from the last room in the homestead, yet allows the households to select the siting of the facilities. Gram Vikas also encourages households to plant banana trees over soak pits to mitigate issues with flood/back-logging that may arise during monsoon season. In order to ensure high-quality construction of the facilities, households hire a local, skilled mason to install the toilet and bathing room, but contribute their own unskilled labour and locally available materials to complete the superstructure. Gram Vikas helps procure and provides external materials, including PVC pipes and porcelain pans. This process is supervised and overseen by a Gram Vikas Field Officer.

Figure 3.4 MANTRA-promoted sanitation and hygiene block. Improved, onsite toilet and enclosed, on-site bathing room infrastructure promoted via Gram Vikas' MANTRA programme. ***Photo:*** *Gram Vikas via S. Thomas*

While the village is constructing these household sanitation and hygiene facilities, a water tower and community meeting hall, tank, and piped water distribution system – bringing water to every household in the village – are also being constructed through village level cooperation and collective action (Figure 3.5). During this phase, Gram Vikas also works with and provides capacity building trainings via village-based groups, such as women's self-help groups which focus on microfinance, health, and hygiene trainings; leadership development trainings for elected officials; school-based activities; and masonry (part of Gram Vikas' livelihoods programme) and hardware trainings. All households must have an improved, on-site toilet and enclosed, on-site bathing room before the village enters the final stage of the programme, the Completion Phase.

Figure 3.5 MANTRA village water tower, tank. The village water tower constructed with support from Gram Vikas houses a tank on the top floor, and a village meeting room on the bottom floor.
Photo: *M. Delea*

Figure 3.6 Improved water, piped on-premise. The village-level water distribution system pipes water to three taps in each household. MANTRA requires one tap to be placed in the toilet (left), and one in the enclosed bathing room (centre). The household decides where the third tap is placed, but it is typically located in the kitchen (right).
Photos: *M. Delea*

Phase III. Completion Phase. During the Completion Phase, the water system is turned on. Access to piped water is provided to each household via three taps – one in the toilet, one in the bathing room, and a third in the kitchen or household stand pipe (Figure 3.6).

3.4.2.2 Why examine collective efficacy when evaluating the MANTRA programme intervention?

Gram Vikas spends up to 8-12 months carrying out its Motivational Phase, which serves to target the programme toward villages that have demonstrated early signs of cooperative behaviour and collective action. Therefore, the programming approach employed during the subsequent Implementation Phase by Gram Vikas through MANTRA assumes villages have the

ability and autonomy to set and pursue communal goals that the entire community will contribute toward. Consequently, MANTRA performance targets are predicated on cooperative behaviour and collective action. Unlike other social constructs, such as social capital and social cohesion, CE pertains to and draws on behavioural control perspectives regarding a group's ability to set and pursue common goals as well as group performance and persistence. Therefore, this construct was deemed to be more appropriate to examine via the MANTRA evaluation than alternative social constructs.

While the matched cohort design selected for this study did not allow for causal inference, it did allow for an investigation into whether CE-related factors were independently associated with intervention status. It was also possible to assess CE in control communities to determine innate levels of CE in villages in which the MANTRA programme had not been fully implemented.

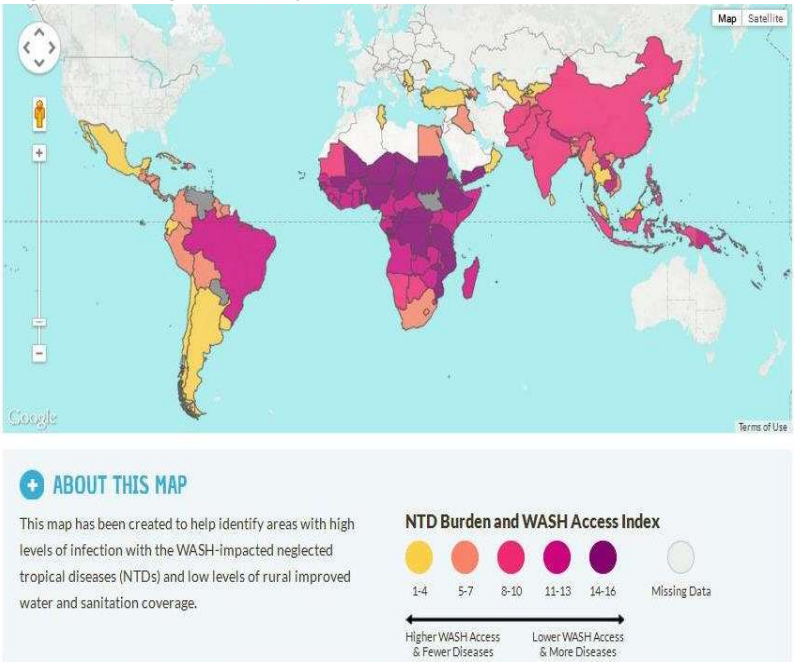
3.5 Sanitation, hygiene, and neglected tropical diseases in Ethiopia

3.5.1 Intervention context – Amhara, Ethiopia

With a population of 107.5 million people, approximately 80% of whom live in peri-urban or rural areas [28], only 6% of all households in Ethiopia use an improved sanitation facility that is not shared; an additional 9% use facilities that are shared, but otherwise improved [29]. Thirty-nine percent of households residing in rural areas have no sanitation facility. The vast majority (94%) of rural households that do use a sanitation facility use an unimproved latrine (55% of which are pit latrines without a slab or open pit). These conditions are not suitable for safely managing faeces, and keeping them isolated from contact with humans and other vectors. When sanitation facilities are in place, they are often not maintained, as needed. A recent outcome evaluation of community-led total sanitation and hygiene (CLTSH), the intervention approach the Government of Ethiopia is scaling nationally, provides evidence of poor operations and maintenance [30]. This evaluation indicated that, amongst the 2,035 households surveyed in 16 districts (*woredas*), only 20% of those in intervention *woredas*, and 21% of households in control *woredas* had repaired their latrines in the past one year.

Water scarcity and resulting water insecurity is a common issue in Ethiopia as well [31]. Over half of Ethiopians living in rural areas spend 30 minutes or longer collecting drinking water [29]. Recent droughts have only exacerbated these conditions. As a result, hygiene practices remain poor in much of rural Ethiopia. These WASH conditions have contributed to enteric infections, NTDs, and malnutrition in Ethiopia. For instance, diarrhoea contributes to more than one in every 10 (13%) child deaths in the country [32]. Enteric infections caused by poor WASH conditions and practices contribute to child malnutrition [14, 33]. In Ethiopia, 38% of children aged less than five years are stunted, meaning they have a height-for-age z-score two standard deviations from the median of the reference population [29]. Another 18% are severely stunted (i.e., height-for-age z-score three standard deviations from the median reference population). Twenty-four percent of Ethiopian children aged less than five years are underweight, meaning they have a weight-for-age z-score two standard deviations below the median reference population [29]. Along with other enteric infections, NTDs such as trachoma, soil-transmitted helminthiasis, and schistosomiasis remain highly prevalent in Ethiopia, with over 74 million people in the country at risk of at least one NTD [34]. See Figure 3.7.

Figure 3.7. Neglected Tropical Diseases and WASH



Source: *Water, sanitation, & Hygiene Online Manual Resource.*

Within Ethiopia, Amhara National Regional State is a region in which WASH conditions are poor and several NTDs are hyper-endemic [35]. Findings from the baseline of our *Andilaye* trial indicate that while progress was made to improve sanitation coverage and utilisation through CLTSH, behavioural slippage (i.e., relapse to unimproved water, sanitation and hygiene [WASH] practices) is common [36]. This is the region in which the *Andilaye* trial, the anchor study of the Ethiopian collective efficacy work, is being conducted.

Data collected in Ethiopia were used to explore the dimensionality of CE, as a latent construct. CE is a potential antecedent of importance for the adoption of cooperative behaviours being promoted by numerous national-level programmes in Ethiopia. Since 2006, the Federal Democratic Republic of Ethiopia's Federal Ministry of Health (FMoH) has been implementing community-based interventions at scale within its national Health Extension Programme (HEP). For example, FMoH has leveraged the HEP, and included a community-led total sanitation and hygiene (CLTSH) intervention as one of the 16 components of its Health Extension Package [37]. While FMoH's CLTSH approach has undergone several iterations of revision, the Ministry has been committed to implementing demand-side, community-based interventions to improve WASH conditions within Ethiopia. Given the results of recent evaluations of CLTSH in Ethiopia [30], our team is in the process of carrying out an ex-ante, parallel cluster-randomised controlled trial evaluating the impact of an enhanced demand-side sanitation and hygiene intervention on sustained behaviour change and health in Amhara, Ethiopia. As a result, there was interest in developing and validating a CE metric that would allow for assessments of CE over time, to determine whether and to what extent there is a causal relationship between CE and the effectiveness of the *Andilaye* intervention, an enhanced community-based, demand-side WASH intervention. Given the timing of the *Andilaye* trial, I am only able to present findings on baseline assessments of CE (see comparison of factor scores in Research Paper 2). However, future analyses of endline data, collected during March-April 2019, will allow an examination of causal attribution via path analyses. This will allow us to test our hypothesis that there is a bi-directional, causal association between CE and intervention effectiveness via path analyses.

3.5.2 Anchor study details: *Andilaye* sanitation and hygiene trial

The *Andilaye* trial is a three-year assessment of the effectiveness of an enhanced, demand-side sanitation and hygiene intervention on sustained behaviour change and health in Amhara, Ethiopia. The purpose of the three-year study is to use formative research findings to inform the integration of NTD-preventive components into existing community-based WASH behaviour change approaches (CLTSH, and the HEP more generally) in Ethiopia, and to evaluate the effectiveness of such an intervention on sustained behaviour change and mental well-being. Monitored WASH-related, NTD-preventive behavioural outcomes are targeted toward specific NTDs of interest, including soil-transmitted helminthiases (STH), intestinal nematode infections spread through contact with larvae-contaminated soil, food, and water; trachoma, an eye infection that can lead to visual impairment, which is caused by the bacterium *Chlamydia*

trachomatis; and schistosomiasis, a trematode infection spread through contact with contaminated stagnant water.

Given my involvement in the design, management, and writing up of the study proposal, field protocols, baseline and midline reports, intervention design content, and other programme-related documents, some of the text I present below is reflected in related study documents. I wrote the majority of the text, however various members of the larger team did contribute and provide feedback, as well.

The *Andilaye* Impact Evaluation is being carried out in Amhara National Regional State, a region of Ethiopia in which WASH conditions are poor, behavioural slippage has been documented, and several NTDs are hyper-endemic. As with the rest of Ethiopia, where CLTSH is being scaled nationally, study communities have either been triggered with CLTSH or are scheduled for triggering in the near future. Despite the absence of key NTD-preventive behavioural WASH promotion, FMoH considers CLTSH its approach for addressing WASH components of NTD programs and promoting other hygiene-related messages for control of enteric diseases.

The study was designed, and is being executed at a time when GoE and FMoH are critically evaluating the nationally scaled HEP. As a result, our intervention design considered demand-side sanitation and hygiene intervention approaches that could be considered as refinements within the HEP if they demonstrated impact. However, we were cognizant of the fact that the Health Extension Package utilised by the HEP has become saturated, and that HEWs are constantly having more work added to their plates via the HEP. As a result, our intervention is exploring the engagement and potential of alternative community change agents for intervention delivery at the community level.

As a result of the national Health Extension Programme and its accompanying initiatives, including CLTSH, the majority (92%, 46 of 50) of sub-district clusters engaged in the *Andilaye* trial have undergone at least some initial engagement via community-based interventions. However, our baseline findings indicated that adoption of behaviours promoted via previous community-based interventions was either unsuccessful or unsustainable [36]. Establishing a CE measurement scale in the early phases of the *Andilaye* trial allowed for the measure and assessment of collective efficacy at baseline, prior to the implementation of a community-level demand-side *Andilaye* intervention.

The *Andilaye* team designed this impact evaluation as a parallel cluster-randomised, controlled trail (RCT, with clusters defined as sub-districts [i.e., *kebeles*]) to allow for causal attribution and quantification of intervention effects. The study assesses and tracks changes in sanitation and hygiene behavioural antecedents (e.g., attitudes; perceived self- and collective efficacy; norms) and behavioural outcomes, water and sanitation insecurity, respondent-reported diarrhoea, and mental well-being at multiple time points. Along with general surveying and observation of household members and WASH facilities, one index child[1] from each enrolled household[2] is being tracked closely, providing additional objective measures of improved NTD-preventive WASH practices throughout the duration of the study. Longitudinal tracking increases study power and reveals seasonal, secular, and other temporal variations in intermediate and behavioural outcomes. Longer follow-up also provides more time to assess key health impacts, and offers a longer perspective for examining externalities, cost-effectiveness, and the sustainability of sanitation and hygiene behaviour change. Below, the intervention is summarised, the control arm is described, and allocation rules are outlined. Finally, further justification is provided for the examination of CE in the context of this trial.

3.5.2.1 Details regarding the Andilaye intervention – study arm intervention exposure

Andilaye, Amharic for "togetherness", encompasses the sentiment of the theoretically-informed and evidence-based intervention developed as part of the *Andilaye* Impact Evaluation. The *Andilaye* intervention is a demand-side sanitation and hygiene intervention that incorporates NTD-preventive behavioural promotion; focuses on positive, community-oriented motivators of behavioural change; promotes achievable incremental improvements; and incorporates strategies that facilitate behavioural maintenance (i.e., prevention of behavioural slippage or relapse back to unimproved behaviours). The intervention addresses issues related to over-extension of health extension workers HEWs, and over-saturation of HEP messaging through the engagement of additional community change agents as mechanisms for intervention delivery. The intervention is based on recent work conducted by WSP, and incorporates feedback from relevant stakeholders, including FMoH, ARHB, Zonal Health Departments, *Woreda* Health Offices, WSP, and other key stakeholders, including formative research community members.

[1] 'Index child' standard operational definition: The youngest child in the household who is greater than one year of age, but less than ten years of age.

[2] 'Household' standard operational definition: A person or group of related or unrelated persons who usually live together in the same dwelling unit(s), who have common cooking and eating arrangements, and who acknowledge one head of household.

Andilaye intervention functions were informed by theory and designed to address key drivers and barriers identified from the problem and solution trees derived from our formative research data and stakeholder-generated interpretations thereof. Through a formal intervention mapping process, we identified theoretically-grounded and evidence-based intervention techniques specifically designed to address the behavioural antecedents and determinants represented by the problem and solution tress. Our formative work also indicated that perceptions related to self- and collective efficacy are important mediators of uptake of community-based interventions. As such, the overarching intervention motto, *Together, we can be a strong, caring, healthy community*, and related intervention functions work to improve individual and community-level agency (i.e., enhance self- and collective efficacy). The motto offers an aspirational message that emphasises the need for collective action to make positive change in one's community.

3.5.2.2 Description of the control arm

The comparison group will receive current standard of care sanitation and hygiene programming (i.e., interventions related to FMoH's current CLTSH model). Any other intervention in comparison communities is being limited, and we are working with government partners to ensure this is the case. Table 4.1 outlines differences between the *Andilaye* intervention and the current CLTSH model.

Table 3.1 Sampling, per study arm – *Andilaye* trial

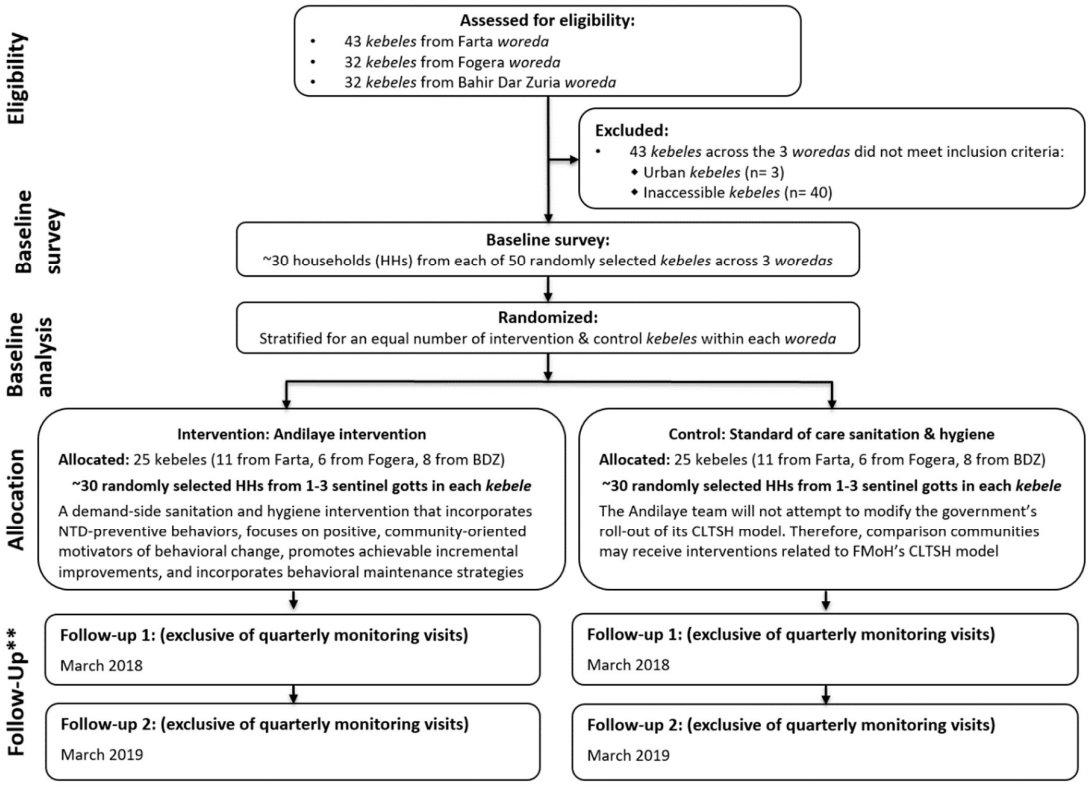
INTERVENTION ARM	CONTROL ARM
***Andilaye* intervention**	**Standard of care sanitation & hygiene (i.e., CLTSH)**
1-3 sentinel *gott*(s) from each of 25 randomly selected *kebeles*, ~30 randomly selected households enrolled in each *kebele*	**1-3 sentinel *gott*(s) from each of 25 randomly selected *kebeles*, ~30 randomly selected households enrolled in each *kebele***
A demand-side sanitation and hygiene intervention that incorporates NTD-preventive behaviours, focuses on positive, community-oriented motivators of behavioural change, promotes achievable incremental improvements, and incorporates strategies that facilitate behavioural maintenance (i.e., prevention of behavioural slippage or relapse back to unimproved behaviours).	The *Andilaye* team will not attempt to modify the government's roll-out of its CLTSH model. Therefore, comparison communities may receive interventions related to FMoH's CLTSH model. We will work with our government partners to minimize other WASH interventions in our comparison communities, to the greatest extent possible.

3.5.2.3 Randomisation: allocation rule for intervention and comparison groups

Following baseline data collection, we used a stratified random design to assign study *kebeles* to either the intervention or comparison arm. Within each stratum (*woreda*), we used a random number generator to generate a random number between zero and one for each *kebele* (cluster), and then ordered each *kebele* by the randomly generated number in ascending order. We then partitioned the communities within each *woreda* into two equal sizes, assigning the

first half of *kebeles* to the intervention arm and the second half to the control arm. We used replacement randomization [38] to secure balance across three key variables (latrine coverage, washing station with soap coverage, and head of household education). Because cluster-randomised trials, particularly trials with a small number of clusters, often have individual-level imbalances between arms, we established *a priori* that the intervention and control mean values for these three variables should be within two standard deviations of the overall variable means. The randomisation process described above was repeated (twice) until these variables were balanced according to that *a priori* criterion. Figure 3.8 provides a summary of the study flow.

Figure 3.8 Flow chart indicating *kebele* eligibility, randomisation, allocation, and follow-up



While allocation occurred at the *kebele* level, we are only collecting data in one to two sentinel *gotts* per *kebele*, purposively selected to minimize spillover. The number of sentinel *gotts* per *kebele* in which data are collected depended solely on the number of eligible and consenting households (i.e., if less than 30 eligible and consenting households were present in one *gott*, data collection took place in a second sentinel *gott* as well in order to obtain the minimal required sample size per *kebele* cluster). Monitoring occurred on a quarterly basis.

After enrollment and baseline survey administration, we randomly allocated *kebeles* to either the intervention arm to receive the intervention, implemented in collaboration with the *Woreda* Health Offices, HEWs, HDAs, and other community change agents, or to the counterfactual (control) arm to receive "standard of care" (current sanitation and hygiene programming, including potential roll-out of the current CLTSH intervention in *woredas* and *kebeles* in Amhara per the existing FMoH implementation protocol). As this study is operating in an area where CLTSH is being rolled out nationally, we are not interfering with established CLTSH roll-out and implementation protocols. While we cannot be sure that our comparison communities will not receive further CLTSH interventions during the course of the trial, any such further implementation of the current CLTSH interventions would only bias effect estimates toward the null. Communities allocated to the comparison arm may receive the *Andilaye* intervention at a later time point.

3.5.2.4 Why examine collective efficacy when evaluating the Andilaye intervention?

We are examining collective efficacy under the auspices of the *Andilaye* trial as part of our examination of behavioural antecedents and other behavioural mediators of *Andilaye* intervention uptake. This inquiry will allow an investigation of causal determinants of behavioural change and maintenance over time, and identification of mechanisms that lead to behavioural change. There is utility in employing such an approach, as it will help determine not only *if*, but *how*, and *why* the intervention was successful, and in which contexts.

As indicated above, our formative work indicated that behavioural control perceptions related to self- and collective efficacy were mediators of prior community-based interventions (e.g., CLTSH). The *Andilaye* motto leverages positive messaging about the community's capacity to engender change, and several of its intervention functions focus on improving behavioural control perceptions at the community (i.e., collective efficacy) and individual (i.e., self-efficacy) levels.

3.6 Chapter references

1. World Health Organization. *Diarrhoeal disease fact sheet*. 2 May 2017; Available from: http://www.who.int/en/news-room/fact-sheets/detail/diarrhoeal-disease
2. *Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015.* Lancet, 2016. **388**(10053): p. 1459-1544.
3. Clasen, T.F., K.T. Alexander, D. Sinclair, S. Boisson, R. Peletz, H.H. Chang, F. Majorin, and S. Cairncross, *Interventions to improve water quality for preventing diarrhoea.* Cochrane Database of Systematic Reviews, 2015(10).
4. Ejemot-Nwadiaro, R.I., J.E. Ehiri, D. Arikpo, M.M. Meremikwu, and J.A. Critchley, *Hand washing promotion for preventing diarrhoea.* Cochrane Database of Systematic Reviews, 2015(9).
5. Jung, Y.T., R.J. Hum, W. Lou, and Y.-L. Cheng, *Effects of neighbourhood and household sanitation conditions on diarrhea morbidity: Systematic review and meta-analysis.* PLOS ONE, 2017. **12**(3): p. e0173808.
6. Prüss-Ustün, A., J. Bartram, T. Clasen, J.M. Colford, O. Cumming, V. Curtis, S. Bonjour, A.D. Dangour, J. De France, L. Fewtrell, M.C. Freeman, B. Gordon, P.R. Hunter, R.B. Johnston, C. Mathers, D. Mäusezahl, K. Medlicott, M. Neira, M. Stocks, J. Wolf, and S. Cairncross, *Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries.* Tropical Medicine & International Health, 2014. **19**(8): p. 894-905.
7. Strunz, E.C., D.G. Addiss, M.E. Stocks, S. Ogden, J. Utzinger, and M.C. Freeman, *Water, Sanitation, Hygiene, and Soil-Transmitted Helminth Infection: A Systematic Review and Meta-Analysis.* PLOS Medicine, 2014. **11**(3): p. e1001620.
8. Grimes, J.E.T., D. Croll, W.E. Harrison, J. Utzinger, M.C. Freeman, and M.R. Templeton, *The roles of water, sanitation and hygiene in reducing schistosomiasis: a review.* Parasites & Vectors, 2015. **8**(1): p. 156.
9. Stocks, M.E., S. Ogden, D. Haddad, D.G. Addiss, C. McGuire, and M.C. Freeman, *Effect of Water, Sanitation, and Hygiene on the Prevention of Trachoma: A Systematic Review and Meta-Analysis.* PLOS Medicine, 2014. **11**(2): p. e1001605.
10. Wagner, E.G. and J.N. Lanoix, *Excreta disposal for rural areas and small communities*, in *World Health Organization monograph series no. 39*. 1958, World Health Organization: Geneva, Switzerland.
11. Sowa, S., J. Sowa, L.H. Collier, and W. Blyth, *Trachoma and allied infections in a Gambian village.* Spec Rep Ser Med Res Counc (G B), 1965. **308**: p. 1-88.
12. West, S., B. Munoz, M. Lynch, A. Kayongoya, Z. Chilangwa, B.B. Mmbaga, and H.R. Taylor, *Impact of face-washing on trachoma in Kongwa, Tanzania.* Lancet, 1995. **345**(8943): p. 155-8.
13. Wolf, J., P.R. Hunter, M.C. Freeman, O. Cumming, T. Clasen, J. Bartram, J.P.T. Higgins, R. Johnston, K. Medlicott, S. Boisson, and A. Prüss-Ustün, *Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression.* Tropical Medicine & International Health, 2018. **23**(5): p. 508-525.
14. Freeman, M.C., J.V. Garn, G.D. Sclar, S. Boisson, K. Medlicott, K.T. Alexander, G. Penakalapati, D. Anderson, A.G. Mahtani, J.E.T. Grimes, E.A. Rehfuess, and T.F. Clasen, *The impact of sanitation on infectious disease and nutritional status: A systematic review and meta-analysis.* International Journal of Hygiene and Environmental Health, 2017. **220**(6): p. 928-949.
15. Luby, S.P., M. Rahman, B.F. Arnold, L. Unicomb, S. Ashraf, P.J. Winch, C.P. Stewart, F. Begum, F. Hussain, J. Benjamin-Chung, E. Leontsini, A.M. Naser, S.M. Parvez, A.E. Hubbard, A. Lin, F.A. Nizame, K. Jannat, A. Ercumen, P.K. Ram, K.K. Das, J. Abedin, T.F.

Clasen, K.G. Dewey, L.C. Fernald, C. Null, T. Ahmed, and J.M. Colford, Jr., *Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial.* The Lancet Global Health, 2018. **6**(3): p. e302-e315.

16. Null, C., C.P. Stewart, A.J. Pickering, H.N. Dentz, B.F. Arnold, C.D. Arnold, J. Benjamin-Chung, T. Clasen, K.G. Dewey, L.C.H. Fernald, A.E. Hubbard, P. Kariger, A. Lin, S.P. Luby, A. Mertens, S.M. Njenga, G. Nyambane, P.K. Ram, and J.M. Colford, Jr., *Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial.* The Lancet Global Health, 2018. **6**(3): p. e316-e329.
17. Clasen, T., S. Boisson, P. Routray, B. Torondel, M. Bell, O. Cumming, J. Ensink, M. Freeman, M. Jenkins, M. Odagiri, S. Ray, A. Sinha, M. Suar, and W.-P. Schmidt, *Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial.* The Lancet Global Health, 2014. **2**(11): p. e645-e653.
18. Carter, R.C., *Can and should sanitation and hygiene programmes be expected to achieve health impacts?* Waterlines, 2017. **36**(1): p. 92-103.
19. Cumming, O. and V. Curtis, *Implications of WASH Benefits trials for water and sanitation.* The Lancet Global Health, 2018. **6**(6): p. e613-e614.
20. Fuller, J.A. and J.N. Eisenberg, *Herd Protection from Drinking Water, Sanitation, and Hygiene Interventions.* Am J Trop Med Hyg, 2016. **95**(5): p. 1201-1210.
21. Garn, J.V., S. Boisson, R. Willis, A. Bakhtiari, T. al-Khatib, K. Amer, W. Batcho, P. Courtright, M. Dejene, A. Goepogui, K. Kalua, B. Kebede, C.K. Macleod, K.I.M. Madeleine, M.S.A. Mbofana, C. Mpyet, J. Ndjemba, N. Olobio, A.L. Pavluck, O. Sokana, K. Southisombath, F. Taleo, A.W. Solomon, and M.C. Freeman, *Sanitation and water supply coverage thresholds associated with active trachoma: Modeling cross-sectional data from 13 countries.* PLOS Neglected Tropical Diseases, 2018. **12**(1): p. e0006110.
22. Oswald, W.E., A.E. Stewart, M.R. Kramer, T. Endeshaw, M. Zerihun, B. Melak, E. Sata, D. Gessese, T. Teferi, Z. Tadesse, B. Guadie, J.D. King, P.M. Emerson, E.K. Callahan, D. Flanders, C.L. Moe, and T.F. Clasen, *Active trachoma and community use of sanitation, Ethiopia.* Bulletin of the World Health Organization, 2017. **95**(4): p. 250-260.
23. *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. 2017, World Health Organization (WHO) and the United Nations Children's Fund (UNICEF): Geneva, Switzerland.
24. International Institute for Population Sciences (IIPS), *National Family Health Survey (NFHS-4), 2015-16: India*. 2017, IIPS: Mumbai.
25. ST & SC Development, M., and Backward Classes Welfare Department, . 2018; Available from: http://stscodisha.gov.in/
26. Reese, H., P. Routray, B. Torondel, G. Sclar, M.G. Delea, S.S. Sinharoy, L. Zambrano, B. Caruso, S.R. Mishra, H.H. Chang, and T. Clasen, *Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India.* BMJ Open, 2017. **7**(3).
27. Reese, H., P. Routray, B. Torondel, S.S. Sinharoy, H.H. Chang, and T.F. Clasen, *Effect of a combined household-level piped water and sanitation intervention on child diarrhea, acute respiratory infection, soil-transmitted helminth infection, and nutritional status: a matched cohort study in rural Odisha, India.* Bulletin of the World Health Organization, 2018. **Submitted**.
28. *World Population Prospects Revision*. 2017, United Nations DESA/Population Division: Geneva, Switzerland.

29. Central Statistical Agency (CSA) [Ethiopia] and ICF, *Ethiopia Demographic and Health Survey 2016*, in *Key Indicators Report*. 2016, CSA and ICF: Addis Ababa, Ethiopia and Rockville, Maryland.
30. BDS - Center for Development Research, *Outcome evaluation of community-led total sanitation and hygiene (CLTSH) program in Ethiopia 2012-2015*. 2016, BDS: Addis, Ababa.
31. Stevenson, E.G.J., L.E. Greene, K.C. Maes, A. Ambelu, Y.A. Tesfaye, R. Rheingans, and C. Hadley, *Water insecurity in 3 dimensions: An anthropological perspective on water and women's psychosocial distress in Ethiopia.* Social Science & Medicine, 2012. **75**(2): p. 392-400.
32. World Health Organization, *CHERG-WHO Methods and Data Sources for Child Causes of Death 2000-2012*, in *Global Health Estimates Technical Paper*. 2014, WHO/HIS/HSI/GHE: Geneva, Switzerland.
33. Black, R.E., L.H. Allen, Z.A. Bhutta, L.E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, and J. Rivera, *Maternal and child undernutrition: global and regional exposures and health consequences.* The Lancet, 2008. **371**(9608): p. 243-260.
34. World Health Organization Regional Office for Africa, *The Expanded Special Project for elimination of Neglected Tropical Diseases (ESPEN)*, in *2017 Annual Report*. 2018, World Health Organization Regional Office for Africa: Barzzaville, Congo.
35. International Trachoma Initiative. *Trachoma Atlas*. 2018; Available from: http://www.trachomaatlas.org/.
36. Garn, J.V., M. Delea, and M. Freeman, *Baseline report: The impact of enhanced, demand-side sanitation and hygiene promotion on sustained behavior change and health in Amhara, Ethiopia*. 2017, Emory University: Atlanta, GA.
37. Federal Ministry of Health (FMoH), *National Health Extension Program*. 2003, FMoH: Addis Ababa, Ethiopia.
38. Lachin, J.M., *Statistical properties of randomization in clinical trials.* Controlled Clinical Trials, 1988. **9**(4): p. 289-311.

Chapter 4: Collective efficacy: Development and validation of a measurement scale

4.1 Chapter overview

Context-specific collective efficacy scales were developed for both the Gram Vikas MANTRA impact evaluation in Odisha, India, and the *Andilaye* trial in Amhara, Ethiopia. This chapter presents information related to the five-step, sequential exploratory and confirmatory approach used to develop, refine, and validate all theoretically-grounded and evidence-based collective efficacy scales. Research paper 2 goes into detail regarding this process, highlighting work from the *Andilaye* trial. As previously mentioned, data captured during the *Andilaye* trial were highlighted in Research Paper 2 given the dataset was more nuanced. The following circumstances contributed to this: 1) donors funded a formal six-month formative research phase, during which additional time and resources were dedicated to examine additional CE factors; 2) a more comprehensive survey was used; and 3) lessons learnt from the Gram Vikas MANTRA work in India were applied to the *Andilaye* work. Details related to the development, refinement, and validation of the collective efficacy scale developed for the rural Odisha context are presented in Chapter 5.

4.2 Framing of research questions

Before assessing whether and to what extent CE is an antecedent of collective behaviour, it was necessary to develop, refine, and validate a CE scale. Therefore, the aim of the research presented in this chapter was to develop a metric to quantitatively assess CE in a manner that would allow for future assessment of related factors as potential antecedents of collective behaviour, and uptake of improved WASH practices. To achieve this aim, a systematic and widely accepted scale development process [1] was adapted and used.

4.2.1 Research Paper 2 objectives

The primary objective of Research Paper 2 was to elucidate the underlying structure of CE through the development, refinement, and validation of CE measurement metrics. For the particular study highlighted in the research paper presented in this chapter (i.e., the Ethiopia *Andilaye* CE study), CE measurement models were created to assess pre-intervention behavioural control perceptions at baseline to evaluate the effect of CE factors on the uptake of the *Andilaye* intervention at endline.

The specific objectives of this research were to:

1. Examine the underlying structure of collective efficacy and its constituent sub-constructs;
2. Ascertain whether there was evidence to support the hypothesis that CE could be measured quantitatively with high construct validity; and
3. Determine whether there were differences in the measurement of CE between various sub-groups.

4.2.2 Research questions

Three specific research questions were investigated in this study:

RQ1. Which sub-constructs (e.g., constituent domains, factors/dimensions, and facets) are salient for measuring collective efficacy in the rural Ethiopia context?

RQ1a. Do the factor solutions generated by our CE scale support our theorised conceptualisation of CE (i.e., do the statistics generated by the factor solutions support our theorised construct dimensionality)?

RQ1b. If alternative factor solutions emerge from EFA-derived solutions, are they substantiated by existing theory and evidence?

RQ2. Are the psychometric characteristics of the resulting CE measurement models compelling in terms of their ability to demonstrate construct validity, or the degree to which the scale measures what it purports to measure?

RQ3. Are there important differences in the measurement of CE between Ethiopian men and women, and households in which an individual with leadership status in a community structure resides vs. those in which no leaders reside?

4.2.3 Related hypotheses

In the context of these research questions, the following hypotheses were tested:

Hypothesis 1: The statistics generated by the factor solutions derived from our empirical data will support our theorised construct dimensionality

Hypothesis 2: The CE measurement models produced via factor analytic methods will demonstrate high construct validity

Hypothesis 3: Given their status, mobilisation, and inclusion within their communities, men will have higher perceptions of behavioural control and related factors than women, leaders than non-leaders

4.3 Summary of collective efficacy scale development process – From India to Ethiopia

A five-step sequential exploratory approach was used as the general scale development, refinement, and validation process for all collective efficacy scales produced via this thesis. Justification for the selection of this process is present in section *4.4 Development of methodology*. Research Paper 2 provides a detailed account of the methodology employed. Below, I provide a brief summary of the process and each step, and explain how and why context-specific CE scales were developed from data collected via the two larger anchor studies indicated in Chapter 3 (i.e., the Gram Vikas MANTRA evaluation and the *Andilaye* trial).

4.3.1 General scale development process

Figure 4.1 provides a visual representation of the scale development, refinement, and validation process.

Figure 4.1 Five-step CE scale development, refinement, and validation approach

Adapted from Spector 1992

In summary, the first phase of the process entails defining the CE construct. This includes a review of the theoretical and empirical literature, and established frameworks related to various social constructs to identify the different concepts (i.e., sub-constructs) used to define the various constructs [1, 2]. Sub-constructs were then grouped into their respective dimensions, and the various dimensions were grouped into three higher-level domains. These various groups of sub-constructs were then pieced together to create a hypothesised CE framework.

Next, the CE survey instrument was designed. In order to develop a first draft of the CE survey instrument, an item pool was established from existing social construct surveys and scales. This included items from The World Bank Group's Social Capital Assessment Tool [3], Catholic Relief Services Social Capital Survey, CARE's WE-MEASR women's empowerment tool, and items from the collective efficacy tool developed by Sampson and colleagues [4]. Once the item pool was established, we coded items under each CE dimeson, removed repetitive items, designed new items to fill gaps under certain dimensions, and re-structured all survey prompts to ensure they worked well with a 5-point Likert-type response format [1].

Once we had a draft CE survey in hand, it was necessary to conduct cognitive, in-depth interviews [5] to gauge the participant's comprehension and understanding of each survey item, or prompt. We employed 'think-aloud' and probing techniques to obtain this information related to the participants' understanding of the meaning of the survey items (i.e., assess face validity) and general feedback regarding the appropriateness of context-specific examples. This process helped ensure that each survey item tapped to the desired sub-construct within our framework, and allowed us to determine whether there was early evidence regarding the substantive aspects of construct validity [6].

This cognitive validation process was performed for each of the two distinct studies. In India, we conducted ten cognitive interviews in both intervention and control villages. The interviews were conducted in Oriya, the local dialect, by trained research assistants. They were audio-recorded and transcribed directly into English. Field notes, debriefings, and team discussions resulted in daily refinements, which were tested during subsequent interviews conducted during subsequent days. An applied version of content analysis [7] guided the refinement process. In Ethiopia, an iterative cognitive validation process was carried out, during which an initial round of four cognitive interviews (three women, one man) were carried out. A second phase of cognitive interviews (four women, three men) was subsequently carried out to pilot the revised survey.

After refined, context-specific surveys were developed, they were employed in each respective study. See Research Papers 2 (Ethiopia) and 3 (India) for details related to survey administration. Steps 4 and 5 of the scale development, refinement, and validation process comprise the analytical phase of the process. As indicated in section 4.4 and in Research Papers 2 and 3, a factor analytic approach, comprised of exploratory and confirmatory factor analyses, was selected as the psychometric analytical methods. Hypotheses regarding men's and women's factor scores were tested via Multiple Indicator Multiple Causes (MIMIC) modelling. Further details related to these methods are provided below and in Research Paper 2.

The CE inquiry carried out for this thesis initiated in India. Lessons learnt through that process were applied to the Ethiopia CE study. Additional formative work was carried out in Ethiopia as well. This additional formative work consisted of four focus group discussions (two with women, two with men) conducted to explore concepts related to our hypothesised CE framework and real-life examples thereof.

4.3.2 Development of context-specific scales

Colleagues developing water insecurity scales have advocated for the development of context-specific scales [8-10]. Like them, those of us who worked on the development of the CE scale presented in Research Papers 2 and 3 recognise that collective behavioural factors are often highly dependent on the local context. Socially-influenced CE sub-constructs may be influenced largely by context-specific political economies and social schemas. For instance, factors related to empirical and normative expectations regarding cooperative behaviour that likely inform perceptions about agentic concepts such as self- and collective efficacy are steeped in rich historical and cultural traditions that may not be comparable on a global scale.

Therefore, instead of developing a CE scale in one context (e.g., India) and carrying it over to the other (e.g., Ethiopia), I discussed with my colleagues, and we agreed to create two separate, context-specific scales. I believe this is an important decision to have made in the infancy of these CE scales. There was considerable overlap in the surveys and scales between contexts, and even work conducted by a Master's student in Cambodia who I helped advise indicated numerous commonalities with the scales developed in India and Ethiopia. As work on CE moves forward and more context-specific scales are developed, we will have more information to determine whether a standardised, global CE scale can maintain high construct validity.

Despite the preference for context-specific scales, the 26-item parsimonious scale presented in Research Paper 2, or the longer 50-item survey presented therein can certainly serve as the foundation for a survey that programme implementers and researchers can use presently. When the CE scales put forward in this paper are used, they must be adapted to the local context. Items related to endogenous and exogenous community structures should be contextualised, as should other tangible examples presented in some of the prompts. I also highly encourage cognitive validation of any tool prior to broad scale deployment.

4.4 Development of methodology

The five-step scale development process utilised for this thesis work was adapted from Paul E. Spector's summated rating scale construction process [1]. Spector's scale development process was selected as a guide for this work because it provides a clear, step-wise approach that emphasises the need to create a specific construct definition and establish substantive justification during the first two steps of the process. These steps were viewed as desirable because they would allow the team to develop a theoretically-grounded hypothesised CE framework that could later be tested via a factor analytical approach. Spector also underscores

the need to have respondents ”critique the scale” [1]. In other words, he emphasises the need to perform cognitive validation to ensure consistency between the intended meaning of the survey items and participants' understanding regarding what the survey items convey (i.e., assess face validity). This resonated with the scale development team given we felt strongly that formative work was necessary prior to large scale deployment of the draft survey instrument. This formative work would allow us to: 1) ensure all survey items were locally appropriate, 2) give a voice to individuals similar to our study participants, and 3) allow them to provide feedback on the survey item prompts. This process would help confirm that each survey item tapped to the desired sub-construct within the hypothesised CE framework. While there are several methodologists who stress the importance of establishing substantive justification of the hypothesised framework [11] and performing cognitive validation [12], some methodologists focus more on the psychometric analysis steps of the scale development and validation process, and often under-emphasise these early formative steps.

Below, additional details regarding specific methodological decisions are presented to provide justification for more specific methodological decisions made during the scale development, refinement, and validation process.

4.4.1 Psychometric analyses and collective efficacy scale development

Several psychometric methods are available to facilitate the measurement of latent constructs that cannot be directly observed or measured. Prior to the execution of psychometric analyses, several different methods were considered, and a factor analytic approach was selected for the purposes of this research. Below, a summary of factor analysis is provided along with a brief explanation regarding the rationale for selecting a factor analytic approach over other methods.

4.4.1.1 Factor analysis as a positivist approach for CE scale development and validation

Factor analysis was selected as the psychometric method for scale development for this body of work. In a factor analytic framework, latent constructs are measured through the analysis of related manifest, observed variables (e.g., survey items), or indicators that represent certain aspects of the latent construct [13]. Factor analyses represent multivariable statistical procedures that aim to elucidate the number of latent factors and the relationships between those factors that explain the variation and covariation of the manifest variables measured through methods such as surveys [11]. Measurement models are used to estimate the relationships between observed measures and latent factors. A typical measurement model is the common factor model, which assumes that each indicator in a set of observed measures

taps to, or reflects an underlying latent factor, or more specifically, a linear function of one or more common factors and one unique factor [11, 14, 15].

As set out in work presented by Bollen and summarised by Steinmetz [14]:

> Latent variables explain the variance in a given indicator and the covariance amongst a set of items that tap to that latent factor. This relationship is expressed mathematically through the following common factor equation:

$$\boldsymbol{X_i = \lambda_i \xi + \delta_i} \qquad \textbf{(4.1)}$$

> *where the observed indicator item X_i represents the ith item in a survey instrument, and is explained by a latent variable ξ, which influences the outcome with factor loading λ_i, and is affected by a measurement error δ_i [11, 14].*

> Spearman proposed this unidimensional (i.e., single-factor) model, which assumes all observed item indicators measure one common factor [14]. When the equation is extended to include location parameters, (e.g., indicator intercepts [τ_i]), factor loadings (λ_i) represent regression coefficients that indicate the pattern of item-factor relationships [11] with the latent variable ξ, with residual δ_i [14].

$$\boldsymbol{E(X_i) = \tau_i + \lambda_i E(\xi) + E(\delta_i)} \qquad \textbf{(4.2)}$$

> Thurstone then extended the unidimensional model, and proposed a multidimensional (i.e., multiple-factor) model [15]. This model structure expresses the variation and covariation amongst observed item indicators as a function of factors and residuals:

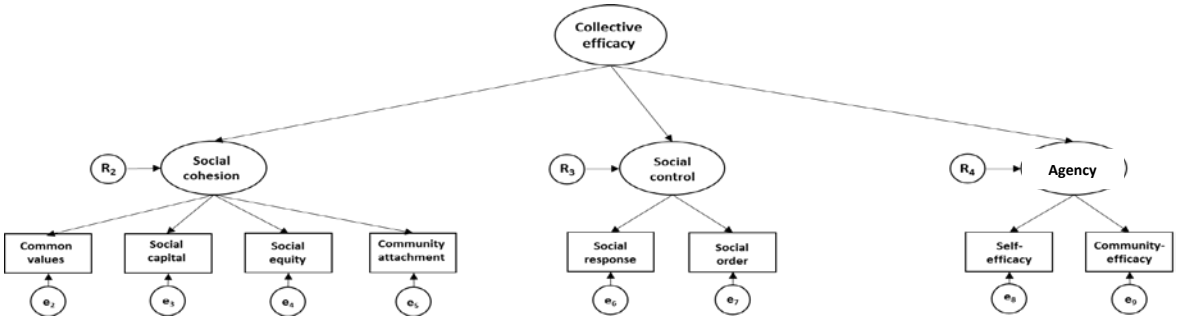
$$\boldsymbol{y_{i1} = v_1 + \lambda_{11}\eta_{i1} + \lambda_{12}\eta_{i2} + \cdots + \lambda_{1m}\eta_{im} + \varepsilon_{i1}} \qquad \textbf{(4.3)}$$

> *where v_j represents the intercept, λ_{jk} are factor loadings, η_{ik} are factor values, and ε_{ij} are residuals with zero means and correlations of zero with the factor [16].*

In accordance with a positivist approach, we turned to a multidimensional factor analytic approach to develop a measurement scale for collective efficacy, and examine its construct validity [17]. While our CE scale development approach was largely positivist, we did incorporate qualitative research components into our larger examination of CE in order to assess the cognitive validity of our scale, iteratively refine it prior to administration, and interpret the quantitative results (e.g., factor loadings, factor scores).

There are two main types of analyses that are based on the common factor model – exploratory factor analysis (EFA) and confirmatory factor analyses (CFA) [11]. These types of analytical methods utilise matrix algebra to generate statistics used to reveal the underlying factor structure related to a latent construct [18]. In the factor analytic context, a factor is a mathematic descriptor summarising the relationships that exist amongst a set of items. Each factor is represented by a vector in a geometric space that contains as many dimensions as factors revealed in the relevant factor solution [13]. Figure 4.2 provides a visual depiction of this, using the initial hypothesised, pre-analysis collective efficacy factor structure.

Figure 4.2 Hypothesised collective efficacy model framework



As its name indicates, EFA is an exploratory, descriptive analytical method employed to determine the number of common factors in a measurement model, and identify which measured variables are reasonable indicators of the latent construct of interest [11]. The goal of EFA is to reduce the set of input indicators required to measure the latent construct of interest [11]. EFA represents an assessment of a correlation or covariance matrix. The relevant matrix is examined in order to assess the size and differential magnitude of factor loadings to make determinations regarding a latent construct's underlying (i.e., factor) structure (i.e., identify indicator items, with salient loadings, that are important for the measure of the latent construct).

CFA, on the other hand, is an evaluation approach that allows for direct testing of hypothesised factor solutions to assess the appropriateness of these factor structures (i.e., determine construct validity, falsify hypothesised models). This method is often used to validate a pre-specified factor solution by testing how well the designated number of factors and pattern of indicator-factor loadings reproduces the sample matrix of the measured variables [11]. While the method cannot prove that a measurement model is "correct", it can provide evidence that indicates a failure to reject or falsify the hypothesised model. In other words, CFA can be used to validate a measurement scale, or more specifically, assess construct validity (i.e., determine

whether the hypothesised scale measures what it intends to measure). It also has the ability to estimate the relationships amongst variables while adjusting for measurement error [11].

4.4.1.2 Rationale for selection of a factor analytic approach

There are three main psychometric measurement theories, which have produced various methods. These theories are outlined and compared below.

Classical test theory (CTT). CTT, also known as "true score theory" [19] is based on the presumption that the measurement of a latent construct requires only that numbers are assigned according to some rule. For example, CTT models use covariance or correlation matrices to derive test scores given the theory assumes that the raw scores generated from test assessments represent measurements of the latent construct [19].

This theory is concerned with overall test scores, which reflect true scores, and assumes that a score derived from an assessment of a latent construct is the sum of an error-free "true score" and an error score. While CTT recognises the presence of errors in measurements, and conceptualises error as a random variable, respondent and test characteristics cannot be disentangled from each other, but rather, they can only be assessed in the context of each other [20]. Some might argue that a lack of consensus regarding cut points for factor retention, and heavy or exclusive reliance on mathematically-based Kaiser-Guttman rule (i.e., eigenvalue > 1.0) is another limitation of the theory [21]. Despite these limitations, CTT represents a long-standing, and well-recognised psychometric theory and related analytical approach. Item response and Generalisable theory represent more recent and sophisticated models, yet have their own uses and limitations that should be considered before one settles on a methodological approach for psychometric analyses.

CTT includes linear modelling approaches, such as the suite of analytical methods that comprise factor analyses. Factor analyses represent procedures that utilise covariance and correlation matrices to explain common variance in order to evaluate the extent to which items measure a latent factor. They are the methods used to reveal the underlying dimensions emerging from the data. The purpose of the factor analytic approach is to determine which latent factors create commonality in the data [22].

Why factor analysis and not principal components analysis (PCA)? PCA and EFA are related, but distinct [23] in that rely on a different set of quantitative methods. Unlike factor analytic methods, PCA methods are not based on the common factor model [11]. PCA does not

differentiate between common and unique variance. Instead, it seeks to account for the variance in observed measures as opposed to explaining the correlations amongst them [11]. Therefore, PCA is better suited for item or data reduction, while EFA is better suited to reveal and explain the underlying structure, or relationships between items and factors. While the literature will indicate that PCA is a method used by some to develop scales, the scales developed through this approach "are typically not very good from a psychometric point of view" given PCA may not account for a considerable amount of variance for a factor [23].

Item response theory (IRT). IRT is comprised of a group of measurement models used to explain the connection between observed item responses on a scale and an underlying construct. With IRT, mathematical equations that employ a non-linear monotonic function are used to describe associations between a subject's given level of the latent construct and the probability of a particular response being selected [22]. This theory allows for joint scaling of the respondent and the survey item.

IRT is concerned with item scores and ordering items along a latent trait, as opposed to being concerned with test scores and explaining variance, which is the case with CTT. IRT modelling approaches include non-linear modelling approaches [20]. Rasch, Mokken, and graded response modelling, are some IRT methods. Rasch and Mokken models are both special cases of item response models. Rasch scales represent models in which the item parameter represents the difficulty of the items, and the person parameter represents the ability or attainment level of the respondent. This type of model assumes all items have the same response function, meaning they are equivalent in terms of discrimination, and therefore result in identical factor loadings for all items. Rasch modelling obtains data that fit a specified model, whereas other statistical modelling approaches modify, accept, and reject parameters based on how well they fit the data [24]. Mokken models represent unidimensional scales that consists of items ordered according to their difficulty, which are intended to measure the same underlying construct. There are similarities between Rasch and Mokken models, but Rasch models assume all items have the same item response function [22].

Unlike CTT methods, IRT methods do not assume equal precision, or measurement invariance across the range of test scores. Measurement invariance assumes that the probability of selecting a given item response category is comparable across groups, given similar levels of the latent construct being measured [25].

Generalisable theory (G theory). Simply put, G theory is a statistical theory used to evaluate the reliability of behavioural measurements [26]. G theory emerged to address limitations regarding undifferentiated error in CTT, which some methodologists criticise as being overly crude characterisations of potential and actual sources of measurement error [27]. In response, Generalisable theory extends earlier analysis of variance approaches for assessing reliability by estimating score variation attributed to different sources of systematic and unsystematic error, such as variation arising from each person, facet, or person-facet interaction [27].

Based on this information, I determined that CTT was the most appropriate methodological approach for this thesis research. This decision was based off of two primary considerations: 1) one objective of this thesis research was to elucidate which sub-constructs (i.e., constituent domains, factors/dimensions, facets) were salient for measuring CE, and the structure thereof; and 2) my interest in measuring CE through its constituent factor scores. A factor analytic approach was of particular interest given CFA also provides a more flexible approach than IRT in that it allows for: 1) multidimensional measurement models, 2) continuous covariates and categorical variables to be used, 3) the direct effect of the covariate to be modelled on the latent factor, and 4) the method to incorporate error theory (i.e., measurement error covariances) [28]. While there were advantages to IRT methods when they first emerged, mathematical solutions are now in place to address previous issues the factor analytic approach had with handling categorical and non-normal data (e.g., estimators such as weighted least-squares with mean and variance adjustments [WLSMV] are now available to handle such data) [16]. Given these considerations, I felt a factor analytic modelling approach was appropriate for performing psychometric analyses, construct validation, and measurement invariance evaluation.

4.4.2 Gendered approach to collective efficacy examination

Both measurement scale development and validation were of primary concern for these analyses, as was ascertaining whether there are important differences in CE measurement characteristics amongst men and women. As such, a two-pronged approach was employed to compare CE measures between men and women. In summary, single-group, gender-specific analyses were carried out separately on men's and women's input matrices to identify the mechanisms through which CE operates for each gender, while multiple-group CFA and MIMIC modelling were performed to examine differences in factor loadings and differential item functioning (DIF). These initial analytical steps would permit an examination of the mechanisms

through CE operates differ by gender (i.e., the underlying structures of collective efficacy differ between men and women).

More specifically, subsequent to generating single-group, gender-specific CE measurement models, we examined the factor solutions to determine whether there were important differences regarding the number and nature of the factors indicated by the two models. Given considerable overlap of factors and items between gender-specific models, a single, parsimonious CE scale was examined further. For this examination, multiple-group analyses that allowed both men's and women's input matrices to be analysed simultaneously were carried out to compare factor loadings between genders. Finally, additional hypothesis tests were performed to examine measurement invariance and test for DIF between men and women in the presence of other pertinent model covariates.

DIF, or measurement *non*-invariance challenges measurement invariance, and occurs when people from different groups (e.g., men, women) with similar levels of the latent construct have different probabilities of responding to an item in a certain way [29]. It was important to assess the measurement invariance of our CE scale, and determine whether there was evidence of DIF between male and female respondents. DIF needs to be identified and addressed because it influences the probabilities of selecting response options related to those items, therefore biasing item responses and potentially confounding the interpretation of the observed variables [25, 30]. In other words, items with DIF may threaten the validity of a scale to measure levels of the underlying latent construct amongst members of different sub-groups.

The gendered approach to CE examination revealed important differences between CE mechanisms and factor scores generated by men's and women's datasets. These findings have programme implications related to both intervention content and implementation approaches. These and other related considerations are examined and discussed in Research Paper 2.

4.4.3 Sample size considerations

Given factor analytic methods are not intended to be used to draw inferences, sampling theory is not necessary for determining whether a sample is of sufficient size to obtain accurate factor solution parameter estimates (e.g., pattern [factor loadings] and structure coefficients). Alternatively, the characteristics of the sample population and the data are important for determining whether the sample is of sufficient size [31]. General rules of thumb for determining sufficient sample size for factor analyses, which draw on empirical evidence and take into account the complexities of factor analyses include participant to item ratio, the ideal

of which would be in the range of 10:1 – 20:1, though ratios of as low as 5:1 reflect the minimal threshold of acceptability [32].

4.4.4 Justification for exploratory factor analysis procedures

Research Paper 2 presents information regarding the exploratory and confirmatory factor analyses that yielded three collective efficacy scales. Below, additional information is provided to further justify related methods and modelling decisions.

4.4.4.1 Extraction method

Factor extraction represents the process by which model parameters such as pattern and structure coefficients and factor inter-correlations are estimated [31]. Conventional EFA relies on analyses of a Pearson correlation matrix to estimate these parameters. The Pearson correlation matrix maintains two assumptions: 1) the data have been measured on an equal interval scale, and 2) a linear relationship exists between the variables [33].

The use of polytomous rating scales (i.e., ordered categorical response scales of more than two levels), such as the five-point, Likert-type response format used for this study, typically results in the violation of these assumptions [34]. Pearson correlations tend to underestimate the strength of the relationships between ordinal items, and result in spurious multidimensionality and biased factor loadings [35]. However, evidence suggests that employing a polychoric correlation matrix – a completely standardised variance-covariance matrix and extension of the tetrachoric correlation matrix – instead of a Pearson correlation matrix results in a solution that more accurately reproduces the measurement model used to generate the data [36]. For these reasons, and given the ordinal nature of the response scale accompanying our collective efficacy scale, we utilised a robust WLSMV estimation method based on assessments of polychoric correlation matrices [33, 37]. A sandwich estimator, applied through the use of complex EFA, was employed to adjust for non-independence of observations within 50 *kebele* clusters [38].

4.4.4.2 Rotation and rationale

We hypothesised a multifactorial model with correlated items. As such, it was necessary to identify a technique for rotating the initial factor matrix in order to obtain a simple structure. An oblique rotation was indicated due to hypothesised item correlation, and Promax was selected as the specific oblique rotation method for these analyses. Several factors influenced the selection of the Promax rotation, including the fact that the rotation raises loadings to a power of four, which produces higher correlations amongst factors and attains a simple structure, and its ability to reach a solution more quickly than the Direct Oblimin rotation [39].

4.4.4.3 Factor retention and item elimination

Mathematically-focused factor extraction methods have a tendency to under- or over-estimate the number of factors in a solution [11, 39, 40]. The results of scree tests are often ambiguous (e.g., no clear shift in the slope) and subject to interpretation [11]. As a result, I used a combination of mathematical (i.e., eigenvalue-based Kaiser-Guttman rule), heuristic (i.e., scree plot), statistical (i.e., model fit statistics), and substantive justification to guide factor retention. That said, I was not able to perform more rigorous procedures (e.g., parallel analysis - [41, 42]) to confirm that I extracted the correct number of factors, as these analyses are not available for categorical data in Mplus [16].

4.4.5 Factor score calculation methods and future application

It was important to calculate quantitative measures of CE, in the form of factor scores, for this thesis research, as doing so would allow for future application of CE measures. However, these measures are not just useful for research, as they may be used by implementers for programme targeting. For instance, programme implementers often perform some type of rapid assessment or vetting of communities to determine whether they are appropriate for programme engagement. The CE scales presented in Research Paper 2 can be adapted and utilised to help improve targeting of communities for programme intervention. Communities scoring relatively low on social disorder and high on social response, common vision, associational participation, and community organisation may be more open to and ready for change, and diffusion of innovations amongst its members [43]. From an evaluation perspective, the CE scales and resulting factor scores may also serve as diagnostic tools to examine community-level factors associated with intervention effectiveness or lack thereof.

Given the potential for broad application of the CE scales, it was important to demonstrate to a larger programme audience, that may be less suited to perform sophisticated modelling and weighting of factor scores that only simple summations or averages of factor scores may render the CE factor scores useful. Rather simple, unsophisticated (i.e., coarse) methods were used to calculate factor scores to make the process more accessible. Given factor loadings indicated a range of magnitudes, thus indicating the relative proximities of the relationships of each item to its respective factor, weighted factor scores were also presented in order to provide slightly more sophisticated data analyses for the purposes of peer-reviewed publication. In addition to Figure 1 presented in Research Paper 2, Table 4.1 included at the end of this chapter present both unweighted and weighted factor scores.

4.5 Collective efficacy scale development and validation

4.5.1 Preamble for Research Paper 2

Research Paper 2 presents details regarding the five-step scale development and validation process utilised to construct all CE measurement models. The paper highlights this process, using the data collected in Ethiopia via a baseline survey for the *Andilaye* trial. The aim of this paper was to explore the underlying structure of collective efficacy in rural Ethiopia. A gendered approach was employed, in which separate CE scales were developed for men and women in order to examine the specific mechanisms through which CE operates for men and women. Given there was considerable overlap between these gender-specific scales, a parsimonious CE scale that included factors reflected in both gender-specific scales was then identified and tested. Item functioning between men and women as well as respondents with leadership roles in the community and those without was then assessed on the parsimonious model to determine whether different groups of individuals with similar levels of the latent construct have different probabilities of responding to an item in a certain way [29]. This paper concludes with a discussion of findings regarding the factors emerging in the CE scales, and focuses on implications for the design, targeting, and evaluation of community-based interventions.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Maryann Delea
Principal Supervisor	Dr. Thomas Clasen
Thesis Title	Social constructs, behaviour change, and the uptake of community-based WASH interventions: Metrics and analytical approaches for measuring collective efficacy and other social constructs

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?	International Journal of Environmental Research and Public Health		
When was the work published?	28 September 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of and designed this study. I worked with colleagues to develop the survey instrument and execute cognitive interviews. I performed all data analyses, and developed the first draft of the paper. Co-authors provided important contributions.

Student Signature: [redacted] Date: 26/09/2018

Supervisor Signature: [redacted] Date: 10-18-2018

International Journal of
Environmental Research and Public Health

MDPI

Article

Collective Efficacy: Development and Validation of a Measurement Scale for Use in Public Health and Development Programmes

Maryann G. Delea [1,2,*], Gloria D. Sclar [2], Mulat Woreta [3], Regine Haardörfer [4], Corey L. Nagel [5], Bethany A. Caruso [2], Robert Dreibelbis [1], Abebe G. Gobezayehu [3], Thomas F. Clasen [1,2] and Matthew C. Freeman [2]

1 Department of Disease Control, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK; robert.dreibelbis@lshtm.ac.uk (R.D.); thomas.f.clasen@emory.edu (T.F.C.)
2 Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA; gloria.sclar@emory.edu (G.D.S.); bethany.caruso@emory.edu (B.A.C.); matthew.freeman@emory.edu (M.C.F.)
3 Emory Ethiopia, Bahir Dar, Addis Ababa, Ethiopia; mworeta@manhep.org (M.W.); abebe.g.gobezayehu@emory.edu (A.G.G.)
4 Department of Behavioral Sciences and Health Education, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA; regine.haardoerfer@emory.edu
5 College of Nursing, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA; cnagel@uams.edu
* Correspondence: maryann.delea@lshtm.ac.uk; Tel.: +44-20-7636-8636

Received: 7 September 2018; Accepted: 26 September 2018; Published: 28 September 2018

Abstract: Impact evaluations of water, sanitation, and hygiene interventions have demonstrated lower than expected health gains, in some cases due to low uptake and sustained adoption of interventions at a community level. These findings represent common challenges for public health and development programmes relying on collective action. One possible explanation may be low collective efficacy (CE)—perceptions regarding a group's ability to execute actions related to a common goal. The purpose of this study was to develop and validate a metric to assess factors related to CE. We conducted this research within a cluster-randomised sanitation and hygiene trial in Amhara, Ethiopia. Exploratory and confirmatory factor analyses were carried out to examine underlying structures of CE for men and women in rural Ethiopia. We produced three CE scales: one each for men and women that allow for examinations of gender-specific mechanisms through which CE operates, and one 26-item CE scale that can be used across genders. All scales demonstrated high construct validity. CE factor scores were significantly higher for men than women, even among household-level male-female dyads. These CE scales will allow implementers to better design and target community-level interventions, and examine the role of CE in the effectiveness of community-based programming.

Keywords: collective efficacy; WASH; behaviour change; gender; behavioural control; collective action; cooperative behaviour; community-based interventions; factor analysis

1. Introduction

It has become commonplace in international development to intervene in communities with interventions that require collective action without first gauging the communities' perceptions regarding their ability and autonomy to engender and maintain change. Such is the case with many

water, sanitation, and hygiene (WASH) interventions, some of which require collective action before first assessing whether reliance on shared agency is a realistic expectation, others of which neglect to address important factors of collective behaviour. This, perhaps, may be an artefact of common programme approaches that tend to address independent, individual and household-level behaviours while aiming for change at higher levels, such as villages, communities, or other collectives of people. Yet, in order to facilitate interdependent adoption of improved collective behaviours, evidence suggests it is important for interventions to address underlying factors that facilitate action and change at those levels [1–4]. Overlooking or underestimating the role of collective behavioural factors, such as behavioural control perceptions (e.g., agency-related factors such as self- and collective efficacy) and social schemas (e.g., social norms) in the uptake of community-based interventions may, in turn, attenuate intervention impact [2,4].

Evidence suggests that collective action is required for WASH interventions to reach the coverage and use levels likely required to realise health gains through "herd protection" [5–7]. Results from rigorously designed and evaluated WASH studies demonstrate lower than expected impact of WASH interventions on health [8–10], in some cases due to poor intervention uptake and sustained adoption. When interpreting these findings, it is important to consider the implementation approaches and intervention techniques that were employed, as well as the level at which these interventions were targeted (e.g., individual, household, group, community). It is pertinent to question whether the WASH sector is considering potentially important behavioural antecedents (i.e., upstream behavioural factors predictive of downstream behavioural, health, and development impacts) in their theories of change, intervention designs, and programme evaluations.

In many sectors, including but not limited to WASH, community-based programmes that target higher order groups (e.g., households, villages, health centres, government ministries) often inadequately address factors of collective behaviour in their intervention design and implementation strategies [11]. Collective efficacy (CE) is one such factor. CE is a latent social construct that encompasses a combination of cognitive and socio-structural factors which facilitate peoples' shared beliefs in their collective power, or ability to come together to execute actions related to a common goal [12,13]. As with other social constructs, CE is complex, and draws on multiple sub-construct domains, such as social cohesion, social control, and cognitive and structural social capital. In addition to shaping a group's decision to pursue common goals, collective efficacy also influences the amount of effort the group spends working toward those goals and the level of persistence expended when group efforts fail to yield desired results [12].

The notion that shifting away from an existing, undesirable behaviour is predicated on a critical mass of group members believing that enough members will cooperate in enacting the new behaviour is well established [1,14,15]. This implies group goal selection. Perceptions regarding CE influence group goal selection and performance [16–19]. In fact, evidence suggests that perceptions regarding efficacy are better predictors of behaviour than prior performance or goal attainment [12].

CE perceptions are influenced by personal attributes, collective dynamics, and situational contexts [12,20]. As such, the mechanisms through which CE operates may differ for men and women. In many societies, existing psychosocial and structural inequalities translate to important differences in the social opportunities men and women have to engage with certain formal and informal community structures [21–24]. These differences are well cited, and have been known to create disparities between men and women in terms of their social inclusion, mobility, civic engagement and associational participation, and position within the communities in which they live [24–26]. These disparities may translate to differences in the mechanisms through which CE operates among men and women, such as the size and strength of their respective social networks, their sense of belonging or social attachment, and perceptions regarding the organisation and responsiveness of the community. In addition, various "gendered structures of constraint" may differentially influence perceived individual- and community-level psychological capabilities among men and women [22,26]. For example, normative expectations and social restrictions that preclude certain types of individuals

(e.g., women) from moving and socialising freely outside of the home and within certain community groups may contribute to disparities among men and women in terms of their perceptions regarding self- and collective efficacy [27–29]. Consequently, existing evidence supports the idea that differences in collective action exist between genders [24].

In public health, and the WASH sector more specifically, we are still seeking to elucidate how complex, group-level factors effect behavioural outcomes and health impacts at individual and collective levels. While social constructs have been explored more broadly when examining various aspects of shared and communal WASH resources [30–32], investigations into collective efficacy in the context of community-based interventions remain scarce. Agentic factors such as self- and collective efficacy are featured extensively in behaviour and behaviour change theory [12,13,33,34]. Yet, only a limited number of studies have used empirical evidence to critically assess collective efficacy, and investigate its underlying structure (i.e., the relationships between the sub-constructs, such as factors and facets that are important to the measure of the construct) [19,35,36].

Researchers studying violent crime in urban Chicago, Illinois, USA found that two sub-constructs—social cohesion and trust among neighbours, and informal social control (i.e., neighbours' willingness to intervene on behaviour for the common good)—define CE as a latent construct [19]. A study conducted in Blacksburg, Virginia, USA suggested that CE related to community computing was comprised of four sub-constructs—activism, informedness, belonging, and association [35]. Among teachers in elementary schools within one urban school district in the USA, CE was found to have a single factor structure that contained items tapping to group competence and task analysis [36]. Findings from these studies are conflicting, and all known psychometric examinations of CE have been conducted on data collected from literate populations in high-income contexts. No known studies have developed and applied CE measurement scales in such a manner that allows for a more thorough examination of the role CE plays in behaviour change and the overall effectiveness of community-based interventions operating in developing contexts.

To address this gap, our team designed a series of studies to develop, validate, and employ scales to assess CE and compare perceptions between genders. For this particular study, we aimed to develop a metric that could be used in the context of a WASH trial, to assess pre-intervention behavioural control perceptions at baseline, and evaluate the effect of CE factors on the uptake of a demand-side sanitation and hygiene intervention at endline. We hypothesise that CE is an important antecedent of the cooperative behaviour and collective action needed to bring about sustained adoption of improved WASH practices at the level required to realise health impacts. Therefore, we used a systematic and widely accepted scale development and validation process [37] to develop and test CE measurement scales, the methods and results of which we present herein. We set out to investigate several research questions related to the measurement of collective efficacy, its constituent sub-constructs, and its factor structure. First, we were interested in elucidating which sub-constructs (e.g., constituent domains, factors/dimensions, facets) were salient for measuring collective efficacy, and the structure thereof, in the rural Ethiopian context. Second, we wanted to examine whether the psychometric characteristics of the resulting CE measurement models were compelling in terms of their ability to demonstrate construct validity, or the degree to which the scale measures what it purports to measure. Finally, we aimed to determine whether there were important differences in the measurement of CE between Ethiopian men and women. The development of these CE scales will allow for enhanced intervention design and targeting, and further examination of potential associations between CE, as measured by our scale, and intervention effectiveness, as measured through changes in WASH-related behavioural outcomes and health impacts.

We set forth three related hypotheses to test as we explored our research questions. First, we hypothesised that the factor solutions derived from our empirical data would support our theorised construct dimensionality. Second, we hypothesised that the CE measurement models produced via factor analytic methods would demonstrate high construct validity. Third, we hypothesised that given their status, mobilisation, and inclusion within their communities,

men would have higher perceptions of behavioural control and related factors than women; leaders would have higher perceptions than non-leaders.

2. Study Overview

This study took place in Amhara National Regional State, Ethiopia as a sub-study of a larger cluster-randomised controlled trial entitled *The impact of enhanced, demand-side sanitation and hygiene promotion on sustained behavior change and health in Amhara, Ethiopia*, or *Andilaye* for short (registered on clinicaltrials.gov, NCT03075436).

2.1. Ethical Approval

The *Andilaye* trial and its sub-studies received ethical approval from Emory University's Institutional Review Board (IRB00076141), the Amhara Regional Health Bureau Research Ethics Review Committee (HRTT0135909), and London School of Hygiene & Tropical Medicine's Observational/Interventions Research Ethics Committee (Ref. 9595). Fieldworkers provided study participants with full details regarding the study prior to inquiring about consent to participate, and took steps to ensure confidentiality for all study participants.

2.2. Data Collection and Analysis

We used a five-step, sequential exploratory and confirmatory approach [38] to develop, refine, and validate a theoretically grounded and evidence-based CE measurement scale [37]. This process relied heavily on a factor analytic approach. Factor analysis is a psychometric method that allows for the measurement of latent constructs that cannot be directly observed or measured. In a factor analytic framework, latent constructs are measured through the analysis of manifest variables (e.g., survey items), indicators that represent certain aspects of the latent construct [39]. The analyses of these data can elucidate the underlying structure of the construct and its constituent sub-constructs.

Factor analysis comprises a suite of analytical methods, including exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA is a descriptive analytical method used to determine the number of common factors in a measurement model, and identify which measured variables are indicators of the latent construct (i.e., identify factor structure) [40]. CFA is an evaluation approach that allows for direct testing and validation of hypothesised factor structures to assess their appropriateness as measurement models (i.e., determine construct validity, falsify hypothesised models). These types of analytical methods use matrix algebra to generate statistics used to reveal a construct's underlying factor structure [41].

Below, we provide details related to each step of our scale development and validation process. Given the progressive nature of this process, we present methods and results for each step, in chronological order, followed by a discussion and conclusions. The methods and results presented below demonstrate the process by which we produced three refined CE measurement scales that demonstrated good construct validity.

3. Collective Efficacy Scale Development and Validation—Methods and Results

3.1. Step 1. Defining Collective Efficacy via a Hypothesised Framework

3.1.1. Step 1. Methods

We leveraged theory and evidence to operationally define CE, and establish a hypothesised CE framework [37,42]. This began with a desk review of the literature during which we identified and extracted information regarding the CE construct and constituent sub-constructs. We then performed an applied thematic content analysis to re-organise emergent sub-constructs into key domains, dimensions/factors, and facets to generate the framework. We constructed definitions for each CE dimension/factor. In factor analytic terms, this step involved establishing operational

definitions for the latent variables [43], and providing substantive justification for the hypothesised framework [44].

3.1.2. Step 1. Results

We established the following operational definition for collective efficacy: A latent construct comprised of a combination of the cognitive and socio-structural components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal. This conceptualisation of CE is grounded in evidence and theory, including Social Cognitive Theory and Social Learning Theory [19,34,45]. The resulting hypothesised CE framework represented a seven-factor conceptualisation, with items tapping to social disorder, social response, social capital, social equity, common values, community attachment, and agency. We hypothesised that these seven factors represented aspects of three domains: informal social control, social cohesion, and behavioural control (Table 1).

3.2. Step 2. Designing the Collective Efficacy Survey

3.2.1. Methods

Next, we generated an item pool by extracting relevant prompts from existing surveys instruments [37]. We prioritised pre-existing, validated tools, when possible. We then coded survey items against our set of CE sub-constructs, removed repetitive items, and designed new items when no relevant prompts existed for a given sub-construct.

We re-structured all survey items such that they worked well with a five-point, Likert type response format [37]. Once our draft tool was developed, it was translated into Amharic, and back-translated into English to ensure the quality of the translations. During December 2016-February 2017, the CE survey was then piloted and iteratively refined through a series of formative research activities, which included two rounds of cognitive interviews and one round of focus group discussions. These formative research activities were conducted in Amharic in the Bahir Dar Zuria district (*woreda*) of Amhara National Regional State, Ethiopia. Interviews and group discussions were audio-recorded and transcribed directly into English.

Trained fieldworkers conducted four in-depth interviews (three women, one man) that employed think-aloud and verbal probing techniques [46] during an initial round of cognitive interviews. This cognitive validation process allowed us to ensure consistency between what we intended to convey through the survey items and participants' understanding regarding the meaning of the survey items (i.e., assess face validity). This process helped us confirm that each survey item tapped to the desired sub-construct within our CE framework. Furthermore, our inclusion of cognitive interviews as a formative scale development activity allowed us to investigate whether there was early evidence regarding the substantive aspects of construct validity [47].

We then conducted four focus group discussions (two with women, two with men) to explore concepts and real-world examples related to our hypothesised CE framework. Based on our qualitative findings, we revised survey items to be more contextually relevant, and designed additional items to ensure each CE sub-construct was adequately measured, and our CE survey had reached conceptual saturation. Finally, we conducted a second round of cognitive interviews (four women, three men) to pilot this revised CE survey and again check participants' understanding of the survey items.

Table 1. Hypothesised collective efficacy framework.

Domain	Dimension/Factor	Definition	Related Facets
Informal social control	Social disorder	General conflict and threats to the existing order—e.g., incivility	Incivility, intolerance, people not living in harmony
	Social response	Community members actively address social issues—e.g., respect differences, celebrate successes, and react to social inequity (below)	Willingness to intervene, community support in times of crisis, collective morals, tolerance, inter-group cooperation
Social cohesion	Common values	Community members share common values, beliefs and ideologies	Order, group cohesion and inclusion, social integration, acceptance, collective norms and ideals, common civic culture
	Social capital *	Residents have strong social networks within the community that establish a sense of trust among community members and leaders and allow for acts of reciprocity	Social networks and social capital, supporting networks and reciprocity, social organisation and groups, associational activity and common purpose, social trust, social bonds, social safety nets, trust and solidarity, volunteer activities
	Social equity	Residents have equal access to resources, services and opportunities within the community and there are safety nets in place in times of crisis	Social solidarity and reductions in wealth disparities, information and communication, contribution to household resources, social justice and equity, ownership of household assets/resources
	Community attachment	Residents feel a sense of connection to their community whether it is through ownership of resources/assets, through social ties or both. Being a part of the community is an aspect of a resident's identity	Place attachment, place identity, sense of belonging
Behavioural control	Agency	Community members' belief that they themselves are capable of achieving an identified goal (i.e., self-efficacy) Community members' belief that their community as a whole is capable of making positive changes (i.e., collective action), and that they ought to be doing so. Here, we also explore community members' perceptions regarding the need for exogenous intervention to achieve common goals	Perceived performance experiences (i.e., enactive mastery), vicarious experiences, physiological arousal (i.e., emotional state/control), self-esteem Collective action and cooperation, participation in collective action, collective behavioural control, empowerment

Note: * Including both structural components (e.g., civic structures engagement/participation) and cognitive components (e.g., trust, reciprocity).

3.2.2. Step 2. Results

The formative survey and scale development work conducted during Step 2 resulted in a 50-item CE instrument. The 50 items were comprised of group-referent statements about interpersonal and ecological aspects of the respondent's community that related to CE as well as self-referent statements about the respondent's own sense of self, agency, autonomy, and level of engagement within his/her community. The content of the measured variables (i.e., items) used to operationalise CE is provided in Table A1. There, we provide detailed information on all 50 CE survey items.

3.3. Step 3. Administering the Collective Efficacy Survey

3.3.1. Step 3. Methods

We trained fieldworkers to administer the CE survey by reading each survey item, followed by each of the five response options (i.e., completely disagree, partially disagree, neither agree nor disagree, partially agree, completely agree). Through two waves of data collection conducted during March–June 2017, fieldworkers visited *Andilaye* study households with the goal of administering the same 50-item CE survey to one man and one woman in randomly selected households. During the initial wave of data collection, the primary female caregiver of the trial's index child (i.e., youngest child aged 1–9 years) was targeted for *Andilaye's* baseline survey. Half of these respondents were randomised to receive our CE survey module. During the second wave, we targeted *Andilaye* study male heads of household. All men were targeted for the CE survey regardless of which survey the respondent from the initial wave of data collection was randomised to receive. This subsequent data collection round was designed to help us examine similarities and differences of CE perceptions between men and women in general, and household-level male-female dyads within a sub-set of study households in particular. The sub-set of dyadic households represented those in which: (1) a woman responded and was randomly allocated to the CE survey during the first wave of data collection; and (2) a man was willing and able to participate in the CE survey during the second wave of data collection.

3.3.2. Step 3. Results

Fieldworkers targeted 1849 CE surveys. Consent was provided by 1846 respondents (i.e., 99% response rate). Fifteen observations were dropped due to data entry errors. The final analytical dataset contained 1831 observations; 1105 from men, 726 from women. At least one individual from 1311 households in 50 sub-district (*kebele*) clusters responded to our CE survey. We obtained CE data on household-level male-female dyads from 520 'dyadic' households. Data were also obtained from 585 men and 206 women residing in 791 'non-dyadic' households. See Table 2 for data on respondent demographics and household characteristics.

Table 2. Respondent demographics, household- and cluster-level characteristics, by gender.

Characteristics	Aggregate		Men		Women	
Number of respondents	1831		1105		726	
Respondent demographics			*n* %			
Median age (IQR)	35	(29–45)	40	(31–47)	31	(27–38)
Relation to head of household						
Respondent is the head of household	1170	64%	1002	91%	168	23%
Spouse	540	29%	4	<1%	536	74%
Other relative	114	6%	92	8%	22	3%
Other non-relative	7	<1%	7	<1%	0	0%
Married	1667	91%	1014	92%	653	90%
Household-level characteristics			*n* %			
Median number of members per household (IQR)	5	(4–6)	5	(4–6)	5	(4–6)
Religion						
Orthodox Christian	1730	95%	1047	95%	683	94%
Muslim	55	3%	33	3%	22	3%
Other	41	2%	20	2%	21	3%
Head of household's education (highest level attained)						
No formal education	1317	72%	801	72%	516	71%
At least some first cycle primary (grades 1–4)	175	10%	104	9%	71	10%
At least some secondary (grades 5–8)	233	13%	139	13%	94	13%
Any high school or above	91	5%	53	5%	38	5%
Refuse or do not know	15	<1%	8	<1%	7	<1%
Access to household latrine (any type)	1394	76%	854	77%	540	74%
Primary drinking water source location						
In compound	68	4%	40	4%	28	4%
Outside compound	1757	96%	1060	96%	697	96%
Household member with leadership in a community structure *	215	16%				
Total number of household-level male-female dyads *	520	40%				

Notes: IQR = inter-quartile range. Five observations from the men's sub-sample (and therefore the aggregate as well) were missing data on the number of members in their households, the head of household's highest educational attainment, religion, and ethnicity; one observation from the men's sub-sample was missing data on marital status. Two observations from the women's sub-sample were missing data on the number of members per household. * 1311 households.

3.4. Step 4. Performing Psychometric Analyses

3.4.1. Step 4. Methods

Data preparation and screening

Initial data cleaning and descriptive analyses were performed in Stata (version 15.0 StataCorp, College Station, TX, USA). We performed subsequent descriptive and all factor analyses (EFA, preliminary CFA, single-group CFA of EFA-derived factor solutions, multiple-group CFA and Multiple Indicators Multiple Causes [MIMIC] modelling) in Mplus software (version 8 Muthén & Muthén, Los Angeles, CA, USA).

To prepare our data for analyses, we first partitioned our CE dataset by gender, and then employed a random-number seed to identify two separate random split-halves for both men and women sub-samples. We designated one random split-half sample for scale development via EFA for each gender; the remaining random hold-out sample was reserved for scale validation via CFA of the EFA-derived factor solutions. This division of the dataset resulted in four split-half samples, two for women, and two for men (i.e., n_{W1}, n_{W2}; n_{M1}, n_{M2}). Univariate analyses performed in Stata and verified results in Mplus examined respondent/household characteristics and item distributions (frequencies and proportions—Table S1) for all 50 CE items on an aggregate level and between genders. We performed Mann-Whitney Rank Sum tests to determine whether there were any significant differences in respondent and household-level characteristics between split-halves.

Preliminary confirmatory factor analysis of the hypothesised CE framework

We decided *a priori* to first test our hypothesised CE framework (Table 1) via a preliminary CFA [48]. Poor model fit statistics for this preliminary CFA would signal that the hypothesised CE framework may need modification in order to produce an appropriate CE measurement framework. In the event model fit statistics indicated poor fitness, we decided *a priori* that we would perform EFA to determine alternative CE factor structures derived from our own empirical data, and conduct CFA again to test and validate the resulting EFA-derived factor structures [40].

For preliminary CFA, we used a robust weighted least-squares with mean and variance adjustment (WLSMV) estimation method [49] based on assessments of polychoric correlation matrices [50,51]. A sandwich estimator was applied to adjust for non-independence of observations within 50 *kebele* clusters. Because it would have been justifiable to conclude our analyses with CFA if the complex preliminary CFA indicated good model fit, we performed these analyses on the full men and women sub-samples (n_M and n_W, respectively). We examined goodness-of-fit indices, assessing both absolute fit (e.g., χ^2:df ratio, root mean square error of approximation [RMSEA]) and incremental, or relative fit (e.g., comparative fit index [CFI], Tucker-Lewis index [TLI]). Standard thresholds of acceptable and good model fitness were employed (i.e., χ^2:df ratio < 3.0; RSMEA of $\leq$0.10 acceptable fit, $\leq$0.05–0.06 good fit; CFI & TLI $\geq$ 0.90 acceptable fit, $\geq$0.95 good fit) [40,52,53]. Factor loadings less than 0.32 were considered non-salient (i.e., not statistically meaningful) [54]. *Post hoc* refinements included the deletion of items with non-salient (factor loadings < 0.32) and/or non-significant (two-tailed $p > 0.05$) factor loadings. We also dropped all factors with less than three items with salient and significant factor loadings, as these factors may have insufficient component saturation, meaning the factor may not have been fully conceptually explained by the emergent items, which could compromise factor interpretation [48].

Exploratory factor analysis

We performed complex EFA on one split-half of data from both men and women sub-samples (n_{W1} and n_{M1}, respectively). As with preliminary CFA, we used a robust WLSMV estimation method based on assessments of polychoric correlation matrices for EFA, and applied a sandwich estimator to adjust for non-independence [49–51]. An oblique rotation was indicated due to hypothesised item correlation, and Promax was selected *a priori* as the specific oblique rotation method for these analyses.

Decision rules related to factor retention were based on a combination of: (1) mathematically based and heuristic descriptive guides (i.e., Kaiser-Guttman rule (eigenvalue > 1.0) [55], scree-plot); (2) goodness-of-fit; and (3) other substantive justification, such as results from cognitive interviews, and theoretical and empirical evidence [40,44,48]. As with preliminary CFA, we employed a holistic approach to evaluate goodness-of-fit indices for EFA. The same thresholds used for preliminary CFA were used for EFA, but we also included an assessment of root mean square residual (RMSR); values below 0.08 indicate reasonable model fit [52].

Oblique rotations produce pattern coefficients that do not fully characterise the relationship between an item and a given factor [44]. Therefore, in order to appropriately interpret EFA results, we evaluated both the factor pattern and factor structure matrices [48]. Structure and pattern coefficients with an absolute value greater than 0.32 were considered salient. Items with factor loadings less than this threshold poorly measured the latent factors, and were eliminated in a step-wise manner [54]. We iteratively re-analysed measurement models subsequent to item reduction [44]. To be retained, factors needed to demonstrate adequate component saturation and sufficient evidence that they were at least adequately measured (i.e., at least three items with factor loadings greater than 0.32, and no or limited item cross-loadings) [56]. Only complex variables (i.e., those with salient factor loadings on more than one factor [cross-loadings]) with strong substantive justification for their cross-loadings were retained. Models that represented the most readily interpretable (i.e., the simplest solution, per Thurstone criteria [57]—outlined in the Supplemental Material) and theoretically justifiable solutions were selected for the refined, gender-specific factor solutions [44].

With regard to the interpretation of EFA results, factor loadings indicate the pattern of item-factor relationships, and are often referred to as pattern coefficients [40]. Factor loadings represent completely standardised estimates of regression slopes for predicting the indicators from the latent variable [40]. While some methodologists caution against the use of thresholds, common guidelines for the interpretation of factor loadings can be used to facilitate interpretation of results (e.g., factor loadings >0.71 excellent, >0.63 very good, >0.55 good, >0.45 fair, and >0.32 adequate) [54].

Single-group confirmatory factor analysis of EFA-derived, gender-specific factor solutions

During CFA of EFA-derived factor solutions, we used split-half hold-out samples (n_{W2}, n_{M2}) to validate EFA-indicated, gender-specific measurement models. The underlying structure used to operationalise the latent factors were those indicated in the factor solution produced via EFA. We identified the scale of every latent factor through the use of marker indicator items, which we identified as the item that demonstrated the highest factor loading on its respective factor, per EFA results [40]. As with our preliminary CFA and EFA, we performed these CFAs using WLSMV with a sandwich estimator to adjust for non-independence. Through *post hoc* model refinements, we eliminated items with non-salient (i.e., factor loading < 0.32) and/or non-significant factor loadings.

We used the same process for holistically examining goodness-of-fit and carrying out *post hoc* model refinements for the CFA of EFA-derived models as those employed during the preliminary CFA. After examining fit statistics, we assessed residuals and modification indices for indications of localised areas of strain (i.e., misfit) in the measurement models. Modification indices greater than 3.84 indicated opportunities for further model refinement and fit improvement, through the estimation of additional parameters, if justified [40,58].

3.4.2. Step 4. Results

Preliminary confirmatory factor analysis of hypothesised CE framework

Results from the preliminary CFA of our hypothesised CE framework indicated that the men's refined model (i.e., with *post hoc* adjustments) demonstrated moderately acceptable absolute fit, but poor incremental or relative fit (χ^2:df ratio = 2.606, RSMEA = 0.038 [0.036–0.040], CFI = 0.911, TLI = 0.904). This suggested that while our hypothesised CE framework represented a plausible structure of the mechanisms through which the CE process operates among men in the Ethiopian context, an alternative framework may have provided a better measure of CE. The women's model did not fit the data well (χ^2:df ratio = 3.409, RSMEA = 0.058 [0.055–0.060], CFI = 0.895, TLI = 0.888), which indicated that the data failed to validate the hypothesised CE framework for women respondents. This suggested that the CE framework required modification in order to reveal the mechanisms through which CE operates for women in rural Ethiopia. See Tables A2 and A3 for additional preliminary CFA results. These findings provided rationale for performing EFA.

Scale development and validation samples, balance of respondent and household characteristics

Our split-half EFA samples consisted of 366 observations from women (i.e., participant to item ratio of over 7:1), and 555 observations from men (i.e., participant to item ration of 11:1). While the participant to item ratio was lower for women, split-half sample sizes were sufficient for both genders according to standard guidance [48,54]. All respondent demographics and household characteristics were balanced across aggregate and gender-specific sub-samples (results not displayed).

Factor extraction and item reduction

Seven factors were extracted during final EFA for both gender-specific EFA-derived factor solutions. We present factors and items indicated in both men's and women's EFA-derived factor solutions in Tables 3 and 4, respectively. Information related to the item reduction processes for both models, and a detailed summary of each factor emerging from the gender-specific EFA-derived

models are provided in the Supplemental Material. Below, we summarise the resulting gender-specific EFA-derived measurement models, by domain.

Men's collective efficacy measurement model

EFA results revealed a seven-factor men's CE measurement model with good model fit (χ^2:df = 1.209, RMSEA = 0.019, RMSR = 0.037). The seven factors included social response, social networks and personal agency, social attachment, common vision, community leadership, associational participation, and community organization. Social response corresponded to the informal social control domain, though it also tapped to certain aspects of cognitive social capital (e.g., trust in community members, reciprocity of knowledge) that may influence social response. The social networks and personal agency factor corresponded to the cognitive social capital domain, though it also tapped to structural social capital, as it reflects the strength and responsiveness of one's social structures. Social attachment and common vision factors corresponded to the social cohesion domain. Community leadership and associational participation factors pertained to the structural social capital domain. These factors and the concepts reflected in their constituent items align with our hypothesised operational definitions of informal social control, social cohesion, and behavioural control (Tables 1 and A4). See Table 3 for information regarding the specific items that tapped to the CE factors in the men's CE measurement model. Modification indices above 3.84 were all relatively low, meaning localised strain was relatively low in all areas identified. No further modifications were deemed theoretically or mathematically justifiable. The standardised estimates of factor loadings from this model were acceptable (Table 3).

After dropping one item (ADVICE) from the male EFA-derived factor solution as a result of less than minimal variance, we conducted CFA on the remaining items tapping to seven factors. *Post hoc* model refinements yielded a refined 31-item, seven-factor solution. The refined CFA model validated the EFA-derived model with minor modifications, and demonstrated good absolute and incremental model fit (χ^2:df ratio = 1.498, RSMEA = 0.030 [0.025–0.035]), CFI = 0.971, TLI = 0.968). The vast majority of items (87%, 27 of 31) comprising the refined solution demonstrated very good to excellent factor loadings (i.e., loadings > 0.630) on a single factor (Table 3).

Women's collective efficacy measurement model

EFA results revealed a seven-factor women's CE measurement model with good model fit (χ^2:df = 1.281, RMSEA = 0.028, RMSR = 0.041). The seven factors included social networks and reciprocity, social disorder, social attachment and personal agency, social response, common vision, associational participation, and community organisation and leadership. The social networks and reciprocity factor corresponded to the cognitive social capital domain, though it also tapped to certain aspects of structural social capital, as it reflected perceptions related to collectives of individuals that promote and protect mutual or personal interests. The informal social control domain included factors related to social disorder and social response. Two factors, social attachment and personal agency and common vision, comprised the social cohesion domain. Associational participation and community organisation and leadership corresponded to the structural social capital domain. These factors and the concepts reflected in their constituent items align with our hypothesised operational definitions of informal social control, social cohesion, and behaviour control. See Table 4 for information regarding the specific items that tapped to the CE factors in the women's measurement model.

We conducted CFA on the items tapping to seven factors, as indicated by EFA. *Post hoc* model refinements yielded a refined 33-item, six-factor solution. The refined CFA model validated the majority of the EFA-derived factor structure, and demonstrated adequate model fit (χ^2:df ratio = 1.574, RSMEA = 0.040 [0.034–0.045]), CFI = 0.962, TLI = 0.958). A majority of items (79%, 26 of 33) on the refined CFA solution demonstrated very good to excellent factor loadings (i.e., >0.630), and 12% (4 of 33) demonstrated good factor loadings (i.e., between 0.550–0.629) (Table 4).

Table 3. Factor loadings for random split-half samples for EFA and CFA of EFA-derived factor solutions, men.

Factors and Associated Items	Item	Final EFA—Factor Pattern Coefficients (n_{M1} = 555)	Final EFA—Factor Structure Coefficients (n_{M1} = 555)	Baseline CFA (n_{M2} = 550)	Refined † CFA (n_{M2} = 550)
Factor 1: Social response (average factor loading = 0.565; average structure coefficient = 0.605; average factor loading on refined CFA = 0.634)					
People in this community live in harmony with each other most of the time.	HARMONY	0.694	0.489	0.440 *	0.438 *
When there is a problem in this community, people come together to discuss how it should be solved.	COMPRSLV	0.654	0.702	0.692 *	0.690 *
People in this community can be trusted.	COMTRUST	0.634	0.741	0.683 *	0.682 *
If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	HLPCRPDZ	0.620	0.657	0.698 *	0.695 *
This is a close−knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.592	0.720	0.791 *	0.789 *
Most people in this community have similar beliefs about what is right and what is wrong.	SIMBLIEF	0.567	0.484	0.409 *	0.408 *
If there is a big dispute between two persons, other people from the community will help in solving the problem.	SLVDISPU	0.510	0.596	0.652 *	0.649 *
People in the community share new knowledge with their neighbour if they learn something new.	SHAREKNO	0.414	0.649	0.726 *	0.724 *
Differences between people, such as the amount of land they own, often causes problems in this community.	DIFPROBS	−0.398	−0.403	−0.278 *	-
Factor 2: Social networks and personal agency (average factor loading = 0.588; average structure coefficient = 0.678; average factor loading on refined CFA = 0.668)					
I have the capacity to achieve my future aims.	SELFEFF	−0.985	−0.841	0.510 *	0.512 *
I have the ability to contribute to this community's development.	SEDEV	−0.600	−0.729	0.675 *	0.674 *
If you suddenly need some money, you can borrow from a person or group in your community.	BORMONEY	−0.479	−0.595	0.759 *	0.758 *
If you and your relatives suddenly had to go away for a day/two, you could count on your neighbours to take care of your children.	NEICAREG	−0.477	−0.623	0.657 *	0.657 *
My neighbours sometimes come to me to share their problems and get help.	COME4HLP	−0.398	−0.602	0.737 *	0.738 *
Factor 3: Social attachment (average factor loading = 0.577; average structure coefficient = 0.723; average factor loading on refined CFA = 0.864)					
I feel attached to this community and its people.	ATTACH	0.795	0.836	0.860 *	0.857 *
People in this community accept me as a member of the community.	ACCEPT	0.673	0.792	0.863 *	0.864 *
Being a member of this community is part of who I am.	IDENTITY	0.631	0.779	0.871 *	0.871 *
People in this community should work together to develop the community.	SHOULDEV	0.455	0.637	0.075	-
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.332	0.571	0.071	-
Factor 4: Common vision (average factor loading = 0.519; average structure coefficient = 0.657; average factor loading on refined CFA = 0.733)					
Most people in this community have similar hopes about the future development of the community.	SIMHOPES	0.702	0.821	0.832 *	0.830 *
People in this community share the same ideas on how village matters should be managed.	COMMGMT	0.651	0.789	0.890 *	0.887 *
Most people in this community have common values, for example, they value hard work.	COMMVALU	0.586	0.697	0.813 *	0.809 *
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.496	0.652	0.702 *	0.760 *
During a crisis situation, such as drought, government services are distributed equally by the community to all households in need.	DISTCRIS	0.344	0.408	0.462 *	0.458 *
People in this community should work together to develop the community.	SHOULDEV	0.336	0.573	0.598 *	0.656 *

Table 3. *Cont.*

Factors and Associated Items	Item	Final EFA—Factor Pattern Coefficients (n_{M1} = 555)	Final EFA—Factor Structure Coefficients (n_{M1} = 555)	Baseline CFA (n_{M2} = 550)	Refined † CFA (n_{M2} = 550)
Factor 5: Community leadership (average factor loading = 0.590; average structure coefficient = 0.713; average factor loading on refined CFA = 0.720)					
Formal administrative leaders, like the *kebele* manager, provide support to this community.	ACTLDR2	0.871	0.814	0.721 *	0.720 *
This community's leaders can be trusted.	TRUSTLDR	0.732	0.823	0.899 *	0.899 *
There are people in this community who show strong leadership.	UNOFLDRS	0.372	0.634	0.733 *	0.734 *
I typically accept advice from others in this community.	ADVICE	0.385	0.581	-	-
Factor 6: Associational participation (average factor loading = 0.702; average structure coefficient = 0.825; average factor loading on refined CFA = 0.913)					
I participate in activities held by any community−based associations, such as the Edir.	PARTCBGP	−0.792	−0.883	0.953 *	0.953 *
I attend meetings of a community−based association, such as the Edir.	ACTCBGP	−0.794	−0.897	0.960 *	0.960 *
I participate in activities held by any government or NGO−initiated community development group, such as the Development Army.	ACTEXOGP	−0.520	−0.694	0.828 *	0.827 *
Factor 7: Community organisation (average factor loading = 0.665; average structure coefficient = 0.697; average factor loading on refined CFA = 0.872)					
The community−based associations, such as the Edir, in this community are very active.	COMACTCG	−0.874	−0.883	0.915 *	0.918 *
The leaders of community−based associations, like Edir leaders, respond to this community's concerns.	ACTLDR1	−0.803	−0.840	0.924 *	0.919 *
People in this community get to choose the leaders of their own community−based associations, such as the Edir leaders.	CHOCGLDR	−0.729	−0.802	0.785 *	0.780 *
In this community, people prioritise their own family's welfare over community development.	OWNWELF	−0.569	−0.575	0.155 *	-
Some households in this community are restricted from community services, such as bed net distribution.	RESTRSER	0.352	0.384	−0.260 *	-

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters; *Extraction*: Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; *Rotation*: Promax; * $p \leq 0.05$; † Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings.

Table 4. Factor loadings for random split-half samples for EFA and CFA of EFA-derived factor solutions, women.

Factors and Associated Items	Item	Final EFA—Factor Pattern Coefficients (n_{W1} = 366)	Final EFA—Factor Structure Coefficients (n_{W1} = 366)	Baseline CFA (n_{W2} = 360)	Refined † CFA (n_{W2} = 360)
Factor 1: Social networks and reciprocity (average factor loading = 0.650; average structure coefficient = 0.725; average factor loading on refined CFA = 0.703)					
In this community, I have friends with whom I can share my problems.	HAVEFRND	0.986	0.851	0.863 *	0.863 *
My neighbours sometimes come to me to share their problems and get help.	COME4HLP	0.921	0.842	0.874 *	0.873 *
If you suddenly need some money, you can borrow from a person or group in your community.	BORMONEY	0.617	0.725	0.688 *	0.689 *
If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	NEICAREG	0.755	0.789	0.599 *	0.598 *
This is a close-knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.542	0.702	0.167	-
I typically accept advice from others in this community.	ADVICE	0.529	0.641	0.616 *	0.616 *
The people of this community will contribute their own money or labour for community development.	CONTRDEV	0.348	0.580	0.684 *	0.683 *
If someone in this community loses a cow or goat, a neighbour will help look for it.	LOSTCOW	0.498	0.672	0.603 *	0.602 *
Factor 2: Social disorder (average factor loading = 0.573; average structure coefficient = 0.546)					
In this community, conflicts like stealing and fighting often occur.	CRIMECON	0.801	0.740	0.366 *	-
In this community, you have to be careful, otherwise your neighbours may cheat you.	CHEATS	0.532	0.500	0.213 *	-
Differences between people, such as the amount of land they own, often causes problems in this community.	DIFPROBS	0.386	0.397	0.900 *	-
Factor 3: Social attachment and personal agency (average factor loading = 0.690; average structure coefficient = 0.793; average factor loading on refined CFA = 0.793)					
Being a member of this community is part of who I am.	IDENTITY	0.907	0.921	0.866 *	0.866 *
I feel proud to be part of this community.	PROUD	0.828	0.906	0.837 *	0.836 *
I feel attached to this community and its people.	ATTACH	0.767	0.850	0.900 *	0.899 *
People in this community accept me as a member of the community.	ACCEPT	0.683	0.806	0.866 *	0.866 *
I have the capacity to achieve my future aims.	SELFEFF	0.521	0.667	0.610 *	0.616 *
I have the ability to contribute to this community's development.	SEDEV	0.436	0.607	0.674 *	0.676 *
Factor 4: Social response (average factor loading = 0.526; average structure coefficient = 0.639; average factor loading on refined CFA = 0.656)					
Most people in this community have similar beliefs about what is right and what is wrong.	SIMBLIEF	0.793	0.594	0.309 *	-
If the people of this community see crime-like activities, they will do something about it.	INTERCRI	0.586	0.579	0.410 *	0.403 *
People in this community can be trusted.	COMTRUST	0.583	0.701	0.744 *	0.736 *
When there is a problem in this community, people come together to discuss how it should be solved.	COMPRSLV	0.493	0.769	0.861 *	0.851 *
People in this community live in harmony with each other most of the time.	HARMONY	0.462	0.587	0.532 *	0.524 *
If there is a big dispute between two persons, other people from the community will help in solving the problem.	SLVDISPU	0.461	0.642	0.642 *	0.632 *
This is a close-knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.443	0.678	0.660 *	0.810 *
If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	HLPCRPDZ	0.385	0.562	0.643 *	0.634 *

Table 4. *Cont.*

Factors and Associated Items	Item	Final EFA—Factor Pattern Coefficients (n_{W1} = 366)	Final EFA—Factor Structure Coefficients (n_{W1} = 366)	Baseline CFA (n_{W2} = 360)	Refined † CFA (n_{W2} = 360)
Factor 5: Associational participation [in community structures] (average factor loading = 0.784; average structure coefficient = 0.795; average factor loading on refined CFA = 0.802)					
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	ACTEXOGP	0.870	0.826	0.809 *	0.808 *
I attend meetings of a community-based association, such as the Edir.	ACTCBGP	0.847	0.854	0.761 *	0.761 *
I participate in activities held by any community-based associations, such as the Edir.	PARTCBGP	0.636	0.704	0.835 *	0.836 *
Factor 6: Common vision (average factor loading = 0.643; average structure coefficient = 0.738; average factor loading on refined CFA = 0.720)					
Most people in this community have similar hopes about the future development of the community.	SIMHOPES	0.898	0.882	0.753 *	0.753 *
People in this community share the same ideas on how village matters should be managed.	COMMGMT	0.718	0.821	0.818 *	0.817 *
Most people in this community have common values, for example, they value hard work.	COMMVALU	0.636	0.729	0.738 *	0.740 *
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.542	0.730	0.765 *	0.765 *
During a crisis situation, such as a drought, government services are distributed equally by the community to all households in need.	DISTCRIS	0.422	0.526	0.527 *	0.526 *
Factor 7: Community organisation and leadership (average factor loading = 0.649; average structure coefficient = 0.768; average factor loading on refined CFA = 0.777)					
The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	ACTLDR1	0.919	0.918	0.835 *	0.835 *
The community-based associations, such as the Edir, in this community is very active.	COMACTCG	0.821	0.802	0.774 *	0.773 *
Formal administrative leaders, like the *kebele* manager, provide support to this community.	ACTLDR2	0.549	0.710	0.668 *	0.668 *
People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.	CHOCGLDR	0.469	0.685	0.822 *	0.822 *
There are people in this community who show strong leadership.	UNOFLDRS	0.489	0.724	0.788 *	0.788 *

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters; *Extraction*: Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; *Rotation*: Promax * $p \leq 0.05$; † Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings.

Table 5. Final collective efficacy scales, and comparison of single-group (men vs. women) CE factors structures.

Factor *	Item	Survey Item (i.e., Indicator Prompt)	Facets Tapped	Scale
Social response	CLOSE	This is a close-knit community (i.e., people in this community have close personal relationships with each other).	Strength of social bonds within collective/community	P, M, W
	COMPRSLV	When there is a problem in this community, people come together to discuss how it should be solved.	Group problem-solving, conflict-resolution	P, M, W
	COMTRUST	People in this community can be trusted.	Trust in collective/community members	P, M, W
	HLPCRPDZ	If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	Propensity to address community-wide issues, conflict-resolution	P, M, W
	SLVDISPU	If there is a big dispute between two persons, other people from the community will help in solving the problem.	Propensity to address sub-community issues, conflict-resolution	P, M, W
	HARMONY	*People in this community live in harmony with each other most of the time.*	*Sense of harmony within the collective/community*	*M*, W ‖,¶
	SIMBLIEF	*Most people in this community have similar beliefs about what if right and what is wrong.*	*Collective morals*	*M*
	SHAREKNO	*People in this community share knowledge with their neighbour if they learn something new.*	*Information sharing, diffusion of knowledge in collective*	*M*
	INTERCRI	*If the people of this community see crime-like activities, they will do something about it.*	*Willingness to intervene*	*W*
Social networks and personal agency	COME4HLP †	My neighbours sometimes come to me to share their problems and get help.	Reciprocity of individual-level problem-solving	P, M, W
	BORMONEY	If you suddenly need some money, you can borrow from a person or group in your community.	Responsiveness of social networks, expectations that help will be given/received by others when in need, cooperating to support one another for one-sided or mutual gain §	P, M, W
	NEICAREG	If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.		P, M, W
	SEDEV †	I have the ability to contribute to this community's development.	Individual-level behavioural control over contribution to collective/group goal attainment	P, M, W ‖
	SELFEFF †	I have the capacity to achieve my future aims.	Individual behavioural control of personal goal attainment	P, M, W ‖
	HAVEFRND	*In this community, I have friends with whom I can share my problems.*	*Availability of support networks for individual-level problem solving*	*W*
	ADVICE	*I typically accept advice from others in this community.*	*Willingness to receive, access to guidance from others*	*W*
	CONTRDEV	*The people of this community will contribute their own money or labour for community development.*	*Common moral principles and codes of behaviour*	*W*
	LOSTCOW	*If someone in this community loses a cow or goat, a neighbour will help look for it.*	*Responsiveness of social networks, expectations help will be received, individuals cooperating to support each other*	*W*

Table 5. *Cont.*

Factor *	Item	Survey Item (i.e., Indicator Prompt)	Facets Tapped	Scale
Community organisation and leadership	ACTLDR1 ‡	The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	Responsiveness, strength of leaders of endogenous community structures to community concerns	P, M, W
	COMACTCG ‡	The community-based associations, such as the Edir, in this community is very active.	Activity level of endogenous community structures	P, M, W
	CHOCGLDR ‡	People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.	Selected representation, civic engagement in endogenous structures	P, M, W
	UNOFLDRS	There are people in this community who show strong leadership.	Presence of individuals demonstrating leadership	P, W, M ‖
	ACTLDR2 §	Formal administrative leaders, like the *kebele* manager, provide support to this community.	Supportive leaders of exogenous community structures	P, W, M ‖
	TRUSTLDR	*This community's leaders can be trusted.*	*Social trust in community leaders.*	*M*
Associational participation	ACTCBGP †,‡	I attend meetings of a community-based association, such as the Edir.	Personal membership/participation, endogenous community structures	P, M, W
	PARTCBGP †,‡	I participate in activities held by any community-based associations, such as the Edir.	Personal involvement/participation in endogenous group activities	P, M, W
	ACTEXOGP †,§	I attend the meetings of any government or NGO-initiated community development group, such as the Development Army.	Personal membership/participation, exogenous community structures	P, M, W
Social attachment	ACCEPT	People in this community accept me as a member of the community.	Social acceptance within the collective/community	P, M, W
	IDENTITY †	Being a member of this community is part of who I am.	Place identity, sense of belonging	P, M, W
	ATTACH †	I feel attached to this community and its people.	Place attachment	P, M, W
	PROUD	*I feel proud to be part of this community.*	*Pride in being a member of the collective/community*	*W*
	COMMGMT	People in this community share the same ideas on how village matters should be managed.	Collective ideals, common civic culture	P, M, W
	SIMHOPES	Most people in this community have similar hopes about the future development of the community.	Common hopes for community goal attainment	P, M, W
Common vision	COMMVALU	Most people in this community have common values, for example, they value hard work.	Shared values, ethics	P, M, W
	COLLEFF	People in this community have the capacity to make positive changes by coming together.	Collective behavioural control; capacity, autonomy	P, M, W
	DISTCRIS	During crisis situations, such as drought, government services are distributed equally by the community to all households in need.	Equal distribution of exogenous resources during crises	P, M, W
	SHOULDEV	*People in this community should work together to develop the community.*	*Normative expectations regarding collective action*	*M*

Notes: M = men's CE scale, W = women's CE scale, P = Parsimonious CE scale. Items in italicised font appeared in only one gender-specific scale—this meant the item was either absent from one gender-specific scale, or it tapped to a different factor and was re-organised for the purposes of generating a parsimonious scale. Factor labels reflect those from the parsimonious CE scale, and differ slightly in the women's and men's CE scales. * Factor titles reflect CE factors in the parsimonious model; † Self-referent item prompts about the respondent's own sense of self, agency, autonomy, and level of engagement within his/her community—all other items reflect group-referent items prompts about interpersonal and ecological aspects of the respondent's community; ‡ Items that refer to endogenous community structures (e.g., community-initiated associations)—local endogenous groups used as examples, but should be adapted to the given local context; § Items that refer to exogenous community structures (e.g., government, NGO-initiated community associations)—local exogenous groups used as examples, but should be adapted; ‖ Items re-organised from gender-specific models to produce a parsimonious framework—reflects one gender-specific model; ¶ Demonstrated DIF, dropped from final parsimonious scale.

Comparison of men and women's CE models, identification of a parsimonious model

As indicated in Table 5, there was considerable overlap between gender-specific CE measurement models. We therefore identified a parsimonious CE scale that reflected those items included in both men's and women's refined CE measurement models (Table 5). These findings suggest that the mechanisms and processes through which CE operates are similar between men and women in rural Ethiopia, but key differences exist, particularly with regard to the number and nature of constituent factor items (Table 5). The gender-specific models represent more saturated, and slightly better fitting models (Table A5). Those models allow for exploration into specific mechanisms through which CE specifically operates for men and women, respectively.

A subsequent CFA that tested the fit of the men's and women's data to the parsimonious CE measurement model [44] demonstrated good model fit (Table A5). These results suggest that the two gender-specific, saturated models and the parsimonious model all demonstrated construct validity, suggesting they are appropriate metrics for measuring CE.

3.5. Step 5. Testing Hypotheses

3.5.1. Step 5. Methods

Multiple-group CFA and MIMIC modelling for assessment of differential item functioning

We performed multiple-group CFA to examine certain aspects (e.g., factor loadings) of measurement invariance between genders. Measurement invariance means that the probability of selecting a given item response category is comparable across groups, given similar levels of the latent construct being measured [59]. This multiple-group CFA differed from previous single-group, gender-specific CFA in that we were able to simultaneously employ input matrices from both men's and women's datasets. Due to the relatively small proportion of households with leadership status, we did not perform multiple-group analyses on this variable.

Next, we performed Multiple Indicator Multiple Causes (MIMIC) modelling to test the validity of our parsimonious CE measurement model in the presence of other relevant covariates, and assess differential item functioning (DIF) [60]. DIF, or measurement *non*-invariance occurs when people from different groups (e.g., men, women) with similar levels of the latent construct have different probabilities of responding to an item in a certain way [61]. Our structural equation MIMIC models consisted of a measurement model component reflected by the refined parsimonious CE model, and a structural model component that specified the direct effects of gender and household leadership covariates on latent factor variables and relevant item indicators. Significant direct effects would indicate DIF between men and women respondents.

The same validation sub-samples used for single-group CFA were used for these analyses, but we aggregated gender-specific sub-samples (n_2). As we constructed our MIMIC models, we first established baseline models that introduced gender and leadership status covariates, but assumed no direct effects of the covariates on any individual CE items. Then, we employed a step-wise, forward selection approach to assess direct effects between these covariates and relevant item indicators. We examined the modification indices, and identified the item indicator with the highest significant, meaningful, and substantively justifiable modification index. We added a direct path between the identified item indicator and relevant covariate. We employed the DIFFTEST option in Mplus to assess whether the additional direct path improved model fit. Given we had a relatively large sample size, it was likely that DIFFTEST statistics would be significant [62], so we evaluated and compared other model fit indices as well.

Collective efficacy factor score calculation

We used two coarse CE factor score calculation methods (i.e., non-refined, un-sophisticated procedures) to generate both average and weighted average CE factor scores [63]. Higher factor scores

represented higher levels of perceived behavioural control over the respective CE factors. We generated CE factor scores for each respondent by summing his/her responses across all items in each factor (i.e., 1 = completely disagree, 2 = partially disagree, 3 = neither agree nor disagree [neutral], 4 = partially agree, 5 = completely agree), and dividing that sum by the number of items tapping to the factor. This approach, however, assumes that all items have the same level of influence, or measurement proximity to their respective latent factor. We have demonstrated that this is not the case. We generated scores in this format to allow for easy comparison of scores, should the tool be used in different contexts. This is not only a simpler method compared to context-specific weighting, but also allows for more appropriate comparison of results outside of the dataset in which the weights were derived [63]. We also calculated weighted average CE factor scores, for which a weight that was equivalent to the item's factor loading was applied to each item score prior to the generation of the average (weighted) factor score. We examined whether there were statistically significant differences in factor scores by gender and leadership status, and between household-level dyads via a regression-based approach with cluster robust standard errors to adjust for within-village clustering. For dyads, we regressed pair-wise factor score differences with cluster robust standard errors [64].

3.5.2. Step 5. Results

Multiple-group confirmatory factor analysis and MIMIC model results

Multiple-group CFA indicated that factor loadings between men and women were similar (Table 6). We present model fit statistics, unstandardised and standardised beta estimates, and standard errors for competing MIMIC models in Table A6. The baseline MIMIC model with latent variables regressed on gender and household leadership covariates, but no direct effects between item indicators (i.e., Model 3 in Table A6) demonstrated good model fit (χ^2:df ratio = 2.124; RMSEA [90% CI] = 0.035 [0.032–0.039]; CFI = 0.965; TLI = 0.960). Only two items from the baseline MIMIC model with latent factors regressed on gender and household leadership status demonstrated modification indices above 3.84. The HARMONY item indicator had the highest, albeit relatively small, modification index (14.342) on gender. This finding indicated that there was DIF between men and women for this item, so we added a direct path between HARMONY and gender.

On the model iteration specifying this direct path (i.e., Model 4 in Table A6), both the unstandardised B and standardised β were salient (i.e., −0.457 and −0.448, respectively), indicating that men scored HARMONY, on average, 0.45 units lower than women. This refined model fit the data well (χ^2:df ratio = 2.090; RMSEA [90% CI] = 0.035 [0.031–0.038]; CFI = 0.966; TLI = 0.961), and DIFFTEST statistics indicated that model fit improved with the inclusion of this direct effects parameter. The modification indices of the resulting model indicated only one additional item with a low modification index (ACTEXOGP, 3.911), and small but non-salient direct effects. Therefore, no further model refinements were made. The final MIMIC model accommodated uniform DIF by incorporating a direct effect between gender and HARMONY, and indirect effects of gender and household leadership status covariates on factor means.

Final collective efficacy measurement metrics

To avoid DIF and ensure construct validity, we dropped HARMONY from CE measurement model, and re-ran a final CFA [65]. As demonstrated in Table 6, the validated model indicated good fit (χ^2:df ratio = 2.191; RMSEA [90% CI] = 0.036 [0.033–0.040]; CFI = 0.966; TLI = 0.960), with all items loading significantly and saliently a single factor. The final, validated parsimonious CE scale included 26 items tapping to six factors: social response, social networks and personal agency, social attachment, common vision, associational participation, and community organization and leadership.

Table 6. Factor loadings and fit indices for multiple-group CFA, baseline and final MIMIC, and final CFA (with MIMIC refinement) models.

Factors and Associated Items	Item	Multiple-Group Men (n_{EM2} = 550)	Multiple-Group Women (n_{EW2} = 360)	Baseline MIMIC Model (n_{E2} = 907) †	Final MIMIC Model ‡ (n_{E2} = 907) †	Final CFA Model § (n_{E2} = 907) †
Factor 1: Social response (average baseline MIMIC model factor loading = 0.673; average final MIMIC model factor loading = 0.671; average final CFA model with MIMIC deletions = 0.708)						
This is a close-knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.776	0.803	0.782 *	0.783 *	0.773 *
When there is a problem in this community, people come together to discuss how it should be solved.	COMPRSLV	0.695	0.847	0.774 *	0.774 *	0.771 *
If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	HLPCRPDZ	0.712	0.638	0.680 *	0.680 *	0.674 *
People in this community can be trusted.	COMTRUST	0.688	0.727	0.703 *	0.703 *	0.691 *
If there is a big dispute between two persons, other people from the community will help in solving the problem.	SLVDISPU	0.649	0.624	0.635 *	0.635 *	0.629 *
People in this community live in harmony with each other most of the time.	HARMONY	0.422	0.531	0.462 *	0.453 *	-
Factor 2: Social networks and personal agency (average baseline MIMIC model factor loading = 0.663; average final MIMIC model factor loading = 0.663; average final CFA model = 0.663)						
If you suddenly need some money, you can borrow from a person or group in your community.	BORMONEY	0.740	0.667	0.724 *	0.724 *	0.724 *
My neighbours sometimes come to me to share their problems and get help.	COME4HLP	0.758	0.748	0.755 *	0.755 *	0.754 *
If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	NEICAREG	0.622	0.562	0.607 *	0.607 *	0.607 *
I have the capacity to achieve my future aims.	SELFEFF	0.533	0.555	0.554 *	0.554 *	0.555 *
I have the ability to contribute to this community's development.	SEDEV	0.700	0.654	0.676 *	0.676 *	0.677 *
Factor 3: Social attachment (average baseline MIMIC model factor loading = 0.885; average final MIMIC model factor loading = 0.885; average final CFA model = 0.884)						
I feel attached to this community and its people.	ATTACH	0.850	0.915	0.872 *	0.872 *	0.871 *
People in this community accept me as a member of the community.	ACCEPT	0.864	0.879	0.892 *	0.892 *	0.894 *
Being a member of this community is part of who I am.	IDENTITY	0.880	0.835	0.890 *	0.890 *	0.888 *
Factor 4: Common vision (average baseline MIMIC model factor loading = 0.737; average final MIMIC model factor loading = 0.737; average final CFA model = 0.737)						
People in this community share the same ideas on how village matters should be managed.	COMMGMT	0.887	0.815	0.854 *	0.854 *	0.854 *
Most people in this community have similar hopes about the future development of the community.	SIMHOPES	0.834	0.766	0.811 *	0.811 *	0.813 *
Most people in this community have common values, for example, they value hard work.	COMMVALU	0.808	0.743	0.784 *	0.784 *	0.784 *
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.741	0.760	0.755 *	0.755 *	0.756 *
During crisis situations, such as a drought, government services are distributed equally by the community to all households in need.	DISTCRIS	0.462	0.510	0.481 *	0.481 *	0.478 *

Table 6. *Cont.*

Factors and Associated Items	Item	Multiple-Group Men (n_{EM2} = 550)	Multiple-Group Women (n_{EW2} = 360)	Baseline MIMIC Model (n_{E2} = 907) †	Final MIMIC Model ‡ (n_{E2} = 907) †	Final CFA Model § (n_{E2} = 907) †
Factor 5: Associational participation (average baseline MIMIC model factor loading = 0.874; average final MIMIC model factor loading = 0.874; average final CFA model = 0.874)						
I attend meetings of a community-based association, such as the Edir.	ACTCBGP	0.979	0.769	0.886 *	0.886 *	0.887 *
I participate in activities held by any community-based associations, such as the Edir.	PARTCBGP	0.933	0.831	0.912 *	0.912 *	0.911 *
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	ACTEXOGP	0.830	0.803	0.825 *	0.825 *	0.825 *
Factor 6: Community organisation and leadership (average baseline MIMIC model factor loading = 0.766; average final MIMIC model factor loading = 0.766)						
The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	ACTLDR1	0.806	0.846	0.829 *	0.829 *	0.829 *
The community-based associations, such as the Edir, in this community are very active.	COMACTCG	0.813	0.778	0.812 *	0.812 *	0.813 *
People in this community get to choose the leaders of their own community-based associations, such as Edir leaders.	CHOCGLDR	0.692	0.810	0.750 *	0.750 *	0.752 *
Formal administrative leaders, like the *kebele* manager, provide support to this community.	ACTLDR2	0.652	0.672	0.669 *	0.669 *	0.667 *
There are people in this community who show strong leadership.	UNOFLDRS	0.762	0.779	0.769 *	0.769 *	0.768 *
Model fit statistics						
χ^2 (df)		1197 (714)		746 (351)	731 (350)	710 (324)
χ^2 contribution from each group (for multiple group CFA)		598.570	598.907	N/A	N/A	N/A
χ^2:df		1.676		2.124	2.090	2.191
RMSEA		0.039		0.035	0.035	0.036
(90% CI)		(0.035–0.042)		(0.032–0.039)	(0.031–0.038)	(0.033–0.040)
CFI		0.963		0.965	0.966	0.966
TLI		0.964		0.960	0.961	0.960

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters * two-tailed $p \leq 0.05$; † Three observations excluded from the MIMIC model due to missing covariate data; ‡ Final MIMIC model reflects refined, parsimonious CE measurement model with latent variables regressed on gender and household leadership status plus the inclusion of a direct path between HARMONY and gender; § Final CFA model reflects refined, parsimonious CE measurement model with HARMONY deleted due to DIF.

Assessment of collective efficacy scores across cohorts

Overall, men scored all CE factors significantly higher than women (Figure 1), suggesting men have higher behavioural control perceptions than women. Men and women in household-level male-female dyads scored all CE factors significantly different, with the exception of social response ($\beta_{unweighted}$ = 0.07, [95% CI: −0.02, 0.16]; $\beta_{weighted}$ = 0.05, [95% CI: −0.02, 0.11]). This indicates CE perceptions may differ significantly within households. Factor scores only differed significantly between respondents with leadership roles and those without on two factors: social networks and personal agency ($\beta_{unweighted}$ = 0.10, [95% CI: 0.03, 0.17]); $\beta_{weighted}$ = 0.07, [95% CI: 0.02, 0.12]), and associational participation ($\beta_{unweighted}$ = 0.15, [95% CI: 0.02, 0.28]; $\beta_{weighted}$ = 0.13, [95% CI: 0.02, 0.24]).

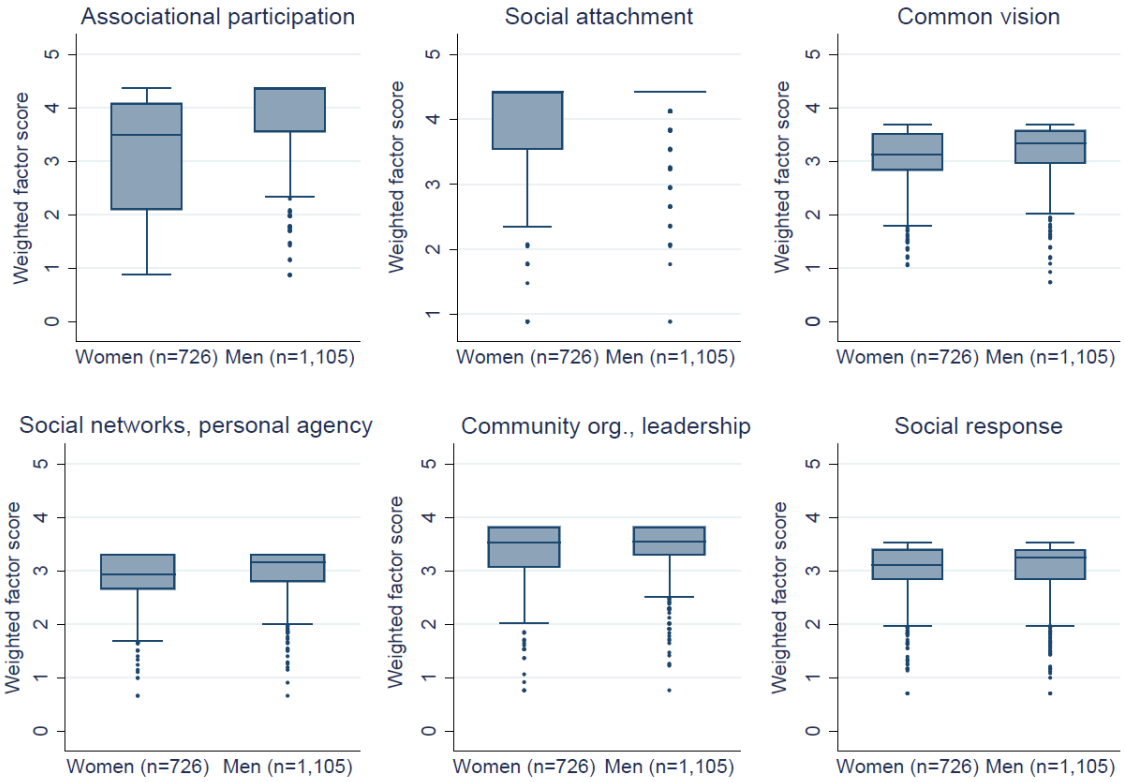


Figure 1. Collective efficacy factor scores (weighted), by respondent gender. Factor scores are visualised as box plots, which depict the distribution of the data through quartiles. The boxes represent the inter-quartile range (i.e., 25% and 75% quartiles comprise the outer edges of the boxes, while the median is indicated by the line inside the box). The lines that extend vertically from either side of the box (i.e., whiskers) indicate the variability of the data outside the upper and lower quartiles. Outliers are plotted as individual points.

4. Discussion

This study contributes to the development of a metric that can be used in community-based health and development programmes, to inform intervention design, identify communities ripe for programmatic targeting, and diagnose factors related to intervention effectiveness. Such a metric may be useful for any programme that targets collective behaviours. It may be particularly beneficial for the WASH sector, as differentials in CE factors may help explain poor uptake of community-based WASH interventions, regression to unimproved behaviours, and lower than expected health gains.

The structures derived through our factor analytical approach reflect rigorously derived measurement models that are grounded in theory and evidence-based. Findings from our exploratory analyses suggested that CE is a complex, multi-dimensional social construct. EFA-derived

factor structures suggest that social response, social networks, social attachment, common vision, associational participation, and community organization and leadership are important factors in the measure of CE among men and women in rural Ethiopia. These factors reflect domains related to social cohesion, informal social control, and cognitive and structural social capital.

Elucidating these CE factors and examining their constituent sub-constructs is important to consider for the design of intervention content and development of implementation strategies. In terms of content, failing to acknowledge and address these CE factors as part of community-based intervention approaches may be problematic for the uptake of such interventions. For instance, a lack of common vision may prove to be a barrier to communities accepting WASH facilities coverage and use targets as communal goals. If the issue of common vision is not addressed alongside mainstream WASH intervention activities, the intervention may fail to stimulate the collective action and cooperative behaviour necessary to achieve programmatic goals at the community level. When collective goals are set, progress toward them may be inhibited in the presence of inadequate or insufficiently nurtured supporting social networks, which may limit the diffusion of innovations within a community [66]. It may be beneficial for interventions to leverage and strengthen social networks to facilitate uptake of improved WASH practices while also addressing personal agency and social inclusion of women within community structures. Addressing these CE factors within the context of community-based interventions may therefore create an environment more conducive to engendering and maintaining positive change. In terms of implementation strategies, it is important to consider CE factors and differentials in factor scores when determining at whom interventions are targeted, and through which mechanisms they are being delivered. Interventions that target programme participants who have low perceptions of self- and collective agency, and are not well positioned to serve as change agents within their households, social networks, and community may prove to be ineffective. Poor associational participation among women may impede adoption and maintenance of improved WASH behaviours, particularly those aimed at infants and young children, when interventions are delivered through community associations or groups. Other mechanisms and intervention techniques, such as one-on-one skills-based household counselling visits may serve to better facilitate behavioural uptake while also enhancing action knowledge and improving self-efficacy perceptions in contexts where women do not readily engage with community groups. Several CE factors may need to be addressed through different types of intervention techniques, which should be considered from the outset of intervention design.

The CE structures revealed by our EFA analyses differed slightly from our hypothesised framework. For example, in our hypothesised CE framework, we conceived that agency would emerge as an independent CE factor/dimension encompassing items related to individual- and collective-level perceptions. However, results from our factor analyses indicated that these agentic concepts are closely tied to social networks and social attachment for men and women, respectively. These findings may suggest that agency perceptions are constituent influencers of these sub-constructs as opposed to being important stand-alone CE factors. On the other hand, agency may not have emerged as an independent factor in our EFA given some aspects of autonomy control beliefs, motivational commitment to communal goals, resilience to adversity, and performance accomplishments were not fully represented within the CE survey. The literature suggests an array of conceptual definitions for CE and its constituent sub-constructs, and our EFA-derived factor structures did not deviate considerably from our hypothesised framework. Therefore, we found sufficient theoretical justification to support our refined, EFA-derived factor structures.

Our EFA exposed some key differences in gender-specific factor structures. We used a two-pronged approach to compare CE among men and women by first examining gender-specific CE mechanisms via single-group measurement models, and then assessing perceptions of common CE factors between genders via comparison of factor scores. One of the most notable distinctions with regard to underlying CE mechanisms (i.e., factor structures) was that personal agency was linked to different sub-constructs in the two gender-specific measurement models. These findings suggest that for women, one's sense of self-agency is linked to one's sense of belonging or social attachment,

while for men, it is linked to expectations regarding the responsiveness of one's social networks. This may be a result of a higher level of social inclusion, mobility, and ability of men to engage more readily with both formal and informal community structures in the rural Ethiopian context. We also observed that community leadership emerged as its own CE factor in the men's CE measurement model, while these items loaded to the "community organisation and leadership" factor in the women's model. This may be an artefact of men having more access to and engagement with people holding such positions within the community.

Our scales produced quantitative measures of CE for all respondents in the form of factor scores for each CE factor. Higher scores reflected higher levels of perceived behavioural control over the respective CE sub-constructs. These findings revealed that, overall, men scored higher on all CE factors than women. This is notable not only for gendered WASH interventions that may be targeting female caregivers as their primary programme participants, but any community-based programme seeking to address gender and women's empowerment, more broadly. For WASH interventions specifically, the sector should note that our findings suggest women may not inherently be in a position to influence adoption of promoted behaviours and practices readily among their peers, across all contexts. In order for such interventions to effectively engage women, they may need to address aspects of perceived agency and social inclusion alongside programme-specific objectives. These findings corroborate existing evidence [24,26,67] that women often suffer from various "gendered structures of constraint", including limitations related to social inclusion, civic engagement, and membership and participation in community structures. CE factor scores suggest that differences in the social inclusion of men and women in rural Ethiopia appear to have created disparities in terms of their perceived individual- and community-level behavioural control perspectives.

We observed differences between men's and women's CE perceptions for a variety of reasons. Becoming aware of the disparities in CE factor scores should influence sampling methodology. Implementers interested in assessing CE and targeting or evaluating programmes that operate at a collective level should measure CE among both men and women to yield data that are representative of the larger programme population. Such a sampling approach will offer a more holistic, less biased appraisal of CE that accounts for heterogeneity of perspectives. This can be done through the employment of our parsimonious scale, which is suited for use among both men and women. CE scale results can be used to inform the design of community-based interventions. Findings from CE surveys may help identify the specific CE factors that need to be addressed, across genders, depending on an intervention's target audience (e.g., women, men, both women and men).

Although our results suggest a more comprehensive and complex underlying CE factor structure than previous studies, these findings do corroborate results from some of the existing literature. When considering the existing empirical evidence, certain sub-constructs seem to transcend contexts, languages, and culture when it comes to the measure of CE. Two of the factors proposed from a study examining CE in the context of community computing in Blacksburg, Virginia, USA [35]—belonging or social attachment and association (associational participation, in our case)—were also indicated in our refined gender-specific and final, parsimonious CE measurement scales. While activism and informedness, the two other factors identified through that study, did not emerge as factors in our analyses, items that represented these concepts were included in our refined gender-specific and final, parsimonious CE scales. Our findings suggest that social cohesion and informal social control domains proposed by authors of a study investigating CE and violent crime in urban Chicago, Illinois, USA [19] are important for the measure of CE, but do not necessarily manifest as factors themselves.

Our more comprehensive factor structures were substantively justified, and suggest that our CE scales include additional factors and items not included in other metrics. This is likely an artefact of the emphasis we placed on the activities conducted during the first two steps of our scale development process. To our knowledge, based on information provided in the literature, previous CE scale development studies did not include or heavily emphasise these scale development steps. As a result, our CE scales reveal more CE factors than the alternatives, and may therefore allow for a more accurate

measure of CE and its effect on important behavioural outcomes at the community level. It seems fitting that a construct as complex as CE would need to draw on a more nuance underlying structure to ensure content validity.

This study had several procedural and analytical strengths and limitations. Our scale development and validation approach was a strength in that it reflected a mixed-methodological process that included focus group discussions and cognitive validation. We actively reflected on these qualitative data, which we used along with other theoretical and empirical evidence to make evidence-based, substantively justified modelling decisions. The size of our female sample may have been a limitation. While our gender-specific split-half sub-samples met common sample size guidelines, our sample size-dependent model fit statistics (e.g., χ^2 p-value) indicated the sample sizes may have only been borderline sufficient for a factor structure as complex as those related to our CE measurement models. While our deliberation of possible factor models was heavily informed by substantive considerations, there is a dearth of existing theoretical and empirical evidence on collective efficacy in low literacy and resource poor settings such as those in which this research was carried out. Our own qualitative evidence aside, it is possible that decisions based on substantive considerations were not appropriate in these contexts. As such, the factor structures identified through our analyses may not necessarily generalise to other populations. Similarly, these findings may not translate to individuals and populations in which mental well-being is poor, and depression and anxiety are common, as these factors may interact with other behavioural determinants to influence behavioural control perceptions.

The external validity of our proposed CE measurement metrics, including their ecological (generalisability to other settings), population (generalisability to other people), and historical (generalisability over time) validity requires further examination [68]. We agree with our colleagues from the water insecurity sector that context-specific scales are advisable [69–71]. We recognise that the underlying structure of CE, as an inferential process, may differ substantially from context to context, as related sub-constructs are largely informed by context-specific political economies and social schemas. For instance, factors related to empirical and normative expectations regarding cooperative behaviour that likely inform perceptions regarding agentic concepts such as self- and collective efficacy are steeped in rich historical and cultural traditions may not be comparable on a global scale. In some contexts, women may be more integrated into endogenous community structures, so gendered differentials in associational participation, social networks, and perceptions regarding social attachment may be less pronounced in those settings. In environments such as India, where caste and class structures are important, different CE-related factors, such as social disorder may have a stronger influence on personal and collective-level behaviour control perceptions [72]. Findings from additional formative research and psychometric assessment efforts both in Ethiopia and elsewhere can; however, enhance our early CE work, further assess the validity of our CE scale, and reveal whether various CE sub-constructs transcend contexts.

5. Conclusions

Our CE scales offer new tools for the examination of collective behaviour factors. These tools can be used for programmatic targeting, intervention design, and diagnostic investigations into the role CE factors play in the uptake of community-based interventions and their impacts on health and development. They also facilitate the generation of evidence related to factors falling along the causal chain, which may explain why biologically plausible health gains are not being achieved by WASH interventions, as expected. Important differences in perceptions related to CE factors among men and women exist. These disparities should be acknowledged and addressed in the design of intervention content and implementation strategies for community-based interventions, particularly those promoting improved WASH behaviours for infants, young children, and their caregivers.

Supplementary Materials: The following are available online at http://www.mdpi.com/1660-4601/15/10/2139/s1, Table S1: Univariate descriptive statistics: Frequency of responses by split-halves and gender.

Author Contributions: Conceptualization, M.G.D.; Methodology, M.G.D., G.D.S., T.F.C., B.A.C., R.H., M.C.F., R.D.; Formal Analysis, M.G.D.; Investigation, M.G.D., G.D.S.; Resources, R.H., C.L.N., R.D., M.W.; Data Curation, M.G.D.; Writing-Original Draft Preparation, M.G.D.; Writing-Review and Editing, G.D.S., M.W., R.H., C.L.N., B.A.C., R.D., A.G.G., T.F.C., M.C.F.; Visualization, M.G.D.; Supervision, T.F.C., M.C.F.; Project Administration, M.G.D., A.G.G., M.C.F.; Funding Acquisition, M.C.F., T.F.C., M.G.D.

Funding: This research was funded by The World Bank Group's Strategic Impact Evaluation Fund (SIEF) contract number [7175829], and the Bill & Melinda Gates Foundation grant number [OPP1008048].

Acknowledgments: The authors would like to thank Resom Berhe for his assistance conducting cognitive interviews, Jedidiah Snyder and Kassahun Zewudie for their assistance coordinating and supervising data collection, and Mulusew Belew for his management support. We are grateful for the support we have received from numerous partners from the Democratic Republic of Ethiopia's Federal Ministry of Health; the Amhara Regional Health Bureau; South Gondar and West Gojjam Zonal Health Departments; the Bahir Dar Zuria, Fogera, and Farta Woreda Health Offices; and the Health Extension Workers, Women's Development Arm Leaders, and Health Development Army members from our study *kebeles*. We would like to acknowledge and thank the cadre of field supervisors and enumerators who captured these data. Last, but certainly not least, we would like to thank our *Andilaye* study participants, who generously gave their time to participate in our formative work and respond to our CE survey.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Administered 50-Item Collective Efficacy Tool.

Hypothesised Domain	Hypothesised Factors	Item Name	Survey Item (i.e., Prompt)	Description of Hypothesised Facets
Informal social control	Social disorder	HARMONY	People in this community live in harmony with each other most of the time.	Sense of harmony within the collective/community
		CHEATS	*In this community, you have to be careful, otherwise your neighbours may cheat you.*	Perceived presence of deceitful individuals
		CRIMECON	*In this community, conflicts like stealing and fighting often occur.*	Perceived presence of incivility, delinquent behaviour
		SAFEATHO *	When I am at home alone, I feel safe from threats of crime.	Feeling of safety while at home, in the community
	Social response	SIMBLIEF	Most people in this community have similar beliefs about what is right and what is wrong.	Collective morals
		INTERCRI	If the people of this community see crime-like activities, they will do something about it.	Willingness to intervene
		SLVDISPU	If there is a big dispute between two persons, other people from the community will help in solving the problem.	Community's propensity to address sub-community-level issues, engage in conflict-resolution
		HLPCRPDZ	If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	Community's propensity to address community-wide issues, engage in conflict-resolution
		SUPMOURN	If someone in this community had a death in their family, the community will come together to support them while they mourn.	Social support & comforting
		COMPRSLV	When there is a problem in this community, people come together to discuss how it should be solved.	Group problem-solving, conflict-resolution
		CONTRDEV	The people of this community will contribute their own money or labour for community development.	Common moral principles & codes of behaviour
		DIFPROBS	*Differences between people, such as the amount of land they own, often causes problems in this community.*	Tolerance
		HAPPYNEI *	I feel happy for my neighbour if they have a good harvest.	Vicarious affective feelings—happiness

Table A1. *Cont.*

Hypothesised Domain	Hypothesised Factors	Item Name	Survey Item (i.e., Prompt)	Description of Hypothesised Facets
Social cohesion	Social capital	COMTRUST	People in this community can be trusted.	Perceived trust in collective/community members
		ADVICE *	I typically accept advice from others in this community.	Willingness to receive, access to guidance from endogenous entities/individuals
		SHAREKNO	People in the community share new knowledge with their neighbour if they learn something new.	Information sharing, diffusion of knowledge within the collective
		CLOSE	This is a close-knit community (i.e., people in this community have close personal relationships with each other).	Strength of social bonds within collective/community
		OWNWELF	*In this community, people prioritise their own family's welfare over community development.*	Commitment to collective development, goal attainment
		LOSTCOW	If someone in this community loses a cow or goat, a neighbour will help look for it.	Perceived responsiveness of social networks, expectations that help will be given to/received by others when in need, individuals cooperating to support one another for either one-sided or mutual gain §
		BORMONEY	If you suddenly need some money, you can borrow from a person or group in your community.	
		NEICAREG	If you & your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	
		HAVEFRND *	In this community, I have friends with whom I can share my problems.	Availability of support networks for individual-level problem-solving
		COME4HLP *	My neighbours sometimes come to me to share their problems and get help.	Reciprocity of individual-level problem-solving
		UNOFLDRS	There are people in this community who show strong leadership.	Perceived presence of individuals demonstrating attributes of leadership
		COMACTCG †	The community-based associations, such as the Edir, in this community is very active.	Activity level of endogenous community structures
		ACTLDR1 †	The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	Responsiveness, strength of the leaders of endogenous community structures to community concerns
		ACTLDR2 ‡	Formal administrative leaders, like the *kebele* manager, provide support to this community.	Supportiveness of the leaders of exogenous community structures
		TRUSTLDR	This community's leaders can be trusted.	Perceived social trust in community leaders
		CHOCGLDR †	People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.	Civic engagement in endogenous community structures, community-selected representation
		COPARTCG †,‡	Most people in this community participate in community associations.	Community engagement in endogenous and exogenous community structures
		ACTCBGP *,†	I attend meetings of a community-based association, such as the Edir.	Personal associational membership/participation, endogenous community structures
		PARTCBGP *,†	I participate in activities held by any community-based associations, such as the Edir.	Personal involvement/participation in activities organised by endogenous community structures
		ACTEXOGP *,‡	I attend the meetings of any government or NGO-initiated community development group, such as the Development Army.	Personal associational membership/participation, exogenous community structures
		PAREXOGP *,‡	I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	Personal involvement/participation in activities organised by exogenous community structures

Table A1. *Cont.*

Hypothesised Domain	Hypothesised Factors	Item Name	Survey Item (i.e., Prompt)	Description of Hypothesised Facets
Social cohesion	Social equity	COMMGDEC	When community groups make decisions, they are pleasing and good for most of the households in this community.	Social equity prioritised during community decision-making processes
		BRIBELDR	*Sometimes people need to bribe community leaders in order to get things done.*	Corruption among community leaders
		DISTCRIS	During a crisis situation, such as drought, government services are distributed equally by the community to all households in need.	Equal distribution of exogenous resources during crises
		RESTRSER	*Some households in this community are restricted from community services, such as bed net distribution.*	*Social injustice, restrictions to resources*
	Common values	COMMVALU	Most people in this community have common values, for example, they value hard work.	Shared values, ethics
		SIMHOPES	Most people in this community have similar hopes about the future development of the community.	Common hopes for community goal attainment, performance
		COMMGMT	People in this community share the same ideas on how village matters should be managed.	Collective ideals, common civic culture
	Community attachment	ACCEPT	People in this community accept me as a member of the community.	Social acceptance within the collective/community
		ATTACH *	I feel attached to this community and its people.	Place attachment
		PROUD *	I feel proud to be part of this community.	Pride in being a member of the collective/community
		IDENTITY *	Being a member of this community is part of who I am.	Place identity, sense of belonging
Behavioural control	Agency	SELFEFF *	I have the capacity to achieve my future aims.	Perceived individual-level behavioural control over personal goal attainment
		SEDEV *	I have the ability to contribute to this community's development.	Perceived individual-level behavioural control over contribution to collective/group goal attainment
		COLLEFF	People in this community have the capacity to make positive changes by coming together.	Perceived community-level behavioural control; capacity and autonomy control beliefs
		EXOASSIS	*This community needs assistance from others outside the community in order to make positive changes.*	Perceived reliance on exogenous support to facilitate goal attainment
		SHOULDEV	People in this community should work together to develop the community.	Normative expectations regarding collective action

Notes: Indicated sub-constructs reflect those conceptualised via our hypothesised collective efficacy framework. Items presented in italicised font were hypothesised to have an inverse relationship with CE. Given the various conceptualisations of these latent constructs, substantive justification existed for the re-conceptualisation articulated in our EFA-derived factor structures. * self-referent item prompts about the respondent's own sense of self, agency, autonomy, and level of engagement within his/her community; all other items reflect group-referent items prompts about ecological aspects of the respondent's community; † items that refer to endogenous community structures (e.g., community-initiated/organised community associations)—local endogenous structures used as examples, but should be adapted to the given local context; ‡ items that refer to exogenous community structures (e.g., government or NGO-initiated/organised community associations)—local exogenous structures used as examples, but should be adapted to the given context; § measured through different scenarios reflecting different levels of urgency/need.

Table A2. Factor Loadings from Preliminary CFA of Our Hypothesised CE Framework, Men.

Factors and Associated Items	Item	Initial Prelim. CFA (n_M = 1105)	Refined † Prelim. CFA (n_M = 1105)
Factor 1: Agency (average factor loading on refined CFA = 0.735)			
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.804 *	0.799 *
I have the ability to contribute to this community's development.	SEDEV	0.775 *	0.778 *
People in this community should work together to develop the community.	SHOULDEV	0.747 *	0.744 *
I have the capacity to achieve my future aims.	SELFEFF	0.612 *	0.617 *
This community needs assistance from others outside the community in order to make positive changes.	EXOASSIS	0.011	-
Factor 2: Common values (average factor loading on refined CFA = 0.831)			
Most people in this community have common values, for example, they value hard work.	COMMVALU	0.779 *	0.778 *
People in this community share the same ideas on how village matters should be managed.	COMMGMT	0.877 *	0.877 *
Most people in this community have similar hopes about the future development of the community.	SIMHOPES	0.838 *	0.839 *
Factor 3: Social response (average factor loading on refined CFA = 0.573)			
If the people of this community see crime-like activities, they will do something about it.	INTERCRI	0.446 *	0.444 *
When there is a problem in this community, people come together to discuss how it should be solved.	COMPRSLV	0.718 *	0.719 *
If there is a big dispute between two persons, other people from the community will help in solving the problem.	SLVDISPU	0.710 *	0.712 *
If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	HLPCRPDZ	0.689 *	0.686 *
The people of this community will contribute their own money or labour for community development.	CONTRDEV	0.660 *	0.660 *
If someone in this community had a death in their family, the community will come together to support them while they mourn.	SUPMOURN	0.520 *	0.521 *
I feel happy for my neighbour if they have a good harvest.	HAPPYNEI	0.445 *	0.442 *
Most people in this community have similar beliefs about what is right and what is wrong.	SIMBLIEF	0.399 *	0.399 *
Differences between people, such as the amount of land they own, often causes problems in this community.	DIFPROBS	−0.267 *	-
Factor 4: Social order (average factor loading on refined CFA = 0.539)			
People in this community live in harmony with each other most of the time.	HARMONY	0.818 *	0.984 *
In this community, you have to be careful, otherwise your neighbours may cheat you.	CHEATS	−0.304 *	−0.305 *
In this community, conflicts like stealing and fighting often occur.	CRIMECON	−0.284 *	−0.327 *
When I am at home alone, I feel safe from threats of crime.	SAFEATHO	0.150 *	-
Factor 5: Social capital (average factor loading on refined CFA = 0.671)			
People in this community can be trusted.	COMTRUST	0.631 *	0.631 *
I typically accept advice from others in this community.	ADVICE	0.591 *	0.593 *
People in the community share new knowledge with their neighbour if they learn something new.	SHAREKNO	0.646 *	0.648 *
This is a close-knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.677 *	0.676 *
If someone in this community loses a cow or goat, a neighbour will help look for it.	LOSTCOW	0.535 *	0.533 *
If you suddenly need some money, you can borrow from a person or group in your community.	BORMONEY	0.605 *	0.609 *
If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	NEICAREG	0.588 *	0.588 *
There are people in this community who show strong leadership.	UNOFLDRS	0.660 *	0.656 *
The community-based associations, such as the Edir, in this community are very active.	COMACTCG	0.622 *	0.626 *

Table A2. *Cont.*

Factors and Associated Items	Item	Initial Prelim. CFA (n_M = 1105)	Refined † Prelim. CFA (n_M = 1105)
The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	ACTLDR1	0.615 *	0.614 *
Formal administrative leaders, like the *kebele* manager, provide support to this community.	ACTLDR2	0.619 *	0.610 *
This community's leaders can be trusted.	TRUSTLDR	0.766 *	0.764 *
People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.	CHOCGLDR	0.521 *	0.520 *
In this community, I have friends with whom I can share my problems.	HAVEFRND	-	-
My neighbours sometimes come to me to share their problems and get help.	COME4HLP	0.633 *	0.637 *
Most people in this community participate in community associations.	COPARTCG	0.700 *	0.698 *
I attend meetings of a community-based association, such as the Edir.	ACTCBGP	0.881 *	0.885 *
I participate in activities held by any community-based associations, such as the Edir.	PARTCBGP	0.888 *	0.892 *
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	ACTEXOGP	0.804 *	0.804 *
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	PAREXOGP	0.781 *	0.781 *
In this community, people prioritise their own family's welfare over community development.	OWNWELF	0.139 *	-
Factor 6: Social equity (average factor loading on refined CFA = 0.573)			
When community groups make decisions, they are pleasing and good for most of the households in this community.	COMMGDEC	0.728 *	-
During a crisis situation, such as a drought, government services are distributed equally by the community to all households in need.	DISTCRIS	0.435 *	-
Sometimes people need to bribe community leaders in order to get things done.	BRIBELDR	−0.226 *	-
Some households in this community are restricted from community services, such as bed net distribution.	RESTRSER	−0.223 *	-
Factor 7: Community attachment (average factor loading on refined CFA = 0.865)			
I feel attached to this community and its people.	ATTACH	0.813 *	0.814 *
I feel proud to be part of this community.	PROUD	0.936 *	0.935 *
Being a member of this community is part of who I am.	IDENTITY	0.868 *	0.869 *
People in this community accept me as a member of the community.	ACCEPT	0.842 *	0.841 *

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters; *Extraction*: Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; *Rotation*: Promax; * two-tailed $p \leq 0.05$; † Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings.

Table A3. Factor Loadings from Preliminary CFA of Our Hypothesised CE Framework, Women.

Factors and Associated Items	Item	Initial Prelim. CFA (n_W = 726)	Refined † Prelim. CFA (n_W = 726)
Factor 1: Agency (average factor loading on refined CFA = 0.723)			
People in this community have the capacity to make positive changes by coming together.	COLLEFF	0.812 *	0.807 *
I have the ability to contribute to this community's development.	SEDEV	0.703 *	0.703 *
People in this community should work together to develop the community.	SHOULDEV	0.692 *	0.689 *
I have the capacity to achieve my future aims.	SELFEFF	0.691 *	0.692 *
This community needs assistance from others outside the community in order to make positive changes.	EXOASSIS	0.231 *	-
Factor 2: Common value (average factor loading on refined CFA = 0.811)			
Most people in this community have common values, for example, they value hard work.	COMMVALU	0.763 *	0.762 *
People in this community share the same ideas on how village matters should be managed.	COMMGMT	0.874 *	0.878 *
Most people in this community have similar hopes about the future development of the community.	SIMHOPES	0.794 *	0.792 *
Factor 3: Social response (average factor loading on refined CFA = 0.623)			
If the people of this community see crime-like activities, they will do something about it.	INTERCRI	0.474 *	0.475 *
When there is a problem in this community, people come together to discuss how it should be solved.	COMPRSLV	0.888 *	0.888 *
If someone in this community had a death in their family, the community will come together to support them while they mourn.	SUPMOURN	0.732 *	0.740 *
The people of this community will contribute their own money or labour for community development.	CONTRDEV	0.699 *	0.700 *
If there is a big dispute between two persons, other people from the community will help in solving the problem.	SLVDISPU	0.681 *	0.679 *
If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.	HLPCRPDZ	0.659 *	0.659 *
I feel happy for my neighbour if they have a good harvest.	HAPPYNEI	0.523 *	0.526 *
Most people in this community have similar beliefs about what is right and what is wrong.	SIMBLIEF	0.331 *	0.316 *
Differences between people, such as the amount of land they own, often causes problems in this community.	DIFPROBS	0.009	-
Factor 4: Social order (dropped from factor model given only 1 item with salient factor loading)			
People in this community live in harmony with each other most of the time.	HARMONY	0.666 *	-
In this community, you have to be careful, otherwise your neighbours may cheat you.	CHEATS	0.053	-
In this community, conflicts like stealing and fighting often occur.	CRIMECON	−0.040	-
When I am at home alone, I feel safe from threats of crime.	SAFEATHO	0.242 *	-

Table A3. *Cont.*

Factors and Associated Items	Item	Initial Prelim. CFA (n_W = 726)	Refined † Prelim. CFA (n_W = 726)
Factor 5: Social capital (average factor loading on refined CFA = 0.661)			
People in this community can be trusted.	COMTRUST	0.644 *	0.636 *
I typically accept advice from others in this community.	ADVICE	0.601 *	0.605 *
People in the community share new knowledge with their neighbour if they learn something new.	SHAREKNO	0.630 *	0.628 *
This is a close-knit community (i.e., people in this community have close personal relationships with each other).	CLOSE	0.749 *	0.744 *
In this community, people prioritise their own family's welfare over community development.	OWNWELF	0.335 *	0.342 *
If someone in this community loses a cow or goat, a neighbour will help look for it.	LOSTCOW	0.580 *	0.585 *
If you suddenly need some money, you can borrow from a person or group in your community.	BORMONEY	0.636 *	0.635 *
If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	NEICAREG	0.602 *	0.602 *
There are people in this community who show strong leadership.	UNOFLDRS	0.726 *	0.723 *
The community-based associations, such as the Edir, in this community are very active.	COMACTCG	0.690 *	0.692 *
The leaders of community-based associations, like Edir leaders, respond to this community's concerns.	ACTLDR1	0.770 *	0.770 *
Formal administrative leaders, like the kebele manager, provide support to this community.	ACTLDR2	0.660 *	0.656 *
This community's leaders can be trusted.	TRUSTLDR	0.713 *	0.702 *
People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.	CHOCGLDR	0.741 *	0.742 *
In this community, I have friends with whom I can share my problems.	HAVEFRND	0.756 *	0.759 *
My neighbours sometimes come to me to share their problems and get help.	COME4HLP	0.768 *	0.772 *
Most people in this community participate in community associations.	COPARTCG	0.624 *	0.624 *
I attend meetings of a community-based association, such as the Edir.	ACTCBGP	0.584 *	0.587 *
I participate in activities held by any community-based associations, such as the Edir.	PARTCBGP	0.579 *	0.580 *
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	ACTEXOGP	0.749 *	0.756 *
I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.	PAREXOGP	0.729 *	0.733 *
Factor 6: Social equity (dropped from factor model given only 2 items with salient factor loadings)			
When community groups make decisions, they are pleasing and good for most of the households in this community.	COMMGDEC	0.650 *	-
Sometimes people need to bribe community leaders in order to get things done.	BRIBELDR	−0.227 *	-
During a crisis situation, such as a drought, government services are distributed equally by the community to all households in need.	DISTCRIS	0.509 *	-
Some households in this community are restricted from community services, such as bed net distribution.	RESTRSER	−0.105 *	-
Factor 7: Community attachment (average factor loading on refined CFA = 0.884)			
I feel attached to this community and its people.	ATTACH	0.889 *	0.888 *
Being a member of this community is part of who I am.	IDENTITY	0.882 *	0.882 *
People in this community accept me as a member of the community.	ACCEPT	0.875 *	0.875 *
I feel proud to be part of this community.	PROUD	0.889 *	0.890 *

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters; *Extraction*: Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; *Rotation*: Promax * two tailed $p \leq 0.05$; † Refined CFA reflects post hoc model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings.

Table A4. Comparison of Model Fit Statistics from Preliminary CFA of Our Hypothesised CE Framework and CFA of EFA-Derived CE Factor Solutions.

Fit Statistic	Preliminary CFA of Hypothesised CE Framework				CFA of EFA-Derived Factor Solutions			
	Men		Women		Men		Women	
	Initial CFA	Refined CFA *	Initial CFA	Refined CFA *	Initial CFA	Refined CFA *	Initial CFA	Refined CFA *
χ^2	2312.227	1991.259	2810.783	2488.535	721.300	618.517	858.865	755.374
df	1106	764	1154	730	504	413	607	480
χ^2:df	2.091	2.606	2.436	3.409	1.431	1.498	1.415	1.574
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
RMSEA (90%	0.031	0.038	0.044	0.058	0.028	0.030	0.034	0.040
CI)	(0.030–0.033)	(0.036–0.040)	(0.042–0.047)	(0.055–0.060)	(0.023–0.032)	(0.025–0.035)	(0.029–0.039)	(0.034–0.045)
CFI	0.911	0.911	0.897	0.895	0.970	0.971	0.964	0.962
TLI	0.905	0.904	0.891	0.888	0.966	0.968	0.960	0.958
WRMR	1.962	2.089	2.124	2.344	1.047	1.005	1.137	1.168

Notes: WLSMV estimation with sandwich estimator to adjust for non-independence of observations with 50 *kebele* clusters; * Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings; **Abbreviations:** df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; WRMR = Weighted Root Mean Square Residual.

Table A5. Model Fit Statistics for CFA of Refined Single-Group and Parsimonious (Overlapping Items) Models.

Fit Statistic	Refined, Single-Group CFA Models *		Fit of Male, Female Data to Parsimonious Model	
	Men	Women	Men	Women
	Saturated Model, All Items	Saturated Model, All Items	Overlapping Items Only	Overlapping Items Only
χ^2	618.517	755.374	536.910	503.073
df	413	480	309	309
χ^2:df	1.498	1.574	1.738	1.628
p-value	<0.001	<0.001	<0.001	<0.001
RMSEA (90% CI)	0.030 (0.025–0.035)	0.040 (0.034–0.045)	0.037 (0.031–0.042)	0.042 (0.035–0.048)
CFI	0.971	0.962	0.965	0.967
TLI	0.968	0.958	0.961	0.963
WRMR	1.005	1.168	1.148	1.070

Notes: WLSMV estimation with sandwich estimator to adjust for non-independence of observations with 50 *kebele* clusters; * Refined CFA models reflect *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.32) or non-significant (two-tailed $p > 0.05$) factor loadings; Abbreviations: df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; WRMR = Weighted Root Mean Square Residual.

Table A6. Competing MIMIC Models: Fit Statistics, Unstandardised B Estimates, Standard Errors, and Standardised β Estimates.

Model	*n*	χ^2 (df)	Δ χ^2 (df)	RMSEA (90% CI)	CFI	TFI	B	S.E.	β
MIMIC models with gender covariate only									
1. Baseline MIMIC model—i.e., CE measurement model with ALL latent factors regressed on gender, no direct effects between gender & item indicators	910	731 (330)	N/A	0.037 (0.033–0.040)	0.966	0.961	-	-	-
2. Model 1 (i.e., ALL latent factors regressed on gender) + direct path between HARMONY item indicator and gender	910	717 (329)	31 *	0.036 (0.032–0.040)	0.967	0.962	−0.453 *	0.081	−0.444 *
MIMIC models with gender AND household leadership status covariates									
3. More saturated baseline MIMIC model with ALL latent factors regressed on gender AND leadership, no direct effects between covariates & items	907	746 (351)	N/A	0.035 (0.032–0.039)	0.965	0.960	-	-	-
4. Model 3 (ALL latent factors regressed on gender and leadership covariates) + direct path between HARMONY item indicator and gender	907	731 (350)	32 *	0.035 (0.031–0.038)	0.966	0.961	−0.457 *	0.081	−0.448 *
5. Model 4 + direct path between ACTEXOGP item indicator and leadership covariate	907	727 (349)	15 *	0.035 (0.031–0.038)	0.967	0.961	0.321 *	0.083	0.299 *
6. More saturated baseline MIMIC model with ONLY latent factors significant on Model 3 regressed on gender and leadership	907	730 (355)	N/A	0.034 (0.031–0.038)	0.967	0.962	-	-	-

Notes: *Matrix*: Polychoric correlations; *Estimation method*: WLSMV with sandwich estimator to adjust for non-independence of observations within 50 *kebele* clusters; Δ χ^2 assessed via DIFFTEST; B: unstandardised estimate; S.E. = standard error; β: standardised estimate; * $p \leq 0.001$, though χ^2 statistics are sensitive to sample size, therefore, DIFFTEST statistics are likely to be significant with large samples; Baseline MIMIC model regressed latent variables on gender and household leadership status; additional direct effect paths between gender and leadership covariates incorporated in subsequent models via step-wise forward selection based on the magnitude of the item indicator's modification index; Model 4 reflects final MIMIC model (i.e., model identified with the best fit and most substantively justified factor structure.

References

1. Mackie, G.; Moneti, F.; Shakya, H.; Denny, E. *What Are Social Norms? How Are They Measured?* UNICEF/University of California, San Diego, Center on Global Justice: San Diego, CA, USA, 2015.
2. Paluck, E.; Shepard, H. The Salience of Social Referents: A Field Experiment on Collective Norms and Harassment Behavior in a School Social Network. *J. Pers. Soc. Psychol.* **2012**, *103*, 899–915. [CrossRef] [PubMed]
3. Zou, X.; Tam, K.; Morris, M.; Lee, S.; Lau, I.; Chiu, C. Culture as Common Sense: Perceived Consensus Versus Personal Beliefs as Mechanisms of Cultural Influence. *J. Pers. Soc. Psychol.* **2009**, *97*, 579–597. [CrossRef] [PubMed]
4. Bandura, A. Health promotion by social cognitive means. *Health Educ. Behav.* **2004**, *31*, 143–164. [CrossRef] [PubMed]
5. Fuller, J.A.; Eisenberg, J.N. Herd Protection from Drinking Water, Sanitation, and Hygiene Interventions. *Am. J. Trop. Med. Hyg.* **2016**, *95*, 1201–1210. [CrossRef] [PubMed]
6. Oswald, W.E.; Stewart, A.E.; Kramer, M.R.; Endeshaw, T.; Zerihun, M.; Melak, B.; Sata, E.; Gessese, D.; Teferi, T.; Tadesse, Z.; et al. Active trachoma and community use of sanitation, Ethiopia. *Bull. World Health Organ.* **2017**, *95*, 250–260. [CrossRef] [PubMed]
7. Garn, J.V.; Boisson, S.; Willis, R.; Bakhtiari, A.; al-Khatib, T.; Amer, K.; Batcho, W.; Courtright, P.; Dejene, M.; Goepogui, A.; et al. Sanitation and water supply coverage thresholds associated with active trachoma: Modeling cross-sectional data from 13 countries. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006110. [CrossRef] [PubMed]
8. Luby, S.P.; Rahman, M.; Arnold, B.F.; Unicomb, L.; Ashraf, S.; Winch, P.J.; Stewart, C.P.; Begum, F.; Hussain, F.; Benjamin-Chung, J.; et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: A cluster randomised controlled trial. *Lancet Glob. Health* **2018**, *6*, e302–e315. [CrossRef]
9. Null, C.; Stewart, C.P.; Pickering, A.J.; Dentz, H.N.; Arnold, B.F.; Arnold, C.D.; Benjamin-Chung, J.; Clasen, T.; Dewey, K.G.; Fernald, L.C.H.; et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: A cluster-randomised controlled trial. *Lancet Glob. Health* **2018**, *6*, e316–e329. [CrossRef]
10. Clasen, T.; Boisson, S.; Routray, P.; Torondel, B.; Bell, M.; Cumming, O.; Ensink, J.; Freeman, M.; Jenkins, M.; Odagiri, M.; et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised trial. *Lancet Glob. Health* **2014**, *2*, e645–e653. [CrossRef]
11. Delea, M.G.; Solomon, H.; Solomon, A.W.; Freeman, M.C. Interventions to maximize facial cleanliness and achieve environmental improvement for trachoma elimination: A review of the grey literature. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0006178. [CrossRef] [PubMed]
12. Bandura, A. Self-efficacy mechanism in human agency. *Am. Psychol.* **1982**, *37*, 122–147. [CrossRef]
13. Bandura, A. Exercise of Human Agency through Collective Efficacy. *Curr. Dir. Psychol. Sci.* **2000**, *9*, 75–78. [CrossRef]
14. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*; Cambridge University Press: New York, NY, USA, 2015.
15. Bicchieri, C. *The Grammar of Society: The Nature and Dynamics of Social Norms*; Cambridge University Press: New York, NY, USA, 2006.
16. Bray, S.R. Collective Efficacy, Group Goals, and Group Performance of a Muscular Endurance Task. *Small Group Res.* **2004**, *35*, 230–238. [CrossRef]
17. Durham, C.C.; Knight, D.; Locke, E.A. Effects of Leader Role, Team-Set Goal Difficulty, Efficacy, and Tactics on Team Effectiveness. *Organ. Behav. Hum. Decis. Process.* **1997**, *72*, 203–231. [CrossRef]
18. Hodges, L.; Carron, A.V. Collective efficacy and group performance. *Int. J. Sport Psychol.* **1992**, *23*, 48–59.
19. Sampson, R.J.; Raudenbush, S.W.; Earls, F. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* **1997**, *277*, 918–924. [CrossRef] [PubMed]
20. Hipp, J.R. Collective efficacy: How is it conceptualized, how is it measured, and does it really matter for understanding perceived neighborhood crime and disorder? *J. Crim. Justice* **2016**, *46*, 32–44. [CrossRef] [PubMed]

21. Mason, K.O. The Status of Women: Conceptual and Methodological Issues in Demographic Studies. *Sociol. Forum* **1986**, *1*, 284–300. [CrossRef]
22. Mabsout, R.; van Staveren, I. Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia. *World Dev.* **2010**, *38*, 783–796. [CrossRef]
23. Meinzen-Dick, R.; Pandolfelli, L.; Dohrn, S.; Athens, J. *Gender and Collective Action: A Conceptual Framework for Analysis*; CAPRi Working Papers 64; International Food Policy Research Institute (IFPRI): Washington, DC, USA, 2006.
24. Agarwal, B. Conceptualising environmental collective action: Why gender matters. *Camb. J. Econ.* **2000**, *24*, 283–310. [CrossRef]
25. SDSN. Report of the SDSN Thematic Group on Challenges of Social Inclusion: Gender, Inequality, and Human Rights. In *Achieving Gender Equality, Social Inclusion, and Human Rights for All: Challenges and Priorities for the Sustainable Development Agenda*; SDSN: New York, NY, USA, 2013.
26. Folbre, N. *Who Pays for the Kids?: Gender and The Structure of Constraint*; Routledge: London, UK; New York, NY, USA, 1994.
27. Ibrahim, S.; Alkire, S. Agency and Empowerment: A Proposal for Internationally Comparable Indicators. *Oxf. Dev. Stud.* **2007**, *35*, 379–403. [CrossRef]
28. Sen, A. *Development as Freedom*; OUP: Oxford, UK, 2001.
29. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef] [PubMed]
30. Bisung, E.; Elliott, S.J.; Schuster-Wallace, C.J.; Karanja, D.M.; Bernard, A. Social capital, collective action and access to water in rural Kenya. *Soc. Sci. Med.* **2014**, *119*, 147–154. [CrossRef] [PubMed]
31. Dean, A.J.; Fielding, K.S.; Lindsay, J.; Newton, F.J.; Ross, H. How social capital influences community support for alternative water sources. *Sustain. Cities Soc.* **2016**, *27*, 457–466. [CrossRef]
32. Person, M.; Delea, M.; Garn, J.; Alexander, K.; Abaire, B.; Freeman, M. Assessing the influence of social capital on water point sustainability in rural Ethiopia. *J. Water Sanit. Hyg. Dev.* **2017**, *7*, 611–622. [CrossRef]
33. Ajzen, I.; Fishbein, M. *Understanding Attitudes and Predicting Social Behavior*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1980.
34. Bandura, A. Social cognitive theory: An agentic perspective. *Annu. Rev. Psychol.* **2001**, *52*, 1–26. [CrossRef] [PubMed]
35. Carroll, J.M.; Rosson, M.B.; Zhou, J. Collective efficacy as a measure of community. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, Oregon, USA, 2–7 April 2005; pp. 1–10.
36. Goddard, R.D.; Hoy, W.K.; Hoy, A.W. Collective Efficacy Beliefs:Theoretical Developments, Empirical Evidence, and Future Directions. *Educ. Res.* **2004**, *33*, 3–13. [CrossRef]
37. Spector, P. *Summated Rating Scale Construction*; Sage Publications, Inc.: Newbury Park, CA, USA, 1992. [CrossRef]
38. Creswell, J.W.; Clark, V.L.P. *Designing and Conducting Mixed Methods Research*; SAGE Publications: Thousand Oaks, CA, USA, 2007.
39. Pett, M.A.; Lackey, N.R.; Sullivan, J.J. *Making Sense of Factor Analysis the Use of Factor Analysis for Instrument Development in Health Care Research*; SAGE: Thousand Oaks, CA, USA; London, UK, 2003.
40. Brown, T.A. *Confirmatory Factor Analysis for Applied Research*; Guilford Publications: New York, NY, USA, 2006.
41. Yong, A.; Pearce, S. A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutor. Quant. Methods Psychol.* **2013**, *9*, 79–94. [CrossRef]
42. Clark, L.A.; Watson, D. Constructing validity: Basic issues in objective scale development. *Psychol. Assess.* **1995**, *7*, 309. [CrossRef]
43. Hooper, D.; Coughlan, J.; Mullen, M. Structural equation modelling: Guidelines for determining model fit. *Electron. J. Bus. Res. Methods* **2008**, *6*, 53–60.
44. Hancock, G.R.; Mueller, R.O. *The Reviewer's Guide to Quantitative Methods in the Social Sciences*; Routledge: New York, NY, USA, 2010.
45. Bandura, A. *Social Learning Theory*; General Learning Press: New York, NY, USA, 1977.
46. Beatty, P.C.; Willis, G.B. Research Synthesis: The Practice of Cognitive Interviewing. *Public Opin. Q.* **2007**, *71*, 287–311. [CrossRef]

47. Castillo-Díaz, M.; Padilla, J.-L. How Cognitive Interviewing can Provide Validity Evidence of the Response Processes to Scale Items. *Soc. Indic. Res.* **2013**, *114*, 963–975. [CrossRef]
48. Henson, R.K.; Roberts, J.K. Use of Exploratory Factor Analysis in Published Research:Common Errors and Some Comment on Improved Practice. *Educ. Psychol. Meas.* **2006**, *66*, 393–416. [CrossRef]
49. Muthen, B.; du Toit, S.; Spisic, D. Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Psychometrika* **1997**, *75*, 40–45.
50. Baglin, J. Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR. *Pract. Assess. Res. Eval.* **2014**, *19*, 1–14.
51. Yang, Y.; Xia, Y. On the number of factors to retain in exploratory factor analysis for ordered categorical data. *Behav. Res. Methods* **2015**, *47*, 756–772. [CrossRef] [PubMed]
52. Hu, L.T.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. A Multidiscip. J.* **1999**, *6*, 1–55. [CrossRef]
53. Muthen, L.K.; Muthen, B. *Mplus: Statistical Analysis with Latent Variables: User's Guide*; Muthén & Muthén: Los Angeles, CA, USA, 2015.
54. Tabachnick, B.G.; Fidell, L.S. *Using Multivariate Statistics*; Pearson/Allyn & Bacon: Boston, MA, USA, 2007.
55. Kaiser, H.F. The Application of Electronic Computers to Factor Analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [CrossRef]
56. Costello, A.B.; Osborne, J.W. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* **2005**, *10*, 1–9.
57. Thurstone, L.L. *Multiple-Factor Analysis*; University of Chicago Press: Chicago, IL, USA; London, UK, 1947.
58. Schreiber, J.B.; Nora, A.; Stage, F.K.; Barlow, E.A.; King, J. Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *J. Educ. Res.* **2006**, *99*, 323–338. [CrossRef]
59. Mellenbergh, G.J. Item bias and item response theory. *Int. J. Educ. Res.* **1989**, *13*, 127–143. [CrossRef]
60. Woods, C.M.; Oltmanns, T.F.; Turkheimer, E. Illustration of MIMIC-Model DIF Testing with the Schedule for Nonadaptive and Adaptive Personality. *J. Psychopathol. Behav. Assess.* **2009**, *31*, 320–330. [CrossRef] [PubMed]
61. Holland, P.W.; Wainer, H. *Differential Item Functioning*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1993.
62. Bentler, P.M.; Bonett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **1980**, *88*, 588–606. [CrossRef]
63. DiStefano, C.; Zhu, M.; Minddrila, D. Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Pract. Assess. Res. Eval.* **2009**, *14*, 1–11.
64. Cameron, A.; Miller, D. A Practitioner's Guide to Cluster-Robust Inference. *J. Hum. Resour.* **2015**, *50*, 317–372. [CrossRef]
65. Camilli, G.; Shepard, L.A. *Methods for Identifying Biased Test Items*; Sage: Thousand Oaks, CA, USA, 1994.
66. Rogers, E. *Diffusion of Innovations*, 3rd ed.; Macmillan Publishing Company, Inc.: New York, NY, USA, 1983.
67. Goetz, A.M. *Getting Institutions Right for Women in Development*; Zed Books: London, UK, 1997.
68. Ferguson, L. External Validity, Generalizability, and Knowledge Utilization. *J. Nurs. Scholarsh.* **2004**, *36*, 16–22. [CrossRef] [PubMed]
69. Wutich, A.; Ragsdale, K. Water insecurity and emotional distress: Coping with supply, access, and seasonal variability of water in a Bolivian squatter settlement. *Soc. Sci. Med.* **2008**, *67*, 2116–2125. [CrossRef] [PubMed]
70. Hadley, C.; Wutich, A. Experience-based measures of food and water security: Biocultural approaches to grounded measures of insecurity. *Hum. Organ.* **2009**, *68*, 451–460. [CrossRef]
71. Stevenson, E.G.J.; Greene, L.E.; Maes, K.C.; Ambelu, A.; Tesfaye, Y.A.; Rheingans, R.; Hadley, C. Water insecurity in 3 dimensions: An anthropological perspective on water and women's psychosocial distress in Ethiopia. *Soc. Sci. Med.* **2012**, *75*, 392–400. [CrossRef] [PubMed]
72. Hathi, P.; Spears, D.; Coffey, D. Can collective action strategies motivate behaviour change to reduce open defecation in rural India? *Waterlines* **2016**, *35*, 118–135. [CrossRef]

International Journal of
Environmental Research and Public Health

MDPI

Supplemental Material

Collective efficacy: Development and validation of a measurement scale for use in public health and development programmes

Maryann G. Delea [1,2,*], Gloria D. Sclar [2], Mulat Woreta [3], Regine Haardörfer [4], Corey L. Nagel [5], Bethany A. Caruso [2], Robert Dreibelbis [1], Abebe G. Gobezayehu [3], Thomas F. Clasen [1,2],

ADDITIONAL METHODS

Additional sampling methodology related to the larger Andilaye trial

For the larger *Andilaye* trial, we employed a structured sampling strategy to randomly select eligible sub-district (*kebele*) clusters and study households. The primary sampling unit for the *Andilaye* trial was the *kebele*; specifically, any rural or peri-urban *kebele* within three districts (*woredas* – Bahir Dar Zuria, Fogera, and Farta) of Amhara, Ethiopia that is accessible throughout the course of the year. The ultimate sampling unit for this study was the household; specifically, any household residing in a targeted, sentinel village (*gott)* within a randomly selected study *kebele*. We randomly selected households for inclusion in the study by using a random number generator to identity approximately 30 households from a sampling frame of all households in the selected *gott* that had at least one child between the ages of 1-9 years. We conducted our CE sub-study amongst the households enrolled in the *Andilaye* trail.

When the primary target respondents for our survey were not present, one of the following adult household members was engaged (listed preferentially): any female household member who serves as a caregiver, any male household member serves as a caregiver, any adult household member.

Additional EFA details

Priority was placed on selecting solutions with model-estimated correlation matrices close in value to the original sample correlation matrices, and residual correlation matrices with values close to zero. Such results indicated that factor solutions with the respective number of retained factors were sufficient for explaining the inter-correlations amongst the observed variables.

Criteria to identify models with simple structure

The following, adapted from Thurstone's criteria [1], were to identify models with simple structure and guide iterative re-analysis of CE measurement models subsequent to item reduction via EFA:

- Each item produces at least one zero loading on some factor.
- Each factor is measured by a set of items with high loadings on the factor, and at least as many zero loadings as there are factors.
- Each item has a high loading on one factor (i.e., primary loading), and a trivial or close to zero loading on all remaining factors.
- Each pair of factors should have a large proportion of zero loadings on both factors (if there are four or more factors total).
- Each pair of factors should have only a few complex variables.

We used established guidance [2] to define a zero loading, which we deemed as any factor loading between -0.100 and +0.100. Our definition of significant loading was informed by empirical evidence as well. Factor loadings with an absolute value greater than 0.320 were considered salient.

Items with factor loadings less than this threshold measured the latent factors poorly, and were eliminated in a step-wise manner [3]. It is worth noting that no broadly accepted guidelines exist for saliency of factor loadings, but pattern coefficients in the range of 0.300-0.400 are often interpreted by analysts as salient in applied research [4]. We defined complex variables as items with factor loadings of |0.300| on more than one factor [4].

Justification of factor extraction approach

Factor retention was not solely based on the Kaiser-Guttman rule (i.e., eigenvalue >1.0 [5]), but also considered heuristic descriptive guides (i.e., scree-plot), goodness-of-fit, and other substantive justification (e.g., results from cognitive interviews, theory and other evidence). The last factor extracted for the men's CE model had an eigenvalue of 1.118; the first factor not retained had an eigenvalue of 1.029. The eigenvalue for the last factor extracted for the women's CE model was 1.336; the first factor not retained had an eigenvalue of 1.068. While the first factor not retained for both men's and women's CE measurement models had values above 1.0 threshold, the retention of those factors was not warranted by strong substantive or statistical justification [4, 6, 7]. Including those factors merely because their eigenvalues were slightly greater than 1.0 would reflect the sole use of a mathematically-based descriptive guide for item retention. Such an approach would go against our pre-analysis plan, disregard heuristic and model fit criteria, and important empirical and theoretical considerations (e.g., results from cognitive interviews, pilot testing of the CE instrument and other prior theoretical and empirical evidence).

In addition, many methodologists have criticised and demonstrated that the Kaiser-Guttman rule can tend to result in overfactoring or underfactoring given sampling error may influence eigenvalues [2, 4, 8]. While identifying and retaining too few factors (i.e., underfactoring) may result in an oversimplified understanding of a construct, retaining too many factors (i.e., overfactoring) may lead to violation of parsimony, which is one primary goal of EFA [4]. Whether over- or underfactoring occurs, the factor solution that results may lead to unreliable factors and/or errors in interpretation [2, 8]. Given the more parsimonious (i.e., eight-factor) measurement models were supported by our knowledge of the existing theoretical and empirical literature base, and other non-mathematically-based criteria, we felt our factor extraction and retention decisions were sufficiently justified.

ADDITIONAL RESULTS

Univariate statistics: CE survey items

Our CE survey included 50 items for factoring (Appendix A). The top five items to which respondents most frequently selected "completely agree" aligned for men and women, though there were some differences with regard to the proportions of those responses between genders (Appendix SA). These items reflected those related to social solidarity or support for one's community members and a sense of pride about being a part of the community: "If someone in this community had a death in their family, the community will come together to support them while they mourn" (94% of men, 91% of women); "I feel happy for my neighbour if they have a good harvest" (96% of men, 92% of women); "I feel proud to be part of this community" (91% of men, 73% of women); "If someone in this community loses a cow or goat, a neighbour will help look for it" (89% of men, 90% of women); "People in this community get to choose the leaders of their own community-based associations, such as Edir leaders" (86% of men, 73% of women). The items to which respondents most frequently selected "completely disagree" also aligned between genders. These items reflected those related to social disorder and inequity: "Sometimes people need to bribe community leaders in order to get things done" (71% of men, 52% of women); "Some households in this community are restricted from community services, such as bed net distribution" (71% of men, 46% of women); "In this community, conflicts like stealing and fighting often occur" (29% of men, 47% of women); "In this community, you have to be careful, otherwise your neighbours will cheat you" (25% of men, 30% of women). In terms of normality of item response distributions, men had 27 items, and women had 15 items with skewness outside of the suggested range (Appendix SA). The WLSMV estimator we employed for

our factor analyses makes no distribution assumptions for observed variables, and only assumes a normal latent distribution underlying each observed categorical variable [9], so no action was taken to address any non-normal item distributions [4].

Interpretation of factor loadings

It is acceptable and appropriate to consider factor loadings that vary in magnitude across the various items tapping to a latent factor, as the magnitude of an item's factor loading reflects the proximity of the relationship between the item and the factor to which it taps [10]. Factor loadings may therefore vary in magnitude across the items tapping to a factor based on the proximity of those relationships [10]. Items that are conceptually less influential (i.e., less proximal) to a given latent factor could demonstrate a lower factor loading without necessarily signaling poor quality of the latent factor and poor validity of the measurement model [11]. An item indicator that almost perfectly reflects a given latent factor should be very highly correlated with it (e.g., as represented by a factor loading in the range of 0.800-0.900). However, other items tapping to the latent factor that are conceptually less important or proximal to the factor can, and theoretically should demonstrate lower factor loadings [10, 12].

Additional preliminary CFA results

We moved forward with *post hoc* model refinements of preliminary CFA models to eliminate non-salient and non-significant factor loadings as well as any factors with insufficient component saturation. For the men's model, this resulted in the elimination of nine items. One item (HAVEFRND) was eliminated because it had less than minimal variance (i.e., a response category with zero observations). Five items were eliminated for non-salient factor loadings (OWNWELF=0.140, SAFEATHO=0.151, RESTRSER=-0.223, BRIBELDR=-0.226, DIFPROBS=-0.267), and one item was eliminated because it had both a non-salient and non-significant loading on its designated factor (EXOASSIS=0.011, p=0.793). After eliminating items that were non-significant and non-salient, one factor (social equity) remained with only two items, which we did not deem sufficient for component saturation. We therefore eliminated that factor and the remaining two items which otherwise demonstrated salient and significant loadings (COMMGDEC=0.717, DISTCRIS=0.428). The standardised estimates of the remaining factor loadings from this model were acceptable (Appendix B), and all remaining factors co-varied significantly. The refined preliminary CFA model of the hypothesized CE framework demonstrated adequate absolute model fit (χ^2:df ratio = 2.606, RSMEA=0.038 [0.036 – 0.040]), but still poor incremental fit (CFI=0.911, TLI=0.904). These results suggest that our hypothesised CE framework represented a plausible structure of the mechanisms through which the CE process operates amongst men in the Ethiopian context. However, poor incremental fit statistics suggested that this may not have been the best fitting model framework.

For the women's model, we eliminated a total of ten items. Four items were eliminated as a result of non-salient factor loadings on the designated factor (RESTRSER=-0.105, BRIBELDR=-0.227, EXOASSIS=0.231, SAFEATHO=0.242). Three items were eliminated due to non-salient and non-significant factor loadings on the designated factor (DIFPROBS=0.009, p=0.868; CRIMECON=-0.040, p=0.543; CHEATS=0.053, p=0.324). Two factors and their three items were eliminated because the factors demonstrated insufficient component saturation (the factor representing social order with its HARMONY item, and the factor representing social equity with its COMMGDEC and DISTCRIS items). The refined preliminary CFA model only marginally reflected the actual hypothesised framework, as two factor loadings were non-salient (social order and social equity). The standardised estimates of the remaining factor loadings from the resulting model were acceptable (Appendix C). Both absolute and incremental fit statistics indicated poor fitness of the resulting women's factor model (χ^2:df ratio=3.409, RSMEA=0.058 [0.055–0.060]; CFI=0.895, TLI=0.888). This means that the data failed to validate the hypothesised CE framework for women respondents, indicating the framework did not reflect the mechanisms through which the CE process operates for women in the rural Ethiopian context.

There was considerable overlap in the items eliminated from both men and women refined, preliminary CFA models. All but one (OWNWELF) of the items eliminated from the men's model were also eliminated from the women's model, and five of the ten items eliminated from the women's model were also eliminated from the men's model (SAFEATHO, RESTRSER, BRIBELDR, DIFPROBS, and EXOASSIS).

Additional EFA and CFA results

Complete EFA results reflect coefficients from both rotated (Promax) pattern and structure matrices along with initial and refined CFA results. While not all factor loadings demonstrated in Tables 3 and 4 are in the range of excellent to very good – though they are still in the acceptable range – we hypothesise that some of those items are conceptually more distal (i.e., marginally less important) to the measurement of the latent factor. We present further details regarding both men and women EFA-derived measurement models in subsequent sub-sections.

Additional details regarding the men's EFA and CFA results

During the EFA analyses, we eliminated three items (HAVEFRND, HAPPYNEI, PROUD) due to less than minimal variance (i.e., no observations in one or more item response category) that prevented the EFA from being processed in MPlus. We also eliminated twelve items, in a step-wise manner: ten items were eliminated because they had no salient loadings on any factor (BRIBELDR, EXOASSIS, SAFEATHO, COPARTCG, CHEATS, INTERCRI, COMMGDEC, CONTRDEV, SUPMOURN, LOSTCOW); one item (PAREXOGP) was eliminated due to evidence of extreme multicollinearity with another related item that loaded to the factor; and one item (CRIMECON) was eliminated because although its pattern coefficient was salient, its structure coefficient was not. This resulted in a 35-item men's CE measurement model (with two complex variables) that tapped to seven factors of CE: social response, social networks and personal agency, social attachment, common vision, community leadership, associational participation, and community organisation.

Factor one, labelled "social response" corresponded to the informal social control domain, though it also tapped to certain aspects of cognitive social capital (e.g., trust in community members, reciprocity of knowledge) that may influence social response. The factor contained nine items that tap to various facets of perceptions regarding the community's propensity to address community- and sub-community level issues, including social disorder (e.g., harmony, problem solving, conflict-resolution, common moral principles and codes of behaviour), support in times of crisis, and tolerance. The concepts reflected in this factor align closely with our hypothesised operational definition of social control, described as an absence of general conflict and threats to the existing order, effective informal social control, tolerance, and intergroup cooperation (Table 1).

Factor two, labelled "social networks & personal agency" corresponded to the cognitive social capital domain, though it also tapped to structural social capital, as it reflects the strength and responsiveness of one's social structures. The factor comprised of five items that relate to issues surrounding supporting networks and individuals cooperating to support one another for either mutual or one-sided gain. Two items related to self-efficacy loaded to this factor. This suggests that for men, one's perspectives regarding personal agency (i.e., individual behavioural control) is linked to perceived expectations that help will be given to or received from others, when needed [13].

Factors three and four corresponded to the social cohesion domain. Factor three, labelled "social attachment" included five items that tap to concepts related to place identity, community acceptance and attachment, and collective agency. Factor four, labelled "common vision" was comprised of six items that reflect shared norms (perceptions of normative expectations regarding contributions to community development) and culture (common values, hopes for the future, ideas about how the community should be managed), social equity (equal distribution of goods in times of crisis), and perceptions regarding community-level agency.

Factors five, six, and seven pertained to the structural social capital domain. Factor five, labelled "community leadership" reflected four items tapping to various aspects of social trust, support, and strength of leadership of formal administrative leaders and both formal and informal community

leaders. Factor six, labelled "associational participation" corresponded to the respondent's personal involvement in established community structures – both exogenously and endogenously organised. The three constituent items reflect both membership (as indicated by meeting attendance) and participation in associational activities. Factor seven, labelled "community organisation" corresponded to various aspects of community organisation, including the activity level of endogenously organised community associations and leaders thereof, community-selected representation, prioritisation of community development, and social justice and equity.

During CFA, we moved forward with *post hoc* model refinements to eliminate non-salient and non-significant factor loadings as well as any factors with insufficient component saturation. Prior to CFA, we eliminated one item (ADVICE) due to less than minimal variance. Subsequent *post hoc* model refinements resulted in the elimination of five additional items. Two items were eliminated for non-significant and non-salient factor loadings (SHOULDEV=0.075, p=0.513; COLLEFF=0.071, p=0.350), and three items were eliminated for non-salient factor loadings (OWNWELF=0.155, RESTRSER=-0.260, and DIFPROBS=-0.278). The standardised estimates of factor loadings from this model were acceptable (Table 3).

Additional details regarding the women's EFA and CFA results

During the EFA analyses, we eliminated one item (HAPPYNEI) due to less than minimal variance. We eliminated twelve additional items in a step-wise manner: six items were eliminated because they had no salient loadings on any factor (RESTRSER, BRIBELDR, COPARTCG, EXOASSIS, COMMGDEC, SHOULDEV); four items were complex variables that cross-loaded on more than one factor without sufficient substantive justification (SAFEATHO, SUPMOURN, SHAREKNO, TRUSTLDR); one item (PAREXOGP) was eliminated due to evidence of extreme multicollinearity with another item that loaded to the factor; and one item (ONWELF) was eliminated because although its pattern coefficient was salient, its structure coefficient was not salient on the factor of interest. This item reduction process resulted in a 37-item women's CE measurement model that tapped to seven factors of CE: social networks & reciprocity, social disorder, social attachment & personal agency, social response, associational participation, common vision, and community organisation & leadership.

Factor one, labelled "social networks & reciprocity" corresponded to the cognitive social capital domain, though it also tapped to certain aspects of structural social capital, as it reflected perceptions related to collections of individuals that promote and protect mutual or personal interests. The factor contained eight items that indicate various aspects of reciprocity demonstrated through social networks, the strength of personal relationships, and the community's propensity to contribute to community development.

Factors two and four corresponded to the informal social control domain, though factor four also tapped to certain aspects of cognitive social capital. Factor two, labelled "social disorder" contained three items that reflect the level of disorder in the community, including conflicts such as stealing, fighting, cheating, and problems caused by intolerance of differences amongst people. Factor four, labelled "social response" contained eight items that tap to various facets of perceptions regarding the community's propensity to address internal issues, including willingness to intervene when crime-like activities are observed, conflict-resolution, common moral principles and codes of behaviour, support in times of crisis, community trust, and strength of relationships.

Factors three and six corresponded to the social cohesion domain. Factor three, labelled "social attachment & personal agency" included six items that tap to concepts related to place identity, community acceptance and attachment, and personal agency. This suggests that, for women, one's sense of self-agency is linked to one's sense of belonging or social attachment. Factor six, labelled "common vision" is comprised of five items that reflect shared culture (common values, hopes for the future, ideas about how the community should be managed), social equity (equal distribution of goods in times of crisis), and perceptions regarding community-level agency.

Factors five and seven corresponded to the structural social capital domain. Factor five, labelled as "associational participation" related to the respondent's personal involvement in established

community structures – both exogenously and endogenously organised. The three constituent items reflect both membership (as indicated by meeting attendance) and participation in associational activities. Factor seven, labelled "community organisation & leadership" corresponded to various aspects of organisation within the community, including the activity level of endogenously organised community associations and leaders thereof, and community-selected representation.

We conducted CFA on the 37 items tapping to seven factors, as indicated by the EFA-derived women's CE factor solution. We moved forward with *post hoc* model refinements to eliminate non-salient and non-significant factor loadings as well as any factors with insufficient component saturation. This resulted in the elimination of five items and one factor. One item was eliminated for a non-significant and non-salient factor loading (CLOSE=0.167, p=0.075), and two items were eliminated for non-salient factor loadings (CHEATS=0.213, and SIMBLIEF=0.309). With the elimination of one non-saliently loading item to the social order factor, the factor itself failed to demonstrate sufficient component saturation, so the factor and its remaining two items (DIFPROBS=0.900, p=0.001 and CRIMECON=0.366, p=0.001) were eliminated from the women's measurement model. The standardised estimates of factor loadings for the resulting six-factor model were acceptable (Table 4). Modification Indices above 3.84 on the women's model were all relatively low, meaning localised strain was relatively low in all areas identified. No further modifications were made.

Additional details regarding comparison of men's and women's CE measurement models

The men's CE measurement model included one more factor (community leadership) than was indicated by the women's CE measurement model. Two of the three items that comprised the leadership factor in the men's model are included in the community organisation factor in the women's measurement model, as there was sufficient substantive justification for those items tapping to that factor.

Comparison of CFA results of our hypothesised CE framework vs. EFA-derived factor solutions

Fit statistics from the preliminary CFA of our hypothesised CE framework and the CFA of the EFA-derived factor solution suggest that slight revisions that were substantively justified resulted in valid CE measurement models for both men and women in the Ethiopian context (Appendix D).

Comparison of fit statistics for CFA of refined, single-group and parsimonious models

Given it is encouraged to consider numerous alternatives before settling on final measurement models [6], we performed a CFA on both men's and women's models that reflected the more parsimonious set of CE indicators (i.e., only those that were completely overlapping between refined and validated CFA models). We present model fit statistics for those models, and compare them to the refined, validated CFA models in Appendix E. These results indicate that both the more saturated and parsimonious models are valid CE measurement metrics. The gender-specific saturated models represent slightly better fitting models.

We present model fit statistics, unstandardised B, standard errors, and standardised β for competing MIMIC models in Appendix F.

ADDITIONAL DISCUSSION

Establishing this CE measurement scale in the early phases of the *Andilaye* trial allowed us to measure and assess collective efficacy at baseline, prior to the implementation of a community-level demand-side sanitation and hygiene intervention. We plan to employ this validated scale again at endline, and compare changes in CE measures between intervention and counterfactual communities over time (pre-, post-intervention). This will allow us to test our hypothesis that there is a bi-directional, causal association between CE and intervention effectiveness.

Further discussion of gender-specific CE measurement models

There were slight differences between gender-specific CE measurement models (31-item, seven-factor solution vs. 33-item, six factor solution for men and women). Major differences between men's and women's CE measurement models involved: 1) the number of factors included in the measurement model, and 2) the manner in which individual-level behavioural control items (SELFEFF, SEDEV) correlated with factors related to social networks versus social attachment for men and women, respectively. The ordering of the CE scale factors also differed between men's and women's measurement models, and social networks & reciprocity emerging as the first factor in the women's model while social response emerging as the first factor in the men's model. These types of differences were expected, and are supported by empirical evidence that suggests women have a higher dependence on social networks and "the commons" than men [14].

The women's CE measurement model included several additional items that tapped to its social network factor that were not included in the men's measurement model. These items reflected additional concepts that reflected facets of reciprocity, communal contribution and collaboration, and solidarity. The women's measurement model also indicated that willingness to intervene in situations of delinquent behaviour was an important item related to social response, and perceptions regarding a sense of pride in being part of one's community was an important item related to social attachment. These items were not indicated in the men's measurement model, though at least in the case of the item that corresponded to pride, the exclusion of that indicator may have to do with less than minimal variance amongst the item responses, as one response category for each of the split-half samples had no observations. The men's measurement model included two items that tapped to its social response factor that were not included in the women's measurement model. These items reflected common understanding regarding right and wrong and information sharing.

The men's measurement model also indicated that perceptions regarding normative expectations about members of the community working together to develop the community was an important item related to the common vision factor. Men's and women's measurement models differed in the sense that the men's CE measurement model indicated that a seventh factor – community leadership – was important for measuring CE. Two of the items that were included in this factor – those indicating supportive formal leadership and strong informal leadership – were included elsewhere in the women's measurement model (community organisation, as supported by sufficient substantive justification). A third item related to perceived trust in the community's leaders was not included in the women's measurement model, but was indicated as an important component for the measurement of community leadership in the men's model.

While we did reveal the underlying CE factor structure for gender-specific models, we also determined that there was considerable overlap between men's and women's CE measurement models. We determined that a parsimonious model that reflected all factors and items in common between the two gender-specific models demonstrated good model fit, and may therefore be used to measure and compare CE between genders. That said, the use of gender-specific CE scales may allow interested researchers to assess the mechanisms through which CE operate, and monitor how measures related to these gender-specific mechanisms change over time, throughout the duration of a development programme or research study.

Significant differences in associational participation factor scores corroborate existing evidence that suggests women may participate less in endogenous and exogenous community structures. This findings indicates that working through formal community structures to enhance women's behavioural control perspectives, including self- and collective efficacy, may not be an appropriate approach. More appropriate approaches may include community-level or household-level intervention activities.

In terms of selecting a CE measurement metric for administration more broadly, it is necessary to determine the aim and objectives of the work at hand, and weigh the benefits of being able to compare CE scores across genders (refined parsimonious CE scale) against being able to assess the mechanisms through which CE operate (gender-specific, saturated CE scales). Our results indicated that CE perceptions differ between men and women, even amongst those living in the same household. Therefore, researchers and programme implementers using an adapted version of our

parsimonious CE scale should either consider obtaining data from men and women within the same household or obtaining CE data from a random selection of men and women within a given community.

Additional discussion regarding factor indeterminacy

While the refined and final validated factor structures championed by this study demonstrate good model fit, and are substantively justified, they reflect only one possible representation of the relationship amongst items in the men's, women's, and parsimonious CE measurement models. As with any EFA, our results were influenced by the structure of the data for the particular sample we ascertained. Other measurement models that fit the data and represent the conceptualisation of CE as well or better than our refined gender-specific and final parsimonious CE measurement models may exist [6]. Through the employment of a randomly selected split-half hold-out sample, we sought to assess the stability of our EFA-derived CE factor structures across an independent sample from the same population, as suggested by numerous methodologists [4, 7, 15].

Further discussion of analytical limitations

Mathematically-focused factor extraction methods have a tendency to under- or over-estimate the number of factors in a solution [2, 4, 8]. The results of scree tests are often ambiguous (e.g., no clear shift in the slope) and subject to interpretation [4]. As a result, we used a combination of mathematical (i.e., eigenvalue-based Kaiser-Guttman rule), heuristic (i.e., scree plot), statistical (i.e., model fit statistics), and substantive justification to guide factor extraction. That said, we were not able to perform more rigorous procedures (e.g., parallel analysis) to confirm that we extract the correct number of factors, as these analyses are not available for categorical data in Mplus [16].

Sufficient component saturation is needed (i.e., two or more items with salient factor loadings) to guarantee appropriate factor interpretation [7]. While some methodologist suggest that as few as two to three items provide sufficient component saturation [17], other more conservative guidelines suggest four or more items with factor loadings of 0.5 or higher, and an average factor loading of 0.700 across all items tapping a factor. All six factors in the final parsimonious CE measurement model had three to five items per factor, all loading ≥ 0.478, indicating sufficient component saturation. With the except of one factor (i.e., "social networks", average factor loading = 0.663), all factors demonstrated average factor loadings of 0.700, signaling that the items were good measures of the factors to which they tapped. All seven factors of the refined, validated men's CE measurement model and all six factors of the refined, validated women's measurement model included three or more items, all with factor loadings greater than 0.500. However, one factor on the women's model, and four factors on the men's model included three items only, which just satisfies moderate [17], but does not more conservative guidelines for component saturation. In addition, two items within the men's measurement model, and one item within the women measurement model reflect factor loadings falling within the salient but only "adequate" range (i.e., 0.400-0.440). More importantly, perhaps, two factors in the refined, validated men's measurement model, and one factor in the refined, validated women's measurement model demonstrated average factor loadings below the ideal 0.700 average (average factor loading on refined CFA: 0.668 and 0.634 on the men's model; 0.656 on the women's model). Interestingly, the factor on the women's model and one factor on the men's model with average factor loadings less than 0.700 represented the social response factor. This suggests that perhaps the items we included in our CE survey for this CE sub-construct may not have included one or more proximal indicators of social response in the Ethiopian context.

Given our EFA results reflect the structure of the sample we ascertained, and the role that sampling error and other systematic error may play in the estimation of factor analytic results, initial EFA findings should be interpreted with caution. These findings should be cross-validated through additional EFA or CFAs using independent datasets [4]. We employed a random split-half hold-out sample for measurement model validation, and the resulting findings were promising, especially our refined final parsimonious CE measurement scale. Still, these findings should undergo further validation with independent datasets, which is planned for another WASH study being evaluated by

members of our research group. Since our results indicated that only minimal component saturation was attained for some CE factors, and more proximal indicators may not have been included for social response and social network factors, additional formative work that further explores these issues is warranted.

SUPPLEMENTAL MATERIAL REFERENCES

1. Thurstone, L.L., *MUltiple-factor analysis*. 1947, Chicago; London: University of Chicago Press.
2. Gorsuch, R.L., *Factor analysis.* 2013.
3. Tabachnick, B.G. and L.S. Fidell, *Using multivariate statistics*. 2007, Boston: Pearson/Allyn & Bacon.
4. Brown, T.A., *Confirmatory Factor Analysis for Applied Research*. 2006: Guilford Publications.
5. Kaiser, H.F., *The Application of Electronic Computers to Factor Analysis.* Educational and Psychological Measurement, 1960. **20**(1): p. 141-151.
6. Hancock, G.R. and R.O. Mueller, *The reviewer's guide to quantitative methods in the social sciences*. 2010, New York: Routledge.
7. Henson, R.K. and J.K. Roberts, *Use of Exploratory Factor Analysis in Published Research:Common Errors and Some Comment on Improved Practice.* Educational and Psychological Measurement, 2006. **66**(3): p. 393-416.
8. Comrey, A.L. and H.B. Lee, *A first course in factor analysis*. Second Edition ed. 1992, Hillsdale, NJ: Lawrence Erlbaum.
9. Li, C.-H., *Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares.* Behavior Research Methods, 2016. **48**(3): p. 936-949.
10. Bollen, K.A., *Structural Equations with Latent Variables*. 1989: Wiley & Sons Inc.
11. Steinmetz, H., *Estimation and comparison of latent means across cultures*, in *Cross-Cultural Analysis : Methods and Applications*, E. Davidov, P. Schmidt, and J. Billiet, Editors. 2011, Taylor & Francis Group: New York, NY. p. 85-116.
12. Steinmetz, H., *Personal communication*. 2018.
13. Forrest, R. and A. Kearns, *Social Cohesion, Social Capital and the Neighbourhood.* Urban Studies, 2001. **38**(12): p. 2125-2143.
14. Agarwal, B., *Conceptualising environmental collective action: why gender matters.* Cambridge Journal of Economics, 2000. **24**(3): p. 283-310.
15. Baglin, J., *Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR.* Practical Assessment, Research & Evaluation, 2014. **19**(5).
16. Muthen, L.K. and B. Muthen, *Mplus : statistical analysis with latent variables : user's guide*. 2015.
17. Pedhazur, E.J. and F.N. Kerlinger, *Multiple regression in behavioral research: explanation and prediction*. 1982: Holt, Rinehart, and Winston.

International Journal of
Environmental Research and Public Health

MDPI

1

Table S1. Univariate descriptive statistics: Frequency of responses by split-halves and gender.

				EFA sub-sample						CFA sub-sample					
		N		N_1						N_2					
		Total sample		Aggregate		Women		Men		Aggregate		Women		Men	
CE item	Item response	N = 1,831		n_{E1} = 921		n_{EW1} = 366		n_{EM1} = 555		n_{E2} = 910		n_{EW2} = 360		n_{EM2} = 550	
HARMONY	***People in this community live in harmony with each other most of the time.***														
	Completely disagree	147	8.03%	66	7.17%	22	6.01%	44	7.93%	81	8.90%	26	7.22%	55	10.00%
	Mildly/partially disagree	76	4.15%	41	4.45%	13	3.55%	28	5.05%	35	3.85%	15	4.17%	20	3.64%
	Neither agree nor disagree	20	1.09%	6	0.65%	5	1.37%	1	0.18%	14	1.54%	10	2.78%	4	0.73%
	Mildly/partially agree	710	38.78%	379	41.15%	112	30.60%	267	48.11%	331	36.37%	82	22.78%	249	45.27%
	Completely agree	878	47.95%	429	46.58%	214	58.47%	215	38.74%	449	49.34%	227	63.06%	222	40.36%
CHEATS	***In this community, you have to be careful, otherwise your neighbours may cheat you.***														
	Completely disagree	491	26.82%	245	26.60%	109	29.78%	136	24.50%	246	27.03%	105	29.17%	141	25.64%
	Mildly/partially disagree	174	9.50%	80	8.69%	37	10.11%	43	7.75%	94	10.33%	42	11.67%	52	9.45%
	Neither agree nor disagree	31	1.69%	15	1.63%	13	3.55%	2	0.36%	16	1.76%	9	2.50%	7	1.27%
	Mildly/partially agree	467	25.51%	241	26.17%	87	23.77%	154	27.75%	226	24.84%	95	26.39%	131	23.82%
	Completely agree	668	36.48%	340	36.92%	120	32.79%	220	39.64%	328	36.04%	109	30.28%	219	39.82%
CRIMECON	***In this community, conflicts like stealing and fighting often occur.***														
	Completely disagree	662	36.16%	319	34.64%	165	45.08%	154	27.75%	343	37.69%	179	49.72%	164	29.82%
	Mildly/partially disagree	146	7.97%	73	7.93%	44	12.02%	29	5.23%	73	8.02%	41	11.39%	32	5.82%
	Neither agree nor disagree	25	1.37%	12	1.30%	11	3.01%	1	0.18%	13	1.43%	6	1.67%	7	1.27%
	Mildly/partially agree	599	32.71%	316	34.31%	77	21.04%	239	43.06%	283	31.10%	72	20.00%	211	38.36%
	Completely agree	399	21.79%	201	21.82%	69	18.85%	132	23.78%	198	21.76%	62	17.22%	136	24.73%
SAFEATHO	***When I am at home alone, I feel safe from threats of crime.***														
	Completely disagree	375	20.48%	197	21.39%	92	25.14%	105	18.92%	178	19.56%	80	22.22%	98	17.82%

	Mildly/partially disagree	113	6.17%	67	7.27%	38	10.38%	29	5.23%	46	5.05%	28	7.78%	18	3.27%
	Neither agree nor disagree	60	3.28%	23	2.50%	17	4.64%	6	1.08%	37	4.07%	30	8.33%	7	1.27%
	Mildly/partially agree	379	20.70%	194	21.06%	88	24.04%	106	19.10%	185	20.33%	82	22.78%	103	18.73%
	Completely agree	904	49.37%	440	47.77%	131	35.79%	309	55.68%	464	50.99%	140	38.89%	324	58.91%
SIMBLIEF	***Most people in this community have similar beliefs about what is right and what is wrong.***														
	Completely disagree	325	17.75%	175	19.00%	57	15.57%	118	21.26%	150	16.48%	47	13.06%	103	18.73%
	Mildly/partially disagree	195	10.65%	100	10.86%	28	7.65%	72	12.97%	95	10.44%	26	7.22%	69	12.55%
	Neither agree nor disagree	128	6.99%	56	6.08%	43	11.75%	13	2.34%	72	7.91%	56	15.56%	16	2.91%
	Mildly/partially agree	651	35.55%	336	36.48%	127	34.70%	209	37.66%	315	34.62%	106	29.44%	209	38.00%
	Completely agree	532	29.06%	254	27.58%	111	30.33%	143	25.77%	278	30.55%	125	34.72%	153	27.82%
INTERCRI	***If the people of this community see crime-like activities, they will do something about it.***														
	Completely disagree	186	10.16%	84	9.12%	31	8.47%	53	9.55%	102	11.21%	35	9.72%	67	12.18%
	Mildly/partially disagree	91	4.97%	52	5.65%	21	5.74%	31	5.59%	39	4.29%	19	5.28%	20	3.64%
	Neither agree nor disagree	83	4.53%	42	4.56%	32	8.74%	10	1.80%	41	4.51%	30	8.33%	11	2.00%
	Mildly/partially agree	563	30.75%	294	31.92%	119	32.51%	175	31.53%	269	29.56%	98	27.22%	171	31.09%
	Completely agree	908	49.59%	449	48.75%	163	44.54%	286	51.53%	459	50.44%	178	49.44%	281	51.09%
SLVDISPU	***If there is a big dispute between two persons, other people from the community will help in solving the problem.***														
	Completely disagree	73	3.99%	38	4.13%	18	4.92%	20	3.60%	35	3.85%	17	4.72%	18	3.27%
	Mildly/partially disagree	42	2.29%	24	2.61%	11	3.01%	13	2.34%	18	1.98%	9	2.50%	9	1.64%
	Neither agree nor disagree	30	1.64%	15	1.63%	14	3.83%	1	0.18%	15	1.65%	12	3.33%	3	0.55%
	Mildly/partially agree	451	24.63%	217	23.56%	86	23.50%	131	23.60%	234	25.71%	89	24.72%	145	26.36%
	Completely agree	1235	67.45%	627	68.08%	237	64.75%	390	70.27%	608	66.81%	233	64.72%	375	68.18%
HLPCRPDZ	***If there is a problem that affects the entire community, for instance, crop disease, people in this community will help each other.***														
	Completely disagree	199	10.87%	101	10.97%	53	14.48%	48	8.65%	98	10.77%	47	13.06%	51	9.27%
	Mildly/partially disagree	64	3.50%	32	3.47%	13	3.55%	19	3.42%	32	3.52%	16	4.44%	16	2.91%
	Neither agree nor disagree	64	3.50%	30	3.26%	22	6.01%	8	1.44%	34	3.74%	24	6.67%	10	1.82%

	Mildly/partially agree	567	30.97%	290	31.49%	108	29.51%	182	32.79%	277	30.44%	88	24.44%	189	34.36%
	Completely agree	937	51.17%	468	50.81%	170	46.45%	298	53.69%	469	51.54%	185	51.39%	284	51.64%
SUPMOURN*[1]	***If someone in this community had a death in their family, the community will come together to support them while they mourn.***															
	Completely disagree	6	0.33%	3	0.33%	2	0.55%	1	0.18%	3	0.33%	2	0.56%	1	0.18%
	Mildly/partially disagree	8	0.44%	6	0.65%	3	0.82%	3	0.54%	2	0.22%	1	0.28%	1	0.18%
	Neither agree nor disagree	12	0.66%	6	0.65%	5	1.37%	1	0.18%	6	0.66%	5	1.39%	1	0.18%
	Mildly/partially agree	88	4.81%	55	5.97%	31	8.47%	24	4.32%	33	3.63%	19	5.28%	14	2.55%
	Completely agree	1717	93.77%	851	92.40%	325	88.80%	526	94.77%	866	95.16%	333	92.50%	533	96.91%
COMPRSLV	***When there is a problem in this community, people come together to discuss how it should be solved.***															
	Completely disagree	83	4.53%	43	4.67%	17	4.64%	26	4.68%	40	4.40%	11	3.06%	29	5.27%
	Mildly/partially disagree	37	2.02%	18	1.95%	7	1.91%	11	1.98%	19	2.09%	9	2.50%	10	1.82%
	Neither agree nor disagree	42	2.29%	16	1.74%	13	3.55%	3	0.54%	26	2.86%	20	5.56%	6	1.09%
	Mildly/partially agree	555	30.31%	286	31.05%	99	27.05%	187	33.69%	269	29.56%	93	25.83%	176	32.00%
	Completely agree	1114	60.84%	558	60.59%	230	62.84%	328	59.10%	556	61.10%	227	63.06%	329	59.82%
CONTRDEV	***The people of this community will contribute their own money or labour for community development.***															
	Completely disagree	113	6.17%	55	5.97%	21	5.74%	34	6.13%	58	6.37%	16	4.44%	42	7.64%
	Mildly/partially disagree	61	3.33%	30	3.26%	19	5.19%	11	1.98%	31	3.41%	12	3.33%	19	3.45%
	Neither agree nor disagree	48	2.62%	26	2.82%	19	5.19%	7	1.26%	22	2.42%	20	5.56%	2	0.36%
	Mildly/partially agree	654	35.72%	349	37.89%	126	34.43%	223	40.18%	305	33.52%	116	32.22%	189	34.36%
	Completely agree	955	52.16%	461	50.05%	181	49.45%	280	50.45%	494	54.29%	196	54.44%	298	54.18%
DIFPROBS	***Differences between people, such as the amount of land they own, often causes problems in this community.***															
	Completely disagree	372	20.32%	195	21.17%	72	19.67%	123	22.16%	177	19.45%	72	20.00%	105	19.09%
	Mildly/partially disagree	99	5.41%	50	5.43%	32	8.74%	18	3.24%	49	5.38%	28	7.78%	21	3.82%
	Neither agree nor disagree	74	4.04%	41	4.45%	37	10.11%	4	0.72%	33	3.63%	23	6.39%	10	1.82%
	Mildly/partially agree	479	26.16%	256	27.80%	89	24.32%	167	30.09%	223	24.51%	78	21.67%	145	26.36%
	Completely agree	807	44.07%	379	41.15%	136	37.16%	243	43.78%	428	47.03%	159	44.17%	269	48.91%

HAPPYNEI *[1]	***I feel happy for my neighbour if they have a good harvest.***																
	Completely disagree	13	0.71%	8	0.87%	8	2.19%	0	0.00%	5	0.55%	4	1.11%	1	0.18%		
	Mildly/partially disagree	4	0.22%	2	0.22%	1	0.27%	1	0.18%	2	0.22%	0	0.00%	2	0.36%		
	Neither agree nor disagree	12	0.66%	7	0.76%	5	1.37%	2	0.36%	5	0.55%	5	1.39%	0	0.00%		
	Mildly/partially agree	78	4.26%	46	4.99%	24	6.56%	22	3.96%	32	3.52%	13	3.61%	19	3.45%		
	Completely agree	1724	94.16%	858	93.16%	328	89.62%	530	95.50%	866	95.16%	338	93.89%	528	96.00%		
COMTRUST	***People in this community can be trusted.***																
	Completely disagree	90	4.92%	42	4.56%	18	4.92%	24	4.32%	48	5.27%	23	6.39%	25	4.55%		
	Mildly/partially disagree	71	3.88%	42	4.56%	19	5.19%	23	4.14%	29	3.19%	14	3.89%	15	2.73%		
	Neither agree nor disagree	58	3.17%	20	2.17%	12	3.28%	8	1.44%	38	4.18%	30	8.33%	8	1.45%		
	Mildly/partially agree	772	42.16%	396	43.00%	127	34.70%	269	48.47%	376	41.32%	115	31.94%	261	47.45%		
	Completely agree	840	45.88%	421	45.71%	190	51.91%	231	41.62%	419	46.04%	178	49.44%	241	43.82%		
ADVICE	***I typically accept advice from others in this community.***																
	Completely disagree	18	0.98%	10	1.09%	8	2.19%	2	0.36%	8	0.88%	4	1.11%	4	0.73%		
	Mildly/partially disagree	20	1.09%	7	0.76%	2	0.55%	5	0.90%	13	1.43%	11	3.06%	2	0.36%		
	Neither agree nor disagree	17	0.93%	7	0.76%	6	1.64%	1	0.18%	10	1.10%	10	2.78%	0	0.00%		
	Mildly/partially agree	646	35.28%	342	37.13%	149	40.71%	193	34.77%	304	33.41%	120	33.33%	184	33.45%		
	Completely agree	1130	61.71%	555	60.26%	201	54.92%	354	63.78%	575	63.19%	215	59.72%	360	65.45%		
SHAREKNO	***People in the community share new knowledge with their neighbour if they learn something new.***																
	Completely disagree	155	8.47%	65	7.06%	19	5.19%	46	8.29%	90	9.89%	35	9.72%	55	10.00%		
	Mildly/partially disagree	77	4.21%	32	3.47%	10	2.73%	22	3.96%	45	4.95%	26	7.22%	19	3.45%		
	Neither agree nor disagree	82	4.48%	37	4.02%	16	4.37%	21	3.78%	45	4.95%	28	7.78%	17	3.09%		
	Mildly/partially agree	630	34.41%	335	36.37%	135	36.89%	200	36.04%	295	32.42%	103	28.61%	192	34.91%		
	Completely agree	887	48.44%	452	49.08%	186	50.82%	266	47.93%	435	47.80%	168	46.67%	267	48.55%		
CLOSE	***This is a close-knit community (i.e., people in this community have close personal relationships with each other).***																
	Completely disagree	61	3.33%	32	3.47%	17	4.64%	15	2.70%	29	3.19%	10	2.78%	19	3.45%		

	Mildly/partially disagree	60	3.28%	32	3.47%	13	3.55%	19	3.42%	28	3.08%	14	3.89%	14	2.55%
	Neither agree nor disagree	36	1.97%	15	1.63%	12	3.28%	3	0.54%	21	2.31%	17	4.72%	4	0.73%
	Mildly/partially agree	684	37.36%	352	38.22%	126	34.43%	226	40.72%	332	36.48%	124	34.44%	208	37.82%
	Completely agree	990	54.07%	490	53.20%	198	54.10%	292	52.61%	500	54.95%	195	54.17%	305	55.45%
OWNWELF	***In this community, people prioritise their own family's welfare over community development.***														
	Completely disagree	119	6.50%	56	6.08%	25	6.83%	31	5.59%	63	6.92%	26	7.22%	37	6.73%
	Mildly/partially disagree	60	3.28%	40	4.34%	27	7.38%	13	2.34%	20	2.20%	11	3.06%	9	1.64%
	Neither agree nor disagree	54	2.95%	26	2.82%	24	6.56%	2	0.36%	28	3.08%	20	5.56%	8	1.45%
	Mildly/partially agree	392	21.41%	207	22.48%	108	29.51%	99	17.84%	185	20.33%	96	26.67%	89	16.18%
	Completely agree	1206	65.87%	592	64.28%	182	49.73%	410	73.87%	614	67.47%	207	57.50%	407	74.00%
LOSTCOW *[1]	***If someone in this community loses a cow or goat, a neighbour will help look for it.***														
	Completely disagree	10	0.55%	5	0.54%	2	0.55%	3	0.54%	5	0.55%	5	1.39%	0	0.00%
	Mildly/partially disagree	6	0.33%	4	0.43%	2	0.55%	2	0.36%	2	0.22%	1	0.28%	1	0.18%
	Neither agree nor disagree	10	0.55%	5	0.54%	4	1.09%	1	0.18%	5	0.55%	5	1.39%	0	0.00%
	Mildly/partially agree	172	9.39%	80	8.69%	26	7.10%	54	9.73%	92	10.11%	26	7.22%	66	12.00%
	Completely agree	1633	89.19%	827	89.79%	332	90.71%	495	89.19%	806	88.57%	323	89.72%	483	87.82%
BORMONEY	***If you suddenly need some money, you can borrow from a person or group in your community.***														
	Completely disagree	153	8.36%	78	8.47%	36	9.84%	42	7.57%	75	8.24%	33	9.17%	42	7.64%
	Mildly/partially disagree	47	2.57%	18	1.95%	9	2.46%	9	1.62%	29	3.19%	16	4.44%	13	2.36%
	Neither agree nor disagree	22	1.20%	11	1.19%	5	1.37%	6	1.08%	11	1.21%	8	2.22%	3	0.55%
	Mildly/partially agree	481	26.27%	243	26.38%	93	25.41%	150	27.03%	238	26.15%	83	23.06%	155	28.18%
	Completely agree	1128	61.61%	571	62.00%	223	60.93%	348	62.70%	557	61.21%	220	61.11%	337	61.27%
NEICAREG	***If you and your relatives suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.***														
	Completely disagree	109	5.95%	48	5.21%	22	6.01%	26	4.68%	61	6.70%	31	8.61%	30	5.45%
	Mildly/partially disagree	42	2.29%	22	2.39%	6	1.64%	16	2.88%	20	2.20%	8	2.22%	12	2.18%
	Neither agree nor disagree	26	1.42%	11	1.19%	4	1.09%	7	1.26%	15	1.65%	5	1.39%	10	1.82%

	Mildly/partially agree	394	21.52%	198	21.50%	80	21.86%	118	21.26%	196	21.54%	75	20.83%	121	22.00%
	Completely agree	1260	68.81%	642	69.71%	254	69.40%	388	69.91%	618	67.91%	241	66.94%	377	68.55%
UNOFLDRS	***There are people in this community who show strong leadership.***														
	Completely disagree	64	3.50%	27	2.93%	9	2.46%	18	3.24%	37	4.07%	15	4.17%	22	4.00%
	Mildly/partially disagree	50	2.73%	23	2.50%	10	2.73%	13	2.34%	27	2.97%	13	3.61%	14	2.55%
	Neither agree nor disagree	112	6.12%	44	4.78%	33	9.02%	11	1.98%	68	7.47%	50	13.89%	18	3.27%
	Mildly/partially agree	758	41.40%	407	44.19%	141	38.52%	266	47.93%	351	38.57%	118	32.78%	233	42.36%
	Completely agree	847	46.26%	420	45.60%	173	47.27%	247	44.50%	427	46.92%	164	45.56%	263	47.82%
COMACTCG*	***The community-based associations, such as the Edir, in this community is very active.***														
	Completely disagree	26	1.42%	10	1.09%	6	1.64%	4	0.72%	16	1.76%	5	1.39%	11	2.00%
	Mildly/partially disagree	19	1.04%	7	0.76%	3	0.82%	4	0.72%	12	1.32%	7	1.94%	5	0.91%
	Neither agree nor disagree	39	2.13%	24	2.61%	16	4.37%	8	1.44%	15	1.65%	13	3.61%	2	0.36%
	Mildly/partially agree	293	16.00%	145	15.74%	76	20.77%	69	12.43%	148	16.26%	88	24.44%	60	10.91%
	Completely agree	1454	79.41%	735	79.80%	265	72.40%	470	84.68%	719	79.01%	247	68.61%	472	85.82%
ACTLDR1	***The leaders of community-based associations, like Edir leaders, respond to this community's concerns.***														
	Completely disagree	38	2.08%	15	1.63%	5	1.37%	10	1.80%	23	2.53%	8	2.22%	15	2.73%
	Mildly/partially disagree	20	1.09%	10	1.09%	7	1.91%	3	0.54%	10	1.10%	5	1.39%	5	0.91%
	Neither agree nor disagree	76	4.15%	42	4.56%	28	7.65%	14	2.52%	34	3.74%	27	7.50%	7	1.27%
	Mildly/partially agree	408	22.28%	215	23.34%	98	26.78%	117	21.08%	193	21.21%	97	26.94%	96	17.45%
	Completely agree	1289	70.40%	639	69.38%	228	62.30%	411	74.05%	650	71.43%	223	61.94%	427	77.64%
ACTLDR2	***Formal administrative leaders, like the kebele manager, provide support to this community.***														
	Completely disagree	125	6.83%	59	6.41%	19	5.19%	40	7.21%	66	7.25%	25	6.94%	41	7.45%
	Mildly/partially disagree	71	3.88%	33	3.58%	14	3.83%	19	3.42%	38	4.18%	15	4.17%	23	4.18%
	Neither agree nor disagree	139	7.59%	64	6.95%	50	13.66%	14	2.52%	75	8.24%	57	15.83%	18	3.27%
	Mildly/partially agree	644	35.17%	332	36.05%	123	33.61%	209	37.66%	312	34.29%	108	30.00%	204	37.09%
	Completely agree	852	46.53%	433	47.01%	160	43.72%	273	49.19%	419	46.04%	155	43.06%	264	48.00%

TRUSTLDR	***This community's leaders can be trusted.***																
	Completely disagree	97	5.30%	45	4.89%	17	4.64%	28	5.05%	52	5.71%	15	4.17%	37	6.73%		
	Mildly/partially disagree	73	3.99%	36	3.91%	16	4.37%	20	3.60%	37	4.07%	19	5.28%	18	3.27%		
	Neither agree nor disagree	101	5.52%	40	4.34%	27	7.38%	13	2.34%	61	6.70%	48	13.33%	13	2.36%		
	Mildly/partially agree	722	39.43%	377	40.93%	137	37.43%	240	43.24%	345	37.91%	110	30.56%	235	42.73%		
	Completely agree	838	45.77%	423	45.93%	169	46.17%	254	45.77%	415	45.60%	168	46.67%	247	44.91%		
CHOCGLDR*	***People in this community get to choose the leaders of their own community-based associations, such as the Edir leaders.***																
	Completely disagree	32	1.75%	12	1.30%	6	1.64%	6	1.08%	20	2.20%	6	1.67%	14	2.55%		
	Mildly/partially disagree	12	0.66%	7	0.76%	6	1.64%	1	0.18%	5	0.55%	3	0.83%	2	0.36%		
	Neither agree nor disagree	49	2.68%	22	2.39%	13	3.55%	9	1.62%	27	2.97%	20	5.56%	7	1.27%		
	Mildly/partially agree	254	13.87%	137	14.88%	76	20.77%	61	10.99%	117	12.86%	66	18.33%	51	9.27%		
	Completely agree	1484	81.05%	743	80.67%	265	72.40%	478	86.13%	741	81.43%	265	73.61%	476	86.55%		
HAVEFRND	***In this community, I have friends with whom I can share my problems.***																
	Completely disagree	90	4.92%	41	4.45%	25	6.83%	16	2.88%	49	5.38%	29	8.06%	20	3.64%		
	Mildly/partially disagree	22	1.20%	12	1.30%	8	2.19%	4	0.72%	10	1.10%	6	1.67%	4	0.73%		
	Neither agree nor disagree	14	0.76%	8	0.87%	8	2.19%	0	0.00%	6	0.66%	6	1.67%	0	0.00%		
	Mildly/partially agree	464	25.34%	239	25.95%	95	25.96%	144	25.95%	225	24.73%	87	24.17%	138	25.09%		
	Completely agree	1241	67.78%	621	67.43%	230	62.84%	391	70.45%	620	68.13%	232	64.44%	388	70.55%		
COME4HLP	***My neighbours sometimes come to me to share their problems and get help.***																
	Completely disagree	88	4.81%	42	4.56%	23	6.28%	19	3.42%	46	5.05%	26	7.22%	20	3.64%		
	Mildly/partially disagree	32	1.75%	15	1.63%	9	2.46%	6	1.08%	17	1.87%	11	3.06%	6	1.09%		
	Neither agree nor disagree	15	0.82%	6	0.65%	5	1.37%	1	0.18%	9	0.99%	8	2.22%	1	0.18%		
	Mildly/partially agree	575	31.40%	284	30.84%	125	34.15%	159	28.65%	291	31.98%	118	32.78%	173	31.45%		
	Completely agree	1121	61.22%	574	62.32%	204	55.74%	370	66.67%	547	60.11%	197	54.72%	350	63.64%		
COPARTCG	***Most people in this community participate in community associations.***																
	Completely disagree	62	3.39%	35	3.80%	21	5.74%	14	2.52%	27	2.97%	16	4.44%	11	2.00%		

	Mildly/partially disagree	35	1.91%	17	1.85%	10	2.73%	7	1.26%	18	1.98%	10	2.78%	8	1.45%
	Neither agree nor disagree	66	3.60%	26	2.82%	20	5.46%	6	1.08%	40	4.40%	29	8.06%	11	2.00%
	Mildly/partially agree	575	31.40%	297	32.25%	121	33.06%	176	31.71%	278	30.55%	104	28.89%	174	31.64%
	Completely agree	1093	59.69%	546	59.28%	194	53.01%	352	63.42%	547	60.11%	201	55.83%	346	62.91%
ACTCBGP	***I attend meetings of a community-based association, such as the Edir.***														
	Completely disagree	263	14.36%	126	13.68%	86	23.50%	40	7.21%	137	15.05%	92	25.56%	45	8.18%
	Mildly/partially disagree	44	2.40%	22	2.39%	15	4.10%	7	1.26%	22	2.42%	20	5.56%	2	0.36%
	Neither agree nor disagree	33	1.80%	16	1.74%	10	2.73%	6	1.08%	17	1.87%	13	3.61%	4	0.73%
	Mildly/partially agree	403	22.01%	205	22.26%	111	30.33%	94	16.94%	198	21.76%	101	28.06%	97	17.64%
	Completely agree	1088	59.42%	552	59.93%	144	39.34%	408	73.51%	536	58.90%	134	37.22%	402	73.09%
PARTCBGP	***I participate in activities held by any community-based associations, such as the Edir.***														
	Completely disagree	204	11.14%	92	9.99%	60	16.39%	32	5.77%	112	12.31%	66	18.33%	46	8.36%
	Mildly/partially disagree	49	2.68%	24	2.61%	14	3.83%	10	1.80%	25	2.75%	19	5.28%	6	1.09%
	Neither agree nor disagree	34	1.86%	15	1.63%	9	2.46%	6	1.08%	19	2.09%	17	4.72%	2	0.36%
	Mildly/partially agree	390	21.30%	206	22.37%	105	28.69%	101	18.20%	184	20.22%	96	26.67%	88	16.00%
	Completely agree	1154	63.03%	584	63.41%	178	48.63%	406	73.15%	570	62.64%	162	45.00%	408	74.18%
ACTEXOGP	***I attend the meetings of any government or NGO-initiated community development group, such as the Development Army.***														
	Completely disagree	315	17.20%	156	16.94%	104	28.42%	52	9.37%	159	17.47%	98	27.22%	61	11.09%
	Mildly/partially disagree	99	5.41%	46	4.99%	28	7.65%	18	3.24%	53	5.82%	32	8.89%	21	3.82%
	Neither agree nor disagree	53	2.89%	24	2.61%	16	4.37%	8	1.44%	29	3.19%	26	7.22%	3	0.55%
	Mildly/partially agree	437	23.87%	235	25.52%	104	28.42%	131	23.60%	202	22.20%	93	25.83%	109	19.82%
	Completely agree	927	50.63%	460	49.95%	114	31.15%	346	62.34%	467	51.32%	111	30.83%	356	64.73%
PAREXOGP	***I participate in activities held by any government or NGO-initiated community development group, such as the Development Army.***														
	Completely disagree	342	18.68%	169	18.35%	104	28.42%	65	11.71%	173	19.01%	108	30.00%	65	11.82%
	Mildly/partially disagree	106	5.79%	50	5.43%	30	8.20%	20	3.60%	56	6.15%	33	9.17%	23	4.18%
	Neither agree nor disagree	55	3.00%	25	2.71%	17	4.64%	8	1.44%	30	3.30%	27	7.50%	3	0.55%

	Mildly/partially agree	461	25.18%		240	26.06%	100	27.32%	140	25.23%		221	24.29%	96	26.67%	125	22.73%
	Completely agree	867	47.35%		437	47.45%	115	31.42%	322	58.02%		430	47.25%	96	26.67%	334	60.73%
COMMGDEC	***When community groups make decisions, they are pleasing and good for most of the households in this community.***																
	Completely disagree	87	4.75%		42	4.56%	24	6.56%	18	3.24%		45	4.95%	22	6.11%	23	4.18%
	Mildly/partially disagree	62	3.39%		31	3.37%	17	4.64%	14	2.52%		31	3.41%	19	5.28%	12	2.18%
	Neither agree nor disagree	131	7.15%		64	6.95%	42	11.48%	22	3.96%		67	7.36%	47	13.06%	20	3.64%
	Mildly/partially agree	707	38.61%		363	39.41%	145	39.62%	218	39.28%		344	37.80%	127	35.28%	217	39.45%
	Completely agree	844	46.10%		421	45.71%	138	37.70%	283	50.99%		423	46.48%	145	40.28%	278	50.55%
BRIBELDR	***Sometimes people need to bribe community leaders in order to get things done.***																
	Completely disagree	1163	63.52%		584	63.41%	187	51.09%	397	71.53%		579	63.63%	193	53.61%	386	70.18%
	Mildly/partially disagree	104	5.68%		61	6.62%	31	8.47%	30	5.41%		43	4.73%	21	5.83%	22	4.00%
	Neither agree nor disagree	207	11.31%		93	10.10%	65	17.76%	28	5.05%		114	12.53%	72	20.00%	42	7.64%
	Mildly/partially agree	185	10.10%		97	10.53%	37	10.11%	60	10.81%		88	9.67%	33	9.17%	55	10.00%
	Completely agree	172	9.39%		86	9.34%	46	12.57%	40	7.21%		86	9.45%	41	11.39%	45	8.18%
DISTCRIS	***During a crisis situation, such as a drought, government services are distributed equally by the community to all households in need.***																
	Completely disagree	185	10.10%		92	9.99%	29	7.92%	63	11.35%		93	10.22%	39	10.83%	54	9.82%
	Mildly/partially disagree	108	5.90%		57	6.19%	23	6.28%	34	6.13%		51	5.60%	17	4.72%	34	6.18%
	Neither agree nor disagree	292	15.95%		147	15.96%	74	20.22%	73	13.15%		145	15.93%	79	21.94%	66	12.00%
	Mildly/partially agree	553	30.20%		273	29.64%	104	28.42%	169	30.45%		280	30.77%	98	27.22%	182	33.09%
	Completely agree	693	37.85%		352	38.22%	136	37.16%	216	38.92%		341	37.47%	127	35.28%	214	38.91%
RESTRSER	***Some households in this community are restricted from community services, such as bed net distribution.***																
	Completely disagree	993	54.23%		496	53.85%	162	44.26%	334	60.18%		497	54.62%	173	48.06%	324	58.91%
	Mildly/partially disagree	192	10.49%		92	9.99%	42	11.48%	50	9.01%		100	10.99%	44	12.22%	56	10.18%
	Neither agree nor disagree	135	7.37%		64	6.95%	37	10.11%	27	4.86%		71	7.80%	48	13.33%	23	4.18%
	Mildly/partially agree	256	13.98%		133	14.44%	59	16.12%	74	13.33%		123	13.52%	53	14.72%	70	12.73%
	Completely agree	255	13.93%		136	14.77%	66	18.03%	70	12.61%		119	13.08%	42	11.67%	77	14.00%

COMMVALU	***Most people in this community have common values, for example, they value hard work.***																
	Completely disagree	36	1.97%	15	1.63%	6	1.64%	9	1.62%	21	2.31%	9	2.50%	12	2.18%		
	Mildly/partially disagree	33	1.80%	15	1.63%	7	1.91%	8	1.44%	18	1.98%	11	3.06%	7	1.27%		
	Neither agree nor disagree	75	4.10%	43	4.67%	33	9.02%	10	1.80%	32	3.52%	29	8.06%	3	0.55%		
	Mildly/partially agree	573	31.29%	287	31.16%	126	34.43%	161	29.01%	286	31.43%	111	30.83%	175	31.82%		
	Completely agree	1114	60.84%	561	60.91%	194	53.01%	367	66.13%	553	60.77%	200	55.56%	353	64.18%		
SIMHOPES	***Most people in this community have similar hopes about the future development of the community.***																
	Completely disagree	56	3.06%	27	2.93%	12	3.28%	15	2.70%	29	3.19%	10	2.78%	19	3.45%		
	Mildly/partially disagree	46	2.51%	27	2.93%	13	3.55%	14	2.52%	19	2.09%	13	3.61%	6	1.09%		
	Neither agree nor disagree	132	7.21%	67	7.27%	48	13.11%	19	3.42%	65	7.14%	48	13.33%	17	3.09%		
	Mildly/partially agree	637	34.79%	328	35.61%	133	36.34%	195	35.14%	309	33.96%	127	35.28%	182	33.09%		
	Completely agree	960	52.43%	472	51.25%	160	43.72%	312	56.22%	488	53.63%	162	45.00%	326	59.27%		
COMMGMT	***People in this community share the same ideas on how village matters should be managed.***																
	Completely disagree	63	3.44%	27	2.93%	13	3.55%	14	2.52%	36	3.96%	18	5.00%	18	3.27%		
	Mildly/partially disagree	54	2.95%	25	2.71%	11	3.01%	14	2.52%	29	3.19%	16	4.44%	13	2.36%		
	Neither agree nor disagree	117	6.39%	62	6.73%	45	12.30%	17	3.06%	55	6.04%	39	10.83%	16	2.91%		
	Mildly/partially agree	635	34.68%	314	34.09%	126	34.43%	188	33.87%	321	35.27%	126	35.00%	195	35.45%		
	Completely agree	962	52.54%	493	53.53%	171	46.72%	322	58.02%	469	51.54%	161	44.72%	308	56.00%		
ACCEPT*	***People in this community accept me as a member of the community.***																
	Completely disagree	39	2.13%	17	1.85%	15	4.10%	2	0.36%	22	2.42%	16	4.44%	6	1.09%		
	Mildly/partially disagree	27	1.47%	13	1.41%	6	1.64%	7	1.26%	14	1.54%	8	2.22%	6	1.09%		
	Neither agree nor disagree	39	2.13%	20	2.17%	16	4.37%	4	0.72%	19	2.09%	17	4.72%	2	0.36%		
	Mildly/partially agree	314	17.15%	164	17.81%	97	26.50%	67	12.07%	150	16.48%	85	23.61%	65	11.82%		
	Completely agree	1412	77.12%	707	76.76%	232	63.39%	475	85.59%	705	77.47%	234	65.00%	471	85.64%		
ATTACH*	***I feel attached to this community and its people.***																
	Completely disagree	31	1.69%	16	1.74%	10	2.73%	6	1.08%	15	1.65%	11	3.06%	4	0.73%		

	Mildly/partially disagree	25	1.37%	7	0.76%	4	1.09%	3	0.54%	18	1.98%	9	2.50%	9	1.64%
	Neither agree nor disagree	25	1.37%	11	1.19%	9	2.46%	2	0.36%	14	1.54%	12	3.33%	2	0.36%
	Mildly/partially agree	327	17.86%	166	18.02%	100	27.32%	66	11.89%	161	17.69%	91	25.28%	70	12.73%
	Completely agree	1423	77.72%	721	78.28%	243	66.39%	478	86.13%	702	77.14%	237	65.83%	465	84.55%
PROUD*	***I feel proud to be part of this community.***														
	Completely disagree	40	2.18%	16	1.74%	12	3.28%	4	0.72%	24	2.64%	18	5.00%	6	1.09%
	Mildly/partially disagree	25	1.37%	8	0.87%	8	2.19%	0	0.00%	17	1.87%	10	2.78%	7	1.27%
	Neither agree nor disagree	18	0.98%	8	0.87%	5	1.37%	3	0.54%	10	1.10%	10	2.78%	0	0.00%
	Mildly/partially agree	219	11.96%	112	12.16%	70	19.13%	42	7.57%	107	11.76%	67	18.61%	40	7.27%
	Completely agree	1529	83.51%	777	84.36%	271	74.04%	506	91.17%	752	82.64%	255	70.83%	497	90.36%
IDENTITY*	***Being a member of this community is part of who I am.***														
	Completely disagree	27	1.47%	13	1.41%	9	2.46%	4	0.72%	14	1.54%	12	3.33%	2	0.36%
	Mildly/partially disagree	25	1.37%	9	0.98%	6	1.64%	3	0.54%	16	1.76%	10	2.78%	6	1.09%
	Neither agree nor disagree	35	1.91%	19	2.06%	15	4.10%	4	0.72%	16	1.76%	14	3.89%	2	0.36%
	Mildly/partially agree	211	11.52%	107	11.62%	67	18.31%	40	7.21%	104	11.43%	68	18.89%	36	6.55%
	Completely agree	1533	83.72%	773	83.93%	269	73.50%	504	90.81%	760	83.52%	256	71.11%	504	91.64%
SELFEFF*	***I have the capacity to achieve my future aims.***														
	Completely disagree	36	1.97%	19	2.06%	7	1.91%	12	2.16%	17	1.87%	9	2.50%	8	1.45%
	Mildly/partially disagree	13	0.71%	5	0.54%	4	1.09%	1	0.18%	8	0.88%	3	0.83%	5	0.91%
	Neither agree nor disagree	38	2.08%	20	2.17%	14	3.83%	6	1.08%	18	1.98%	11	3.06%	7	1.27%
	Mildly/partially agree	375	20.48%	187	20.30%	96	26.23%	91	16.40%	188	20.66%	86	23.89%	102	18.55%
	Completely agree	1369	74.77%	690	74.92%	245	66.94%	445	80.18%	679	74.62%	251	69.72%	428	77.82%
SEDEV	***I have the ability to contribute to this community's development.***														
	Completely disagree	66	3.60%	38	4.13%	22	6.01%	16	2.88%	28	3.08%	15	4.17%	13	2.36%
	Mildly/partially disagree	66	3.60%	25	2.71%	17	4.64%	8	1.44%	41	4.51%	30	8.33%	11	2.00%
	Neither agree nor disagree	92	5.02%	44	4.78%	35	9.56%	9	1.62%	48	5.27%	30	8.33%	18	3.27%

	Mildly/partially agree	544	29.71%	273	29.64%	111	30.33%	162	29.19%	271	29.78%	102	28.33%	169	30.73%
	Completely agree	1063	58.06%	541	58.74%	181	49.45%	360	64.86%	522	57.36%	183	50.83%	339	61.64%
COLLEFF	***People in this community have the capacity to make positive changes by coming together.***														
	Completely disagree	51	2.79%	21	2.28%	10	2.73%	11	1.98%	30	3.30%	11	3.06%	19	3.45%
	Mildly/partially disagree	34	1.86%	13	1.41%	5	1.37%	8	1.44%	21	2.31%	9	2.50%	12	2.18%
	Neither agree nor disagree	119	6.50%	63	6.84%	50	13.66%	13	2.34%	56	6.15%	44	12.22%	12	2.18%
	Mildly/partially agree	545	29.77%	276	29.97%	108	29.51%	168	30.27%	269	29.56%	114	31.67%	155	28.18%
	Completely agree	1082	59.09%	548	59.50%	193	52.73%	355	63.96%	534	58.68%	182	50.56%	352	64.00%
EXOASSIS	***This community needs assistance from others outside the community in order to make positive changes.***														
	Completely disagree	296	16.17%	144	15.64%	31	8.47%	113	20.36%	152	16.70%	45	12.50%	107	19.45%
	Mildly/partially disagree	84	4.59%	48	5.21%	27	7.38%	21	3.78%	36	3.96%	17	4.72%	19	3.45%
	Neither agree nor disagree	105	5.73%	57	6.19%	41	11.20%	16	2.88%	48	5.27%	38	10.56%	10	1.82%
	Mildly/partially agree	507	27.69%	251	27.25%	98	26.78%	153	27.57%	256	28.13%	108	30.00%	148	26.91%
	Completely agree	839	45.82%	421	45.71%	169	46.17%	252	45.41%	418	45.93%	152	42.22%	266	48.36%
SHOULDEV*	***People in this community should work together to develop the community.***														
	Completely disagree	36	1.97%	16	1.74%	6	1.64%	10	1.80%	20	2.20%	9	2.50%	11	2.00%
	Mildly/partially disagree	23	1.26%	9	0.98%	4	1.09%	5	0.90%	14	1.54%	7	1.94%	7	1.27%
	Neither agree nor disagree	44	2.40%	23	2.50%	19	5.19%	4	0.72%	21	2.31%	17	4.72%	4	0.73%
	Mildly/partially agree	319	17.42%	161	17.48%	67	18.31%	94	16.94%	158	17.36%	71	19.72%	87	15.82%
	Completely agree	1409	76.95%	712	77.31%	270	73.77%	442	79.64%	697	76.59%	256	71.11%	441	80.18%

Notes. * Items with skewness outside the suggested range (i.e., > 3.0), distributions of aggregate men's CE item responses

[1] Items with skewness outside the suggested range (i.e., > 3.0), distributions of aggregate women's CE items responses

2

Table 4.1. Collective efficacy factor scores, by gender and household leadership status, and comparison of household-level male-female dyads

Characteristic	Factor 1: Social response	Factor 2: Social networks & agency	Factor 6: Social organisation & ldrship	Factor 5: Associational participation	Factor 3: Social attachment	Factor 4: Common vision
Weighted average CE factor scores						
Aggregate (N=1,831)	3.0 (0.50)	2.9 (0.43)	3.4 (0.46)	3.6 (1.04)	4.2 (0.55)	3.2 (0.51)
Sex						
Men (n=1,105)	3.1 (0.47)*	3.0 (0.40)***	3.5 (0.41)***	3.9 (0.87)***	4.3 (0.39)***	3.2 (0.48)***
Women (n=726)	3.0 (0.53)*	2.9 (0.47)***	3.4 (0.51)***	3.1 (1.21)***	4.0 (0.70)***	3.1 (0.55)***
Household leadership status						
Any leader (n=313)	3.1 (0.52)	3.0 (0.40)*	3.4 (0.47)	3.7 (1.04)*	4.2 (0.52)	3.2 (0.51)
No leader (n=1,510)	3.0 (0.50)	2.9 (0.44)*	3.4 (0.46)	3.5 (1.04)*	4.1 (0.56)	3.2 (0.51)
Household-level male-female dyads (n=520)						
Men	3.1 (0.45)	3.0 (0.41)***	3.5 (0.42)*	3.9 (0.81)***	4.3 (0.39)***	3.3 (0.46)***
Women	3.0 (0.50)	2.9 (0.47)***	3.4 (0.49)*	3.1 (1.13)***	4.0 (0.66)***	3.1 (0.53)***
Average CE factor scores						
Aggregate (N=1,831)	4.3 (0.71)	4.4 (0.64)	4.5 (0.60)	4.1 (1.20)	4.7 (0.62)	4.3 (0.70)
Sex						
Men (n=1,105)	4.3 (0.67)*	4.5 (0.59)***	4.5 (0.54)***	4.4 (0.99)***	4.8 (0.44)***	4.3 (0.66)***
Women (n=726)	4.3 (0.76)*	4.3 (0.70)***	4.4 (0.67)***	3.5 (1.28)***	4.5 (0.79)***	4.1 (0.74)***
Household leadership status						
Any leader (n=313)	4.3 (0.73)	4.5 (0.58)*	4.5 (0.62)	4.0 (1.20)*	4.7 (0.58)	4.2 (1.18)
No leader (n=1,510)	4.3 (0.70)	4.4 (0.65)*	4.5 (0.60)	4.0 (1.20)*	4.7 (0.63)	4.2 (0.70)
Household-level male-female dyads (n=520)						
Men	4.4 (0.64)	4.5 (0.61)***	4.5 (0.55)*	4.5 (0.93)***	4.8 (0.44)***	4.4 (0.63)***
Women	4.3 (0.72)	4.4 (0.69)***	4.4 (0.64)*	3.5 (1.29)***	4.5 (0.75)***	4.2 (0.72)***

Notes. Mean factor scores and standard deviations (SD) are presented. Weighted average CE factor scores reflect those for which a weight that was equivalent to the item's factor loading was applied to each item score prior to the generation of the average (weighted) factor score

*** $p \leq 0.001$, * $p \leq 0.05$ on non-parametric equality of medians tests

4.7 Chapter references

1. Spector, P., *Summated Rating Scale Construction*. 1992: Newbury Park, California.
2. Clark, L.A. and D. Watson, *Constructing validity: Basic issues in objective scale development.* Psychological assessment, 1995. **7**(3): p. 309.
3. Grootaert, C. and T. Van Bastelar, *Understanding and Measuring Social Capital A Multidisciplinary Tool for Practitioners.* 2013.
4. Sampson, R.J., S.W. Raudenbush, and F. Earls, *Neighborhoods and violent crime: a multilevel study of collective efficacy.* Science, 1997. **277**(5328): p. 918-24.
5. Beatty, P.C. and G.B. Willis, *Research Synthesis: The Practice of Cognitive Interviewing.* Public Opinion Quarterly, 2007. **71**(2): p. 287-311.
6. Castillo-Díaz, M. and J.-L. Padilla, *How Cognitive Interviewing can Provide Validity Evidence of the Response Processes to Scale Items.* Social Indicators Research, 2013. **114**(3): p. 963-975.
7. Guest, G., K.M. MacQueen, and E.E. Namey, *Applied Thematic Analysis*. 2011: SAGE Publications.
8. Wutich, A. and K. Ragsdale, *Water insecurity and emotional distress: Coping with supply, access, and seasonal variability of water in a Bolivian squatter settlement.* Social Science & Medicine, 2008. **67**(12): p. 2116-2125.
9. Hadley, C. and A. Wutich, *Experience-based measures of food and water security: Biocultural approaches to grounded measures of insecurity.* Human Organization, 2009. **68**(4): p. 451-460.
10. Stevenson, E.G.J., L.E. Greene, K.C. Maes, A. Ambelu, Y.A. Tesfaye, R. Rheingans, and C. Hadley, *Water insecurity in 3 dimensions: An anthropological perspective on water and women's psychosocial distress in Ethiopia.* Social Science & Medicine, 2012. **75**(2): p. 392-400.
11. Brown, T.A., *Confirmatory Factor Analysis for Applied Research*. 2006: Guilford Publications.
12. De Silva, M.J., T. Harpham, T. Tuan, R. Bartolini, M.E. Penny, and S.R. Huttly, *Psychometric and cognitive validation of a social capital measurement tool in Peru and Vietnam.* Social Science & Medicine, 2006. **62**(4): p. 941-953.
13. Pett, M.A., N.R. Lackey, and J.J. Sullivan, *Making sense of factor analysis the use of factor analysis for instrument development in health care research*. 2003, Thousand Oaks, Calif.; London: SAGE.
14. Steinmetz, H., P. Schmidt, A. Tina-Booh, S. Wieczorek, and S.H. Schwartz, *Testing measurement invariance using multigroup CFA: differences between educational groups in human values measurement.* Quality & Quantity, 2008. **43**(4): p. 599.
15. Thurstone, L.L., *Multiple-factor analysis*. 1947, Chicago; London: University of Chicago Press.
16. Muthen, L.K. and B. Muthen, *Mplus : statistical analysis with latent variables : user's guide*. 2015.
17. Zumbo, B., *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. 1999, Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
18. Yong, A. and S. Pearce, *A beginner's guide to factor analysis: Focusing on exploratory factor analysis.* Tutorials in Quantitative Methods for Psychology, 2013. **9**(2): p. 72-94.
19. Novick, M.R., *The axioms and principal results of classical test theory.* Journal of Mathematical Psychology, 1966. **3**: p. 1-18.
20. Hambleton, R.K., H. Swaminathan, and J.H. Rogers, *Fundamentals of item response theory*. 1991, New York: Sage publications.
21. Thompson, B., *Exploratory and confirmatory factor analysis: understanding concepts and applications.* 2004, American Psychological Association: Washington, D.C.
22. Cappelleri, J.C., A.G. Bushmakin, R.A. Gerber, N.K. Leidy, C.C. Sexton, J. Karlsson, and M.R. Lowe, *Evaluating the power of food scale in obese subjects and a general sample of*

individuals: Development and measurement properties. International Journal of Obesity, 2009. **33**(8): p. 913-922.
23. Jolliffe, I.T., *Principal Component Analysis*. 2002, New York: Springer.
24. Andrich, D., *Controversy and the Rasch Model: A Characteristic of Incompatible Paradigms?* Medical Care, 2004. **42**(1): p. I-7.
25. Mellenbergh, G.J., *Item bias and item response theory.* JIJER</cja:jid> International Journal of Educational Research, 1989. **13**(2): p. 127-143.
26. Shavelson, R.J. and N.M. Webb, *Generalizability Theory: A Primer*. 1991, London: SAGE Publications.
27. Feldt, L.S. and R.L. Brennan, *Reliability*, in *Educational Measurement*, R. Linn, Editor. 1989, American Council on Education and Macmillan: New York.
28. Lu, I.R.R., D.R. Thomas, and B.D. Zumbo, *Embedding IRT in Structural Equation Models: A Comparison With Regression Based on IRT Scores.* Structural Equation Modeling: A Multidisciplinary Journal, 2005. **12**(2): p. 263-277.
29. Holland, P.W. and H. Wainer, *Differential item functioning*. 1993, Hillsdale: Lawrence Erlbaum Associates.
30. Yount, K.M., K. VanderEnde, S. Zureick-Brown, T.H. Minh, S.R. Schuler, and H.T. Anh, *Measuring attitudes about women's recourse after exposure to intimate partner violence: the ATT-RECOURSE scale.* J Interpers Violence, 2014. **29**(9): p. 1579-605.
31. Hancock, G.R. and R.O. Mueller, *The reviewer's guide to quantitative methods in the social sciences*. 2010, New York: Routledge.
32. Stevens, S.S., *On the Theory of Scales of Measurement.* Science, 1946. **103**(2684): p. 677-80.
33. Baglin, J., *Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR.* Practical Assessment, Research & Evaluation, 2014. **19**(5).
34. Timmerman, M.E. and U. Lorenzo-Seva, *Dimensionality assessment of ordered polytomous items with parallel analysis.* Psychological Methods, 2011. **16**(2): p. 209-220.
35. Olsson, U., *On The Robustness Of Factor Analysis Against Crude Classification Of The Observations.* Multivariate Behavioral Research, 1979. **14**(4): p. 485-500.
36. Holgado–Tello, F.P., S. Chacón–Moscoso, I. Barbero–García, and E. Vila–Abad, *Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables.* Quality & Quantity, 2008. **44**(1): p. 153.
37. Yang, Y. and Y. Xia, *On the number of factors to retain in exploratory factor analysis for ordered categorical data.* Behavior Research Methods, 2015. **47**(3): p. 756-772.
38. Muthen, B., S. du Toit, and D. Spisic, *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes.* Psychometrika, 1997. **75**(1): p. 40-45.
39. Gorsuch, R.L., *Factor analysis.* 2015.
40. Comrey, A.L. and H.B. Lee, *A first course in factor analysis*. Second Edition ed. 1992, Hillsdale, NJ: Lawrence Erlbaum.
41. Horn, J.L., *A rationale and test for the number of factors in factor analysis.* Psychometrika, 1965. **30**(2): p. 179-185.
42. Humphreys, L.G. and R.G. Montanelli Jr, *An Investigation of the Parallel Analysis Criterion for Determining the Number of Common Factors.* Multivariate Behavioral Research, 1975. **10**(2): p. 193-205.
43. Rogers, E., *Diffusion of innovations*. Third ed. 1983, New York, New York: Macmillan Publishing Company, Inc.

PART III: COLLECTIVE EFFICACY & THE UPTAKE & INFLUENCE OF A COMMUNITY-BASED WASH PROGRAMME

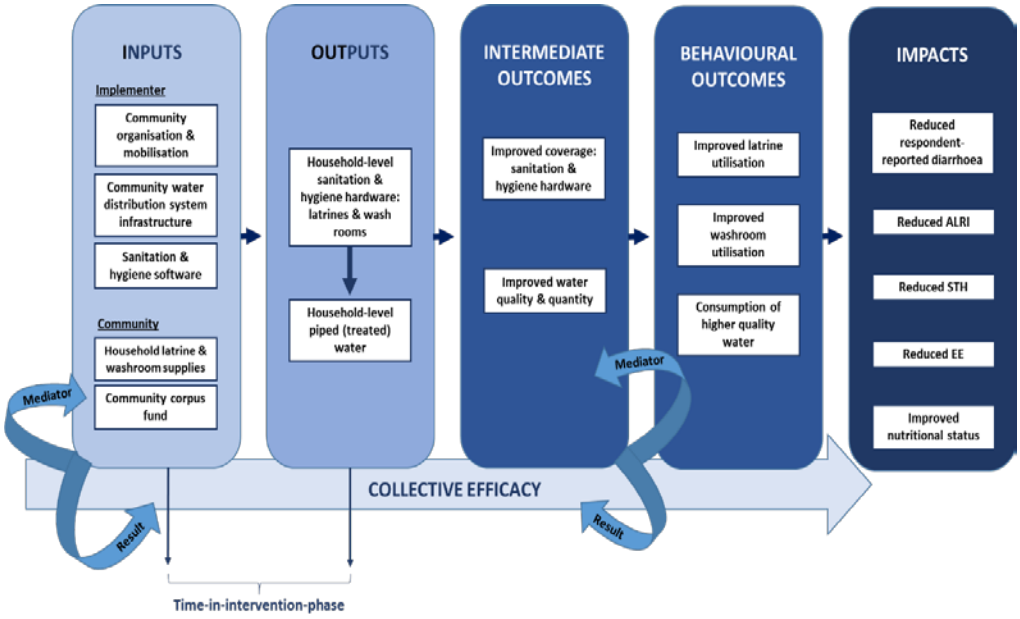
Chapter 5: Collective efficacy and the uptake and influence of community-based interventions predicated on collective action

5.1 Chapter overview

The primary purpose of developing CE scales was to employ them to assess whether and to what extent CE factors affect the uptake of community-based WASH interventions. The research paper presented in this chapter highlights work done in India to examine whether and to what extent CE-related factors are independently associated with behavioural and child nutrition outcomes that fall along the pathway to improved child health and nutrition. Given the limitations of the matched cohort study design, I was not able to assess causal effects between CE, improved behaviours, and downstream health outcomes; plans are in place to perform those analyses with *Andilaye* trial data.

To identify indicators of the MANTRA intervention that fall along the causal pathway, a simplified logic model was developed. Figure 5.1 depicts the logic model, and indicates how CE may influence intervention effectiveness via mediating behavioural outcomes and downstream health impacts.

Figure 5.1 Simplified logic model for the Gram Vikas MANTRA programme



This figure indicates that Gram Vikas leveraged a village water distribution system which piped water into each household to ensure village-level saturation of high-quality household-level sanitation and hygiene hardware (i.e., improved, on-site household toilets and enclosed, on-site bathing rooms).

5.2 Framing of research questions

To date, the WASH community's examination of the community-level, or interpersonal behavioural factors that influence uptake of WASH interventions, particularly community-based interventions, has been limited in scope, as they tend to emphasise individual-level factors [1, 2]. Therefore, my colleagues and I sought to address this gap by examining CE and its association with the uptake and influence of the MANTRA intervention. As indicated in Chapter 3, the research highlighted in this chapter is a sub-study of a larger evaluation of the Gram Vikas MANTRA water supply and sanitation programme.

5.2.1 Research Paper 3 objectives

Research Paper 3 examines whether and to what extent there is evidence of an independent association between CE-related factors, improved WASH behaviours (i.e., facilities coverage and utilisation), and downstream nutritional outcomes amongst children less than five. The objectives of this work were two-fold, to:

1. Develop and assess a theoretically-grounded, evidence-based CE measurement scale that ascertained the underlying structure of CE in the rural Odisha context; and
2. Employ the metric to obtain quantitative measures of CE that would allow for diagnostic evaluation.

5.2.2 Research questions

This CE sub-study was designed to investigate the following research questions:

RQ1. Do the factor solutions generated by our India collective efficacy scale support our theorised conceptualisation of CE (i.e., do the statistics generated by the factor solutions support our theorised construct dimensionality)?

RQ2. Does our India CE scale demonstrate construct validity?

RQ3. Are there statistically significant differences in CE factor scores between men and women, individuals with a leadership role in the village and those without?

RQ4. Is there evidence of an independent association between CE factor scores and various indicators along the causal chain?

RQ4a. Is there evidence of an association between CE factor scores and Gram Vikas MANTRA intervention status?

RQ4b. Is there evidence of an association between CE factor scores and uptake of improved WASH coverage and utilisation practices?

RQ4c. Is there evidence of an association between CE factor scores and nutritional status of children aged less than five years?

5.2.3 Related hypotheses

My colleagues and I hypothesise that CE is an important antecedent of the cooperative behaviour and collective action needed to bring about sustained adoption of improved WASH practices at the level required to realise health impacts. In the context of the research highlighted in this chapter, and the specific research questions under examination, the following hypotheses were tested:

Hypothesis 1: The statistics generated by the factor solutions support our theorised construct dimensionality.

Hypothesis 2: The India CE measurement models produced via factor analytic methods demonstrate high construct validity.

Hypothesis 3: Given their status, mobilisation, and inclusion within their communities, men have higher perceptions of behavioural control and related factors than women, leaders than non-leaders.

Hypothesis 4: Higher CE, as measured via our India scale (i.e., CE-related factor scores), is associated with: 1. MANTRA intervention exposure (i.e., CE perceptions enhanced via MANTRA); 2. higher levels of intervention uptake (e.g., higher prevalence of improved drinking water, piped on-premise; improved, on-site sanitation facilities; and enclosed bathing rooms); and 3. improved nutritional outcomes (height-for-age z-score, weight-for-age z-score, weight-for-height z-score) amongst children aged less than five years.

Hypothesis 5: Even in control villages, there is a positive association between CE-related factor scores (i.e., innate CE) and prevalence of improved WASH facilities coverage and utilisation.

Research Paper 3, included in this chapter, presents details related to these specific inquiries. Of note, unlike the analyses presented in Research Paper 2, gender-specific models were not generated for this inquiry. Instead, one CE measurement model was fitted, and differential item functioning between men and women, and leaders and non-leaders was assessed. Section 5.3 below provides further detail regarding descriptive analyses that are not included in Research Paper 3, but may provide the reader with additional information regarding the behavioural and nutritional outcomes presented therein.

5.3 Presentation of descriptive analyses not highlighted in Research Paper 3

The outcomes of interest for Research Paper 3 include five behavioural outcomes that represent WASH facilities coverage and utilisation, and three nutritional outcomes amongst children aged less than five:

WASH facilities coverage

1. Access to improved drinking water source, piped on-premise;
2. Access to an improved, on-site sanitation facility;
3. Access to an enclosed, on-site bathing room;

WASH facilities utilisation

4. Household-level utilisation of an improved, on-site sanitation facility;
5. Utilisation of enclosed, on-site bathing room;

Nutritional outcomes amongst children aged less than five years

6. Height-for-age z-score (HAZ);
7. Weight-for-age z-score (WAZ); and
8. Weight-for-height z-score (WHZ).

These coverage and utilisation outcomes serve as indicators of behavioural outcomes linked to the key components of Gram Vikas' MANTRA intervention. Details related to measurement methods for each outcome are presented in Research Paper 3. Information related to each outcome is presented below.

A total of 1,734 households provided data from this CE study; 803 from intervention villages, 931 from control villages. Details regarding respondent, household, and village characteristics are presented in Research Paper 3.

Access to improved drinking water source, piped on-premise

A total of 608 households in this sub-study had access to an improved drinking water source, piped on-premise. While only 67% (n=535) of the 803 intervention households had access to an improved drinking water source, piped on-premise, this was significantly higher than control households, only 8% (n=73) of which had access.

Access to an improved, on-site sanitation facility

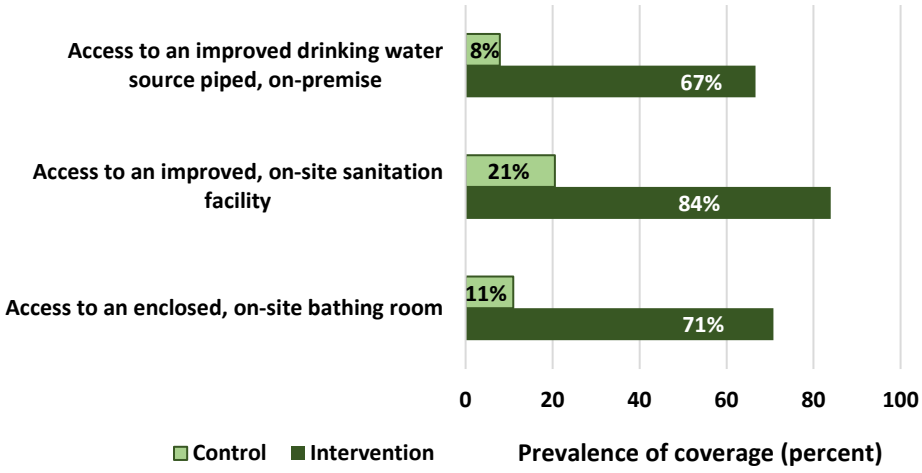
A total of 865 households had access to an improved, on-site sanitation facility. Access was high in intervention villages, as 84% (n=674) of households in this study arm had access to an improved facility. Access in control villages was considerably lower in control villages, as only 21% (n=191) of the households in the study arm in this sub-study had access to an improved, on-site sanitation facility.

Access to an enclosed, on-site bathing room

A total of 670 households had access to an enclosed, on-site bathing room. As with improved sanitation facilities access, access to an enclosed, on-site bathing room was considerably higher in intervention villages compared to control villages, with 71% (n=568) of intervention households having access compared to 11% (n=102) of control households.

Figure 5.2 visualises WASH facilities coverage outcomes, by study arm.

Figure 5.2 WASH facilities coverage outcomes



Sanitation facilities utilisation and other defecation practices

A total of 848 respondents (49%)[1] reported that household members usually use no sanitation facility (i.e., openly defecate), 34 of those indicated open defecation within the household compound. Of those 848 respondents, 86% resided in a control village, 15% resided in an intervention village. Given only 816 respondents reported not having access to a sanitation facility, this means that 32 households with access (4%) do not usually use the facility. Household-level utilisation patterns are described in more detail below.

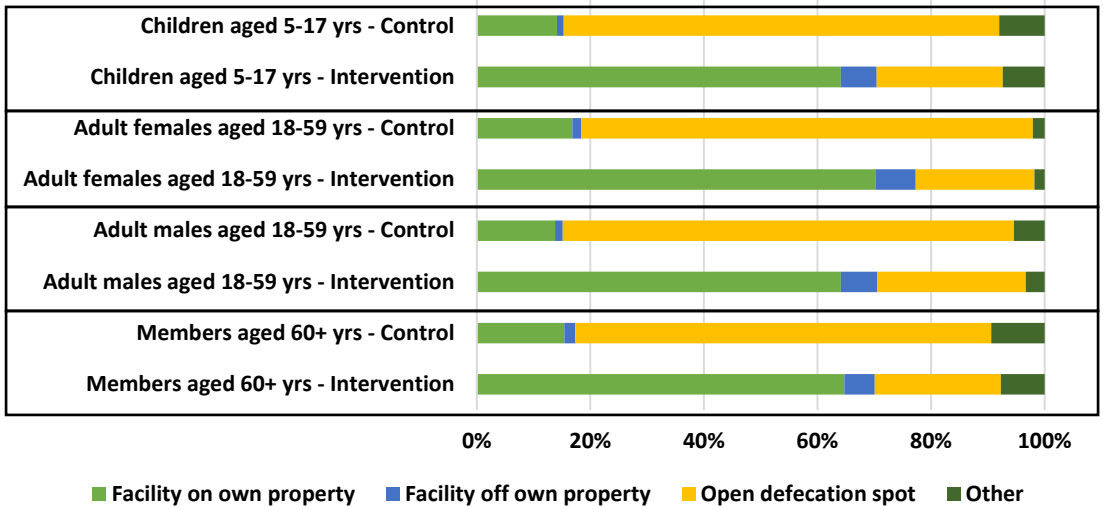
Household sanitation utilisation profile – usual defecation place

Figure 5.3 depicts respondent-reported usual defecation location for various household user groups, according to age and gender. A total of 803 households had one or more members who were aged 60 years or older (i.e., 'older adult' household members; n= 388 intervention, n=415 control); the remaining 926 CE study households with user cohort data reported having no members in this age cohort. Amongst 'older adult' members, 390 (49%) usually openly defecate (86 intervention [22%], 304 control [73%]); 315 (39%) usually use the toilet on the household's own property for defecation (251 intervention [65%], 64 control [15%]); 69 [9%] usually use some other

[1] These respondents included those with access to a sanitation facility.

sanitation facility (30 intervention [8%], 39 control [9%]; and 29 (4%) usually use a toilet located off the household's own property (21 intervention [5%], 8 control [2%]).

Figure 5.3 Usual defecation location, by cohort & study arm



A total of 1,645 CE households had one or more adult male aged 18-59 years; the remaining 89 CE study households had no male household members in this age range. Unfortunately, defecation practice data were missing for 77 of these members. Of the 1,568 households with a member in this cohort and for whom defecation practice data were obtained (n=729 intervention, n=839 control), 858 (55%) usually openly defecate (191 intervention [26%], 667 control [80%]); 583 (37%) usually use the toilet on the household's own property for defecation (467 [64%] intervention, 116 [14%] control); 69 (4%) usually use some other sanitation facility (24 [3%] intervention, 45 [5%] control); and 58 (4%) usually use a toilet located off the household's own property (47 [6%] intervention, 11 [1%] control).

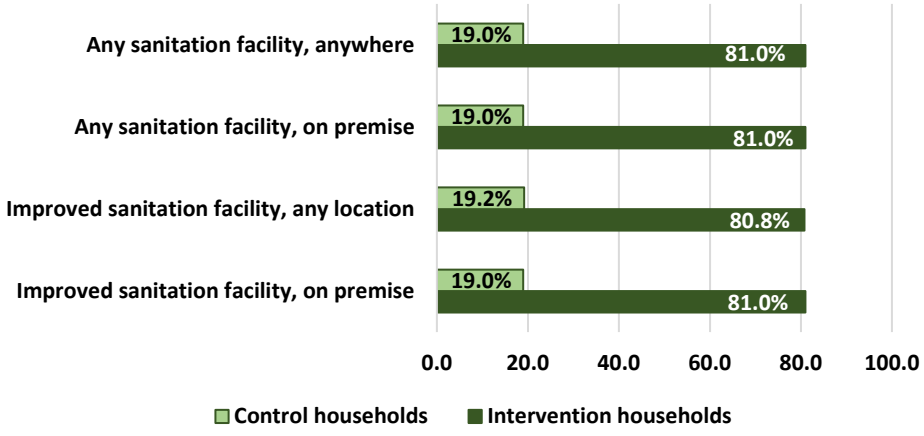
A total of 1,724 households had one or more adult female aged 18-59 years; the remaining 10 CE study households reported having no female household members in this age range. Defecation practice data were missing for eight of these individuals. Of the 1,716 households with a member in this cohort and for whom defecation practice data were obtained (n=793 intervention, n=923 control); 900 (52%) usually openly defecate (166 [21%] intervention, 734 [80%] control); 712 (42%) usually use the toilet on the household's own property for defecation (557 [70%] intervention, 155 [17%] control); 71 (4%) usually use a toilet located off the household's own property (56 [7%]

intervention, 15 [2%] control); 33 (2%) usually use some other sanitation facility (14 [2%] intervention, 19 [2%] control).

A total of 1,037 households had one or more child aged 5-17 years; the remaining 697 CE study households reported having no members in this age group. No defecation practice data were obtained for 11 of these individuals. Of the remaining 1,026 individuals, 536 (52%) usually openly defecate (102 [22%] intervention, 434 [77%] control); 375 (37%) usually use the toilet on the household's own property for defecation (295 [64%] intervention, 80 [14%] control); 79 (8%) usually use some other sanitation facility (34 [7%] intervention, 45 [8%] control); 36 [4%] usually use a toilet located off the household's own property (29 [6%] intervention, 7 [1%] control).

After assessing defecation practices on a household-level, across all user groups, a total of 641 CE study households (37%) were identified as having all household members five years and older usually defecating in any sanitation facility in any location. When disaggregating by study arm, 81% (n=519) of these households resided in intervention villages (Figure 5.4).

Figure 5.4. Household sanitation use profile - Usual defecation place, all members 5+ yrs



Utilisation of enclosed, on-site bathing room facility

Of the 670 households with a bathing room facility, 622 reported that at least one woman in the household used the facility at least sometimes (530 intervention, 92 control).

Nutritional outcomes

Kernel density plots for the three nutritional outcomes of interest, HAZ, WAZ, and WHZ are presented in Figure 5.5. Mean HAZ amongst intervention children aged less than five years was -1.47 (SD 1.18), while mean HAZ amongst control children in the same age cohort was -1.74 (SD 1.12). These results represent a statistically significant difference between children from intervention and control villages at α=0.05 (p=0.015). This suggests that children from intervention villages were less stunted, or less prone to chronic undernutrition due to prolonged illness and/or food deprivation.

Mean WAZ amongst intervention children aged less than five years was -1.35 (SD 1.14), while mean WAZ amongst control children in the same age cohort was -1.57 (SD 1.09). These results represent a statistically significant difference between children from intervention and control villages at α=0.05 (p=0.050). WAZ represents a composite indicator reflecting both acute and chronic undernutrition. Therefore, these findings suggest that children from intervention villages were less underweight than children from control villages.

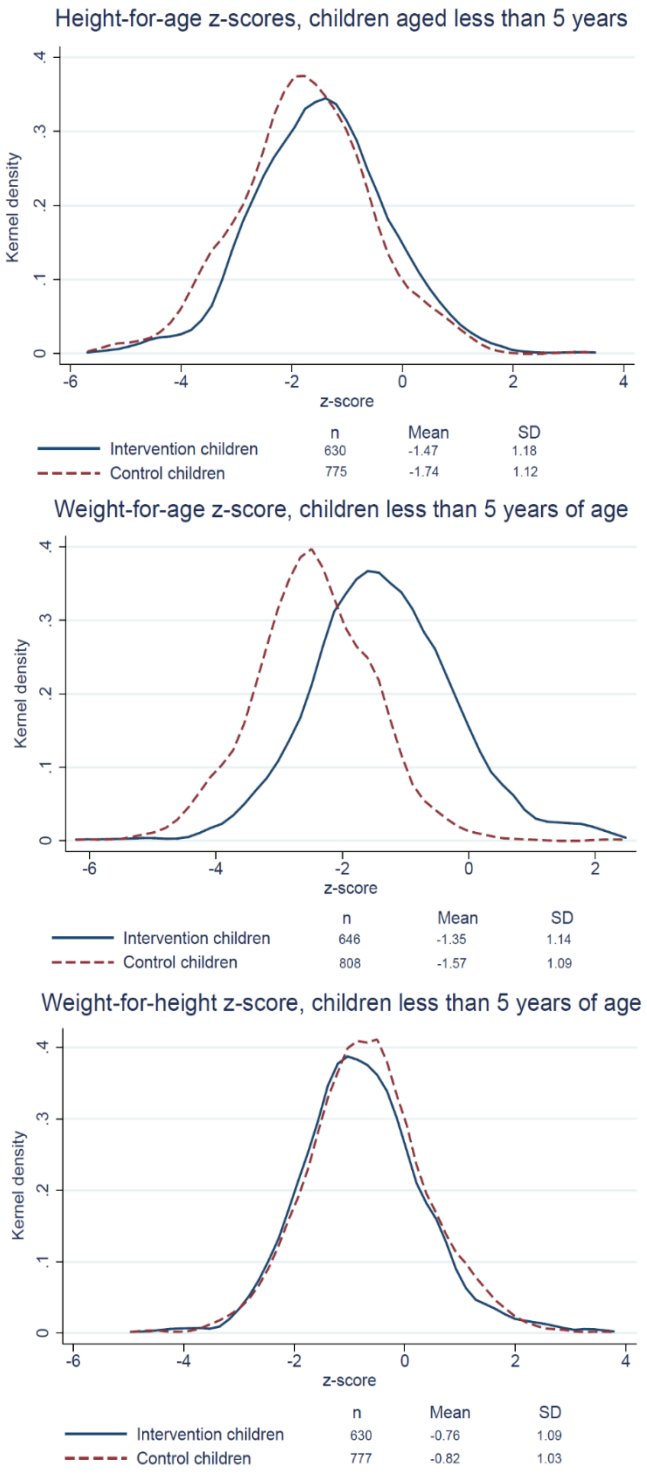


Figure 5.5. Nutritional status of children aged less than 5 years, by study arm – CE study households

Mean WHZ amongst intervention children aged less than five years was -0.76 (SD 1.09), while mean WHZ amongst control children in the same age cohort was -0.82 (SD 1.03). These results do not indicate a statistically significant difference between children from intervention and control villages (p=0.473). These findings suggest that similar amounts of acute undernutrition, or wasting, are observed in intervention and control villages.

5.4 Collective efficacy and the Gram Vikas MANTRA intervention

5.4.1 Preamble to Research Paper 3

Research Paper 3 presents details regarding the modelling of key behavioural and nutritional outcomes of interest (outlined above) and CE factors. CE-related factor scores were entered into models along with other covariate and background variables to determine whether there was evidence of independent associations between CE factors, uptake of improved WASH practices, and downstream nutritional outcomes. Gender and intervention status-based stratified analyses were carried out. Assessing results within the intervention arm allowed for an investigation of whether there was evidence of independent associations between CE factors and indicators of intervention uptake (e.g., behavioural and nutritional outcomes). Given CE is an innate characteristic operating amongst any collective or group of people, it was also fruitful to investigate whether there was evidence of associations between CE factors and uptake of improved WASH behaviours in control villages. Exploring whether there was evidence of these relationships in control villages provided counterfactual assessments while also allowing for an examination of innate properties of CE where no community-based WASH interventions were operating.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
www.lshtm.ac.uk

Registry
T: +44(0)20 7299 4646
F: +44(0)20 7299 4656
E: registry@lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

RESEARCH PAPER COVER SHEET

PLEASE NOTE THAT A COVER SHEET MUST BE COMPLETED FOR EACH RESEARCH PAPER INCLUDED IN A THESIS.

SECTION A – Student Details

Student	Maryann Delea
Principal Supervisor	Dr. Thomas Clasen
Thesis Title	Social constructs, behaviour change, and the uptake of community-based WASH interventions: Metrics and analytical approaches for measuring collective efficacy and other social constructs

If the Research Paper has previously been published please complete Section B, if not please move to Section C

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

**If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.*

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMJ Global Health
Please list the paper's authors in the intended authorship order:	Maryann G. Delea, Gloria D. Sclar, Corey L. Nagel, Sheela S. Sinharoy, Heather Reese, Thomas F. Clasen
Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived of and designed this study. I worked with team members to develop the survey instrument and conduct cognitive interviews during the formative phase. I analysed the data and wrote the first draft of the paper. Co-authors provided contributions.

Student Signature: [redacted] Date: 16/10/2018

Supervisor Signature: [redacted] Date: 10-18-2018

5.5 Research Paper 3: Collective efficacy and the adoption and influence of a community-based water supply and sanitation programme: a cross-sectional study in Odisha, India

Maryann G. Delea [1,2*], Gloria D. Sclar [2], Corey L. Nagel [3], Sheela S. Sinharoy [2], Heather Reese [2], Thomas F. Clasen [1,2]

[1] Department of Disease Control, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom;

[2] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States of America;

[3] College of Nursing, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America;

Status: Prepared for submission to *BMJ Global Health*

ABSTRACT

Introduction

Null findings from water, sanitation, and hygiene (WASH) trials highlight the need to elucidate factors affecting the adoption and impact of WASH interventions. We hypothesise that uptake of improved WASH behaviours and downstream health gains may be associated with collective efficacy (CE) — a group's perceived ability and autonomy to come together and execute actions related to common goals.

Methods

We carried out this cross-sectional study within a matched cohort study conducted in rural Odisha, India comparing households from villages that agreed to implementation of a community water supply and household latrine programme with households from control villages not enrolled in the programme. We employed mixed-methods to develop and validate a CE scale, and used multilevel

modelling to examine evidence of associations between CE factors and intervention status and uptake, and nutritional outcomes amongst children under five.

Results

Compared to controls, intervention households had higher CE factor scores for village leadership (β=0.16, 95% CI: 0.08, 0.25) and agency (β=0.08, 95% CI: 0.01, 0.15), and lower scores for social disorder, conditional on education. Prevalence of improved WASH behaviours was associated with CE factors: improved water piped on-premise was associated with village leadership (aPR=2.25, 95% CI: 1.12, 4.53); improved on-site toilets with social disorder (aPR=0.90, 95% CI: 0.84, 0.97); enclosed bathing rooms with social response (aPR=1.12, 95% CI: 1.02, 1.23); and utilisation of improved sanitation facilities by all family members with agency (aPR=1.17, 95% CI: 1.00, 1.37) and village leadership (aPR=3.86, 95% CI: 1.67, 8.97). Agency, social response, and social disorder factors were associated with nutritional outcomes.

Conclusion

This study demonstrated the ability of our CE scale to pinpoint specific factors that contribute toward or prevent the uptake of improved WASH behaviours. Future experimental studies can help determine whether and to what extent CE is an antecedent of collective action and community-based intervention effectiveness.

SUMMARY BOX

What is already known?

- Collective efficacy is an interpersonal behavioural factor that pertains to a group's ability and autonomy to catalyse and maintain change.
- Collective efficacy influences collective goal setting; group performance, including the amount of effort and resolve the group will expend; and ultimate goal attainment.
- To date, the WASH community's examination of social constructs has been limited in scope, and has not included thorough investigations into relationships between collective efficacy and WASH behaviours.

What are the new findings?

- Collective efficacy-related factors, such as village leadership, social disorder, and agency were independently associated with village participation in a community-based WASH programme that focuses on collective action, social inclusion, and social equity.
- Village leadership, social disorder, social response, and agency factor scores were associated with prevalence of improved WASH facilities coverage and utilisation.
- Agency and village leadership factor scores from adult household members were independently associated with downstream nutritional outcomes amongst children aged less than five years.

What do the findings imply?

- Assessing collective efficacy may help address persistent knowledge gaps related to interpersonal factors contributing to poor uptake of community-based WASH interventions (e.g., failure to address behavioural control perspectives at individual and collective levels), and less than expected health gains.
- Collective efficacy should be incorporated into the targeting, design, implementation, and evaluation of community-based interventions predicated on collective action.
- Our findings imply that future experimental studies should investigate whether collective efficacy is an important interpersonal behavioural antecedent of collective action and effective community-based interventions.

INTRODUCTION

There is accumulating evidence regarding the roles various psychosocial, contextual, and technological factors play in the sustained adoption of improved WASH behaviours and practices [3-6]. Despite existing and emerging literature on the nuances of various behavioural factors and their interactions with other contextual and technological factors, the WASH community continues to overlook some key interpersonal psychosocial influencers of behavioural change and maintenance. The sector also continues to focus largely on examining and addressing individual and household-level factors while higher-level interpersonal factors may be stronger influencers of action at collective levels, the level at which WASH programmes are typically delivered and change may be required to realise health gains [7-9]. Underestimating or overlooking these behavioural factors promotes a “black box” mentality that supports oversimplified conceptualisations of causal

chains. Failing to unpack these issues and collect data on critical mechanisms through which sustained behavioural adoption may occur prevents a more thorough examination of the behavioural antecedents of cooperative behaviour and collective action. Such investigations may be needed to help answer why WASH programmes and interventions are not bringing about desired health impacts.

When examinations into interpersonal factors of WASH uptake have been carried out, they have largely focused on social capital and social norms. Although efficacy, or beliefs regarding either one's own (self-efficacy) or a group's (collective efficacy) ability and autonomy to set, pursue, and achieve goals, is dominant in social theory as an important behavioural factor, it remains largely unstudied within the WASH sector. Failures in the uptake and maintenance of improved WASH behaviours may, at least partially, be ascribed to poor inclusion of efficacy-related factors in the targeting, design, and implementation of related interventions.

Collective efficacy (CE) is a latent social construct pertaining to factors that facilitate a group's perceived ability and autonomy to come together and execute actions related to a common goal [10, 11]. Perceptions regarding CE have been shown to not only influence group goal selection and performance [12-14], but also the amount of effort and resolve a collective will spend working toward those goals [10]. These basic properties of CE are what make it a more appropriate social construct to consider when assessing interpersonal behavioural factors of collective action than social capital or social norms alone.

Gram Vikas MANTRA approach to improving sanitation and hygiene

Gram Vikas is a local non-governmental organisation based in Odisha, India. The organisation's Movement for Action Network for Transformation of Rural Areas (MANTRA) programme was being evaluated during our study of CE. The MANTRA programme approach ensures village-wide saturation of high-quality, on-site sanitation and bathing facilities by leveraging a common good – a village-level, piped water distribution system. Gram Vikas provided technical assistance for the development of the water distribution system that is piped throughout the village, with each household compound receiving three taps. As a condition to providing the water supply, Gram Vikas requires that each household install: 1) a high-quality improved toilet, with one household tap piped into the facility; and 2) an enclosed bathing room, with an additional tap or shower in the facility. The cost of these facilities is supported by government sanitation subsidies. The presence

of improved water piped on premise, an improved on-site sanitation facility with water tap, and bathing room with water tap serves to fully interrupt the traditional setting in which open defecation typically occurs – off-site water sources that allow for post-defecation ritual cleansing. Presence of sanitation and bathing facilities with water access ensures that ritual cleansing related to defecation and menstrual hygiene management can take place privately, and on the household's own compound.

Gram Vikas will not initiate village enrollment into MANTRA before a village demonstrates its unanimous commitment to the programme (see Supplemental Material for further details). Furthermore, MANTRA programme requirements stipulate that the water distribution system should not be turned on or turned over to the village until every household has installed its own improved, on-site sanitation and hygiene block. Consequently, MANTRA's WASH facilities coverage and utilisation targets are predicated on the villager's cooperative behaviour and collective action. As a result of Gram Vikas' programme targeting approach, which limits programme enrollment (i.e., exposure) to villages demonstrating cooperative behaviour during the pre-implementation phase, a village's innate, or underlying CE likely determined its exposure to the MANTRA intervention. In this sense, CE may influence intervention exposure. At the same time, MANTRA's programming approach, intervention components, and activities, may serve to enhance CE-related factors during the implementation phase. Therefore, CE not only influences exposure to the intervention, but likely falls along the pathway from intervention exposure to behavioural and nutritional outcomes. Figure 1 visually depicts how we hypothesise CE may influence intervention uptake and impact through behavioural and downstream health outcomes.

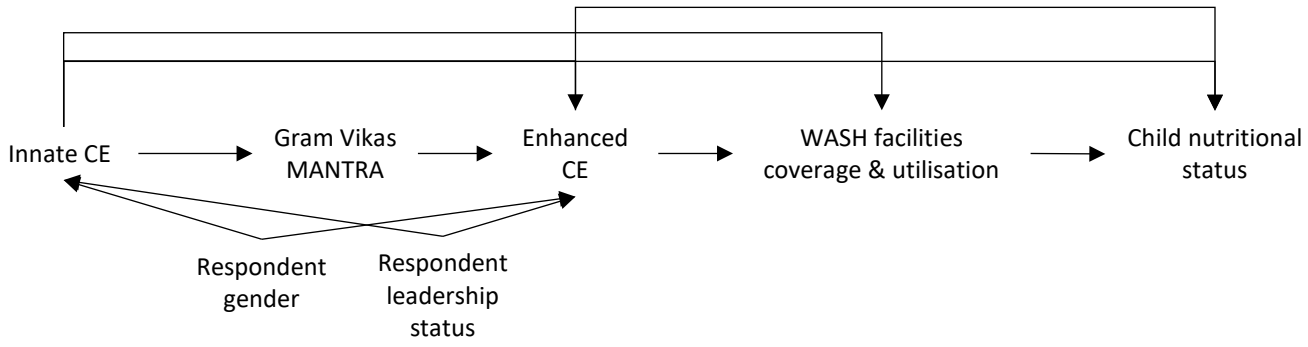


Figure 1. Hypothesised pathway from CE to downstream behavioural & nutritional outcomes

Study objectives and research questions

To date, the WASH community's examination of the interpersonal behavioural factors that influence uptake of WASH interventions, particularly community-based interventions, has been limited in scope. In this study, we sought to address this gap by examining CE, its role in the uptake of improved WASH behaviours, and its association with downstream nutritional status. More specifically, we were interested in exploring evidence of independent associations between CE factor scores and WASH facilities coverage and utilisation, nutritional outcomes amongst children aged less than five years, and intervention status more generally. We aimed to develop a theoretically-grounded and evidence-based CE scale for the rural Odisha context, and use it to examine evidence of independent associations between CE and indicators along the causal chain.

METHODS

Study setting and design

This assessment of CE and its association with behavioural outcomes and nutritional status was a sub-study nested within a matched cohort study evaluating effect of the Gram Vikas MANTRA on health. Details related to intervention and the matched cohort study are described elsewhere [15, 16]. The MANTRA programme and its interventions were developed by Gram Vikas [15]. MANTRA's interventions seek to improve coverage and utilisation of: 1) a household pour-flush toilet with dual soak-away pits, 2) an enclosed bathing room, and 3) household piped water connections in the toilet, bathing room, and elsewhere in the household compound (e.g., kitchen) [15]. For a village to receive the piped water distribution system, every household has to construct its own toilet and bathing room, in addition to other programme requirements. Gram Vikas assists with the development of a piped water system, which becomes operational once every household completes toilet construction. The village is responsible for on-going operation and maintenance costs.

This CE study was conducted in 45 intervention and 45 control villages in Ganjam and Gajapati Districts, eastern Odisha, India. These districts represent the setting in which Gram Vikas developed and refined its MANTRA intervention. The majority of people who live in these districts reside in rural areas. Pre-intervention data indicate that open defecation was highly prevalent in these districts [15]. Intervention villages included in the study were selected from a list of more than 1000

villages, provided by Gram Vikas, in which the MANTRA programme had been implemented at least three years previously; intervention villages were matched to control villages on various demographic and geographic criteria. Gram Vikas reports that a majority of villages that are approached about the programme do not ultimately agree to participate. While control villages were matched to intervention villages, the fact that the intervention villages had agreed to participate in and actually completed the programme may render them different from the controls on unmeasured factors, such as CE factors.

Given Gram Vikas did not obtain any pre-intervention measures of CE or related factors, we designed this study as a post-only, cross-sectional assessment of CE and its associations with behavioural and nutritional outcomes. While this post-only design does not allow us to infer causality from associations, or to determine the directionality of relationships (i.e., intervention contributing to CE or CE contributing to intervention), the presence of associations could motivate future experimental studies to explore causality.

Sampling methodology

The 90 villages enrolled in the matched cohort study comprised the study population for this research. Villages that Gram Vikas formally engaged for their MANTRA intervention, with intervention start dates between 2003-2006, were eligible for inclusion in the matched cohort's intervention arm. Forty-five villages were randomly selected for study enrollment amongst those meeting eligibility criteria. A multi-step restriction, matching, and exclusion process was used to match 45 control villages to the randomly selected intervention villages, as detailed elsewhere [15]. Households with a child under five years of age were eligible for enrollment. All households enrolled in the matched cohort study were targeted for this CE study. CE data were collected from all households in which a capable and willing adult (i.e., individual aged 18 years or older) was available and provided consent to participate in the CE study.

Formative work

We employed a five-step, sequential exploratory and confirmatory approach [17] to develop, refine, and validate a theoretically-grounded and evidence-based CE measurement scale [18]. Full details regarding our CE scale development and validation approach are published elsewhere, though highlight subsequent CE scale development that was carried out in rural Ethiopia [19]. In summary, for the purposes of CE scale development in India, we established a hypothesised

framework that we transitioned into a draft CE survey. We performed a series of in-depth cognitive interviews to gauge participants' understanding of the meaning of survey prompts (i.e., assessed face validity) and appropriateness of response selection. Findings from these in-depth cognitive interviews informed survey refinements, which resulted in a 36-item survey (Table S2).

Data collection

Administration of the CE survey took place during April-September 2016 (summer/rainy seasons in Odisha, India). One adult (i.e., individual aged 18 years or older) per household was randomly selected, using a random number generator, to respond to the CE survey. This strategy allowed for the collection of data on CE perspectives from both men and women. WASH facilities coverage and utilisation, and nutritional status data captured during the same rounds of data collection were collected from the respondents of those household surveys. In the event that no relevant data were collected during the same rounds, data collected from the household during other rounds of data collection were used.

Measures

Collective efficacy

Our refined, 36-item CE survey was comprised of group-referent statements about interpersonal and ecological aspects of the respondent's community that related to CE, and self-referent statements about the respondent's own sense of self, agency, autonomy, and level of engagement within his/her community. Given our intention to develop a CE metric for integrated programmes, no WASH-specific indicators were included in the survey. These items tapped to our hypothesised CE framework. We trained fieldworkers to administer the CE survey by reading each survey item, followed by each of the five response options (i.e., completely disagree, mildly disagree, neither agree nor disagree, mildly agree, completely agree). Respondents were asked to rate their agreement with the statement on this five-point Likert-type scale.

WASH facilities coverage and utilisation

Fieldworkers captured data on access to and utilisation of WASH facilities through standardised survey prompts employed in Demographic and Health Surveys (DHS). These coverage and utilisation outcomes serve as indicators of behavioural outcomes linked to the key components of Gram Vikas' MANTRA intervention. We used WHO/UNICEF Join Monitoring Programme for Water Supply, Sanitation, and Hygiene (JMP) definitions to classify coverage outcomes. Household-level

utilisation of an on-site, improved sanitation facility was ascertained by querying respondents on the usual place of defecation for its household members, by cohort – age for younger and older members, age and gender for adults aged 18-59 years (i.e., children 5-17 years, adult females aged 18-59 years, adult males aged 18-59 years, adult members aged 60 years and older). We assessed usual place of defecation across the cohorts to generate a household-level utilisation profile, and dichotomised the outcome according to use by all user groups. In order for a household to be categorised as one in which all members usually defecate in an on-site improved sanitation facility, the usual defecation place for all user groups must have been a facility on the household's own property. In addition, that facility must have been classified as an improved sanitation facility, per JMP definition, located on-site. Utilisation of an on-site, enclosed bathing room was obtained from the female respondents targeted for the cohort study's fourth round of data collection. Those who indicated having access to an on-site, enclosed bathing room were inquired about the frequency of their use of that facility (i.e., whether it was used for bathing always, sometimes, or never). We later dichotomised this outcome to determine exclusive use.

Nutritional outcomes amongst children aged less than five years

We collected anthropometric data on children under the age of five years using standardised methods developed by WHO [20, 21]. Details regarding the materials and procedures used for obtaining weight and recumbent length (for children aged less than two) or standing height (for children aged two to five years) are published elsewhere [15]. WHO reference standards were then used to translate height and weight measurements into: 1) height-for-age z-score (HAZ), an indicator of chronic undernutrition resulting from prolonged food deprivation and/or illness; 2) weight-for-height z-score (WHZ), an indicator of acute undernutrition resulting from more recent food deprivation or illness; and 3) weight-for-age z-score (WAZ), a composite indicator that reflects both acute and chronic undernutrition [22].

Measures on exclusive breastfeeding and minimum dietary diversity were generated in accordance with WHO/UNICEF guidelines [23]. Exclusive breastfeeding was ascertained for all infants 0-5 months of age. Minimum dietary diversity was obtained for all children 6-59 months of age.

Analytical methods

Psychometric analyses of CE data

Initial data cleaning and descriptive analyses were performed in Stata (version 15.0 StataCorp, College Station, TX, USA). We then employed a factor analytical approach, consisting of exploratory and confirmatory factor analyses, to examine the underlying structure of factors related to collective efficacy, and identify a data-derived CE measurement model for rural Odisha [24]. We subsequently performed Multiple Indicators Multiple Causes (MIMIC) modelling to determine whether various groups of people with the same underlying sense of CE factors scored items differently (e.g., men/women, leaders/non-leaders) [25]. These psychometric analyses were conducted in Mplus software (version 8 Muthén & Muthén, Los Angeles, CA, USA). Given the psychometric analyses are not the focus of this paper, we present further details related to these methods and results in the Supplemental Material.

From our final refined CE measurement model, we generated CE factor scores for each respondent by summing his/her responses across all items in each factor (i.e., 1=completely disagree, 2=mildly disagree, 3=neither agree nor disagree [neutral], 4=mildly agree, 5=completely agree), and dividing that sum by the number of items tapping to the factor [26]. We calculated weighted average factor scores for each CE factor, for which a weight that was equivalent to the item's factor loading was applied to each item score prior to the generation of the average (weighted) factor score.

Multi-level modelling of CE and behavioural and nutritional outcomes

Subsequent to univariate and bivariate analyses, we performed multivariate analyses, in three phases. During each phase, we employed random-intercept mixed effects model frameworks to account for non-independence of observations within village clusters (phases I and II), and within village clusters and households (phase III). During all analytical phases, we employed a backward step-wise elimination modelling approach, using an exit criterion of $p>0.05$ to fit each model. Explanatory variables demonstrating significant associations with outcomes of interest on bivariate analyses at a 10% significance level, per Wald test, were considered for entry into multivariate hierarchical models. Other non-statistical substantive justification supported our inclusion of all CE factor indicators in all multivariate models, regardless of whether they were statistically significant at a 10% level on bivariate analyses.

During phase I, we examined whether there were differences between CE factor scores across study arms and between genders and respondents with leadership status vs. those without. We then examined whether there was evidence of independent associations between CE factor scores and

intervention status, gender, and leadership status (i.e., exposure to the Gram Vikas MANTRA intervention) using maximum likelihood estimation via mixed effects linear regression with robust standard errors.

During phase II, we examined whether there was evidence of independent associations between CE factor scores and behavioural uptake of improved WASH practices, as manifest by installation of water, sanitation, and hygiene facilities, and utilisation thereof. Given CE may have dictated exposure to the MANTRA intervention, but may also have been enhanced by it, we stratified these analyses by study arm and respondent gender, as indicated by results from our phase I analyses. We employed a two-level random intercepts model framework for phase I and II to accommodate the hierarchical data structure, and adjust for village and household/respondent-level covariates. We used mixed effects Poisson regression with robust variance estimation to produce prevalence ratios for WASH facilities coverage and utilisation outcomes.

During phase III, we examined whether there was evidence of independent associations between CE factor scores and down-stream nutritional outcomes amongst children aged less than five years. Given nutritional data were collected on all children in the household aged less than five years, we employed a three-level random intercept model framework for these analyses. We employed maximum likelihood estimation via mixed effects linear regression with robust standard errors for continuous nutritional indicators (i.e., HAZ, WHZ, WAZ).

Ethical approvals

The Gram Vikas MANTRA matched cohort study and its sub-studies received ethical approval from the London School of Hygiene and Tropical Medicine's Observational/Interventions Research Ethics Committee (Ref 9071), and the Kalinga Institute of Medical Sciences at KIIT University, Bhubaneswar, India (KIMS/KIIT/IEC/053/2015) ethics committee. The study was registered on ClinicalTrials.gov (NCT02441699).

RESULTS

Study sample characteristics

A total of 1,802 eligible respondents were available and provided consent to participate in the CE survey, reflecting a 75% response rate amongst all households enrolled in the larger matched cohort study. Forty-seven observations were dropped from the analytical sample due to incomplete

or duplicative identifiers, or missing gender or age data that would impede appropriate designation for sub-group analyses. An additional 21 respondents indicated that they either did not currently live in the village, or resided there less than four days per week. The final CE analytical sample contained 1,734 observations; 1,123 from women, 611 from men. CE data were obtained on 803 respondents living in intervention villages, and 931 respondents living in control villages. A total of 2,086 children aged less than five years from 1,481 households provided data on nutritional outcomes. See Table 1 for further details.

Association between CE factor scores and intervention status, respondent's gender & leadership status

After item reduction carried out via our factor analytic approach, the CE scale we developed and validated reflected a 25-item, six-factor measurement model that included the following factors: social response, social disorder, agency, village leadership, social networks, and social attachment. These factors correspond to behavioural control, informal social control, social cohesion, and cognitive and structural social capital domains. The final CE measurement model demonstrated acceptable model fit ($\chi 2$:df ratio = 2.379; Root Mean Square Error of Approximation [90% CI] = 0.040 [0.0036-0.044]; Comparative Fit Index = 0.924; Tucker-Lewis Index = 0.913), with all items loading significantly and saliently on a single factor. See Supplemental Material for further details regarding results from our psychometric analyses.

CE factor scores differed by intervention status and respondent's gender and village leadership status (Figure 2). Below, we summarise results of our multivariate analyses, highlighting variables that were significantly and independently associated with each factor score.

Table 1. Study sample characteristics,* behavioural and nutritional outcomes of interest

	Aggregate (N=1,734)	Intervention			Control			I/C difference
		All (n=803)	Women † (n=542)	Men † (n=261)	All (n=931)	Women † (n=581)	Men † (n=350)	
	n (%)	n (%)			n (%)			p-value ‡
Respondent-reported characteristic								
Respondent's age - median (IQR)	30 (25-43)	30 (25-45)	28 (25-38)	35 (30-55)	30 (25-40)	28 (25-35)	35 (30-51)	0.735
Number of years respondent has lived in village - median (IQR)	23 (9-36)	24 (8-37)	12 (7-27)	35 (28-46)	23 (10-35)	12 (7-26)	34 (27-45)	0.589
Social response factor score (weighted) - median (IQR)	3.3 (2.7-3.5)	3.3 (2.8-3.5)	3.3 (2.7-3.5)	3.3 (3.1-3.5)	3.3 (2.7-3.5)	3.1 (2.5-3.5)	3.3 (2.8-3.5)	0.164
Social disorder factor score (weighted) - median (IQR)	1.7 (1.3-2.1)	1.6 (1.1-2.0)	1.7 (1.2-2.0)	1.6 (1.0-2.0)	1.8 (1.3-2.1)	1.9 (1.4-2.1)	1.7 (1.2-2.1)	0.001***
Agency factor score (weighted) - median (IQR)	2.7 (2.1-3.2)	2.7 (2.1-3.2)	2.7 (2.0-3.1)	2.9 (2.6-3.4)	2.7 (2.1-3.2)	2.4 (1.9-2.8)	2.8 (2.6-3.4)	0.057
Village leadership factor score (weighted) - median (IQR)	2.6 (2.0-2.9)	2.7 (2.1-3.0)	2.7 (2.2-3.0)	2.7 (2.1-3.0)	2.5 (1.9-2.9)	2.3 (1.9-2.8)	2.6 (1.9-3.0)	<0.001***
Social networks factor score (weighted) - median (IQR)	4.5 (4.0-4.5)	4.5 (4.0-4.5)	4.5 (4.0-4.5)	4.5 (4.0-4.5)	4.5 (3.6-4.5)	4.5 (3.6-4.5)	4.5 (4.0-4.5)	0.134
Social attachment factor score (weighted) - median (IQR)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	4.3 (4.3-4.3)	0.486
Respondent holds a leadership role in the village	65 (4)	39 (5)	22 (4)	17 (7)	26 (3)	13 (2)	13 (4)	0.025**
Respondent's highest educational attainment								0.229
No formal education	655 (38)	294 (36)	229 (42)	65 (25)	361 (39)	264 (45)	97 (28)	
At least some primary (up to year 4)	216 (12)	83 (10)	42 (8)	41 (16)	133 (14)	64 (11)	69 (20)	
At least some secondary (up to year 10)	783 (45)	387 (48)	256 (47)	131 (50)	396 (43)	239 (41)	157 (45)	
Higher than secondary	80 (5)	39 (5)	15 (3)	24 (9)	41 (4)	14 (2)	27 (8)	
Household-level characteristics								
Number of household members - median (IQR)	7 (6-8)	7 (5-8)	6 (5-8)	7 (6-9)	7 (6-8)	6 (5-8)	7 (6-8)	0.585
Wealth quintile §								0.011**
Highest (i.e., richest)	339 (20)	189 (24)	120 (22)	69 (26)	150 (16)	97 (17)	53 (15)	
Fourth	353 (20)	174 (22)	127 (23)	47 (18)	179 (19)	109 (19)	70 (20)	
Middle	347 (20)	146 (18)	94 (17)	52 (20)	201 (22)	126 (22)	75 (21)	
Second	347 (20)	159 (20)	107 (20)	52 (20)	188 (20)	119 (20)	69 (20)	
Lowest (i.e., poorest)	348 (20)	135 (17)	94 (17)	41 (16)	213 (23)	130 (22)	83 (24)	
Caste/tribe of the head-of-household								0.034**
Scheduled caste	277 (18)	97 (13)	71 (15)	26 (11)	180 (23)	108 (22)	72 (24)	
Scheduled tribe	195 (13)	88 (12)	50 (10)	38 (16)	107 (13)	57 (11)	50 (17)	
Other backward caste	617 (41)	294 (41)	195 (40)	99 (42)	323 (41)	217 (43)	106 (36)	
Other caste	428 (28)	242 (34)	170 (35)	72 (31)	186 (23)	118 (24)	68 (23)	
Head-of-household's occupation								0.359
Professional	86 (6)	50 (7)	31 (7)	19 (9)	36 (5)	22 (5)	14 (5)	
Sales/service worker	172 (12)	82 (12)	61 (13)	21 (10)	90 (12)	67 (14)	23 (8)	
Production worker	331 (23)	159 (23)	126 (27)	33 (15)	172 (22)	114 (24)	58 (20)	
Agricultural worker	661 (45)	296 (43)	189 (40)	107 (49)	365 (47)	201 (42)	164 (57)	
Unemployed	217 (15)	108 (16)	68 (14)	40 (18)	109 (14)	78 (16)	31 (11)	

Hindu religion	1,480 (98)	687 (98)	462 (98)	225 (97)	793 (99)	499 (99)	294 (98)	0.686
Village-level characteristics								
Area of village (hectares) - median (IQR)	3.3 (2.1-6.1)	3.5 (2.0-6.1)	3.5 (2.1-6.1)	3.5 (1.8-6.1)	3.3 (2.1-5.2)	3.7 (2.3-5.3)	3.1 (1.9-4.9)	0.798
Density of village (households per hectare) - median (IQR)	49.2 (37.5-63.4)	40.9 (32.9-57.0)	40.9 (32.9-57.0)	42.0 (32.9-57.5)	56.2 (40.9-69.8)	56.2 (41.6-65.6)	56.2 (40.9-69.8)	0.012**
Paved road access to the village	1,440 (83)	677 (97)	456 (97)	221 (96)	763 (95)	487 (97)	276 (92)	0.575
Behavioural and nutritional outcomes of interest								
WASH facilities coverage outcomes								
Improved drinking water source PIPED, on-premise ‖	608 (35)	535 (67)	353 (65)	182 (70)	73 (8)	50 (9)	23 (7)	<0.001***
Interruption of drinking water availability - 24+ hrs in past 2 wks	153 (9)	110 (14)	70 (13)	40 (15)	43 (5)	27 (5)	16 (5)	0.003**
On-site improved household sanitation facility ‖	865 (50)	674 (84)	448 (83)	226 (87)	191 (21)	113 (20)	78 (22)	<0.001***
On-site enclosed bathing room	670 (39)	568 (71)	379 (70)	189 (72)	102 (11)	70 (12)	32 (9)	<0.001***
WASH facilities utilisation outcomes								
HH members 5+ yrs usually defecate in on-site, improved facility	527 (32)	427 (56)	290 (57)	137 (55)	100 (12)	58 (11)	42 (13)	<0.001***
Presence of completed pit	885 (98)	683 (99)	448 (98)	235 (100)	202 (97)	119 (96)	83 (98)	0.154
Presence of a pan that is broken, blocked, or full of leaves/debris	851 (95)	42 (6)	30 (7)	12 (5)	7 (3)	4 (3)	3 (4)	0.081
Exclusive use of an on-site enclosed bathing room ¶	594 (89)	508 (89)	339 (90)	169 (89)	86 (84)	60 (86)	26 (81)	0.160
Presence of working water source in/near bathing room	526 (79)	470 (83)	318 (84)	152 (80)	56 (55)	41 (59)	15 (47)	0.001***
Presence of roof on bathing room	650 (97)	555 (98)	369 (97)	186 (98)	95 (93)	65 (93)	30 (94)	0.035**
Child-level characteristics and nutritional outcomes – Children aged less than five years								
	N=2,086	**n=974**			**n=1,112**			
Child age - $\bar{x}$ (sd)	1.92 (1.41)	1.91 (1.41)			1.94 (1.41)			0.624
	N=71	**n=30**			**n=41**			
Exclusive breastfeeding amongst children aged less than 6 mo.	39 (55)	14 (47)			25 (61)			0.238
	N=1,754	**n=813**			**n=941**			
Minimum dietary diversity or higher (i.e., ≥ 4.0 threshold) ††	1,100 (63)	546 (67)			554 (59)			0.015**
	N=1,882	**n=864**			**n=1,018**			
Reported diarrhoea during the past 7 days ‡‡	67 (4)	27 (3)			40 (4)			0.378
	N=1,405	**n=630**			**n=775**			
Height-for-age z-score (HAZ) - $\bar{x}$ (sd)	-1.62 (1.16)	-1.47 (1.18)			-1.74 (1.12)			0.011**
	N=1,407	**n=630**			**n=777**			
Weight-for-height z-score (WHZ) - $\bar{x}$ (sd)	-0.79 (1.06)	-0.76 (1.09)			-0.82 (1.03)			0.456
	N=1,454	**n=646**			**n=808**			
Weight-for-age z-score (WAZ) - $\bar{x}$ (sd)	-1.47 (1.12)	-1.35 (1.14)			-1.57 (1.09)			0.044**

Notes. * Variables considered for inclusion in one or more multivariate models of targeted behavioural and nutritional outcomes of interest, as indicated by bivariate analyses and/or other substantive justification. †Gender of CE survey respondent - note, this person provided data on his/her own demographics and responded to the CE survey, all other data were provided by the primary female caregiver of the matched cohort study during a separate household survey. ‡ Wald p-value adjusted for non-independence of observations at village level for respondent and household characteristics, adjusted for clustering at village and household level for child nutritional outcomes, unadjusted for other covariates (see Tables 2a, 2b, S1a, S1b for adjusted statistics from multivariate analyses). ** $p \leq 0.05$; *** $p \leq 0.001$. § Captures proportion of households in each quintile of the standardised wealth index; derived from 20 imputations. ‖ Per JMP definitions. ¶ Female respondent's reported use of bathing facility. †† Per WHO/UNICEF guidelines. ‡‡ Defined as at least one occasion of three or more loose stools during 24 hours.

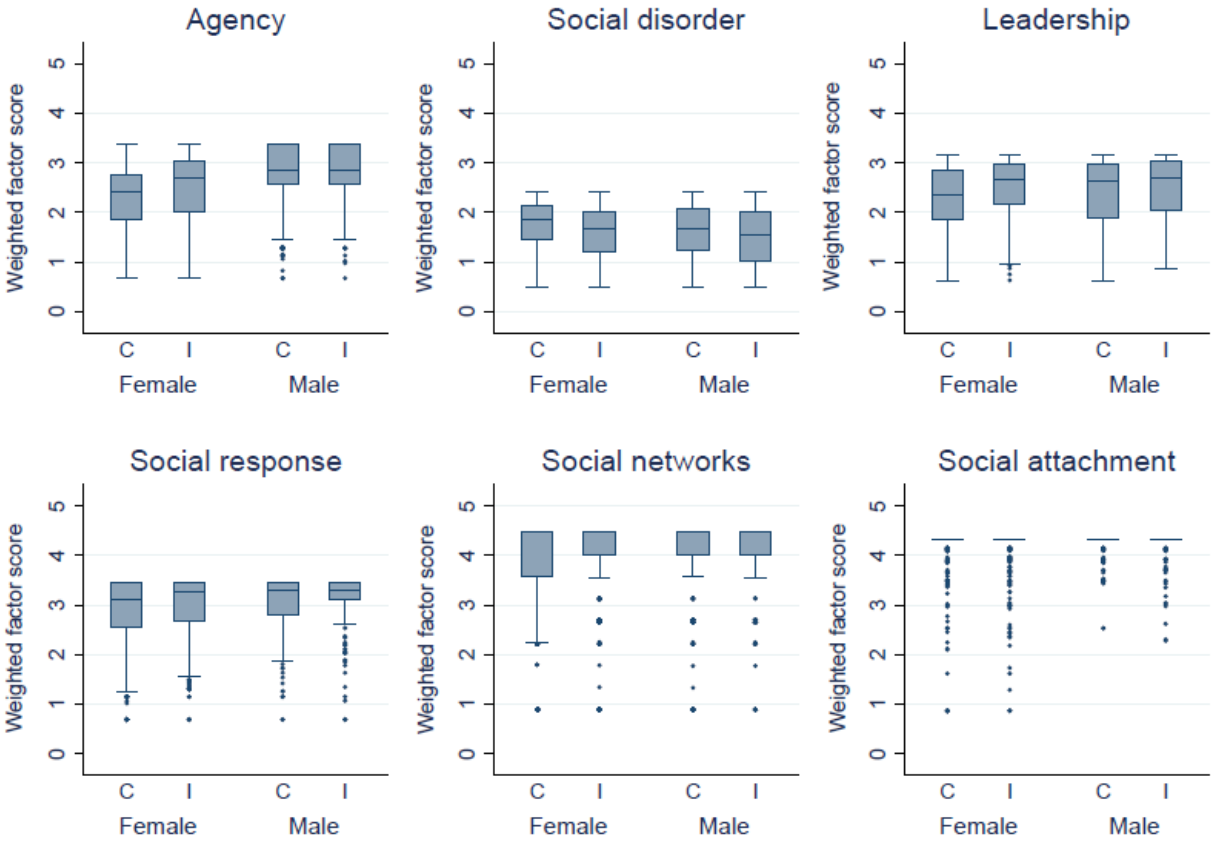


Figure 2. Collective efficacy factor scores (weighted), by respondent gender & study arm. Factor scores are visualised as box plots, which depict the distribution of the data through quartiles. The boxes represent the inter-quartile range (i.e., 25% and 75% quartiles comprise the outer edges of the boxes, while the median is indicated by the line inside the box). The lines that extend vertically from either side of the box (i.e., whiskers) indicate the variability of the data outside the upper and lower quartiles. Outliers are plotted as individual points.

Agency

Agency factor scores were associated with intervention status (p=0.018). Compared to control respondents, those from intervention villages perceived they had more agency over their own future and their capacity to contribute toward the development of their village, and higher behavioural control perceptions about their village's ability to come together to make positive changes (β_1=0.08, 95% CI: 0.01, 0.15). There were also independent associations between agency factor scores and respondent gender (p <0.001) and leadership status (p <0.001), with men scoring agency indicators higher than women (β_2=0.45, 95% CI: 0.38, 0.52), and leaders scoring them higher than non-leaders (β_3=0.37, 95% CI: 0.24, 0.50).

Social disorder

There was evidence of an association between social disorder factor scores and intervention status. Respondents residing in intervention villages provided low levels of agreement with statements suggesting that crime and violence often occurs in their village, differences between people such as wealth or caste often cause problems, and someone in the village may cheat them if they are not careful compared to respondents from control villages. However, the association was conditional on respondent education, and was significant when the respondent's educational attainment was less than or greater than some secondary school (p=0.017).

Village leadership

Village leadership factor scores were associated with intervention status (p<0.001). Respondents from intervention villages perceived that they had informal and formal leaders in their village who were active and could be trusted, and scored village leadership higher than those residing in control villages (β_1=0.16, 95% CI: 0.08, 0.25). There was also an association between village leadership factor scores and the respondent's own leadership status (p=0.004), as respondents with leadership roles scored village leadership higher than those who did not hold a leadership role (β_2=0.21, 95% CI: 0.06, 0.35).

Social response

There was strong evidence of an independent association between social response factor scores and respondent gender (p <0.001), as men tended to perceive that fellow villagers could be trusted, and had a higher propensity to address issues and intervene when needed. They scored social response higher than women (β=0.19, 95% CI: 0.14, 0.24).

Social networks

Social networks factor scores were associated with respondent gender (p<0.001). Compared to women respondents, men perceived that their social networks were more available for and demonstrated reciprocity of individual-level problem solving, and scored social networks indicators higher than women (β_1=0.23, 95% CI: 0.14, 0.32).

Social attachment

Social attachment factor scores were associated with respondent gender (p=0.002). Compared to women, men had higher perceptions regarding social bonds within the village and their own social

acceptance and sense of belonging therein, and scored social attachment indicators higher than women (β_1=0.06, 95% CI: 0.02, 0.09). There was also strong evidence of an independent association between social attachment factor scores and leadership status (p=0.011), as respondents with leadership roles scored attachment indicators higher than those without (β_2=0.06, 95% CI: 0.01, 0.10).

Association between CE factor scores and behavioural and nutritional outcomes of interest

Findings from our stratified multivariate analyses indicated that some CE factor scores were independently associated with behavioural outcomes of interest. Below, we summarise results by behavioural outcome, with further detail provided in Tables 2a and 2b.

WASH facilities coverage outcomes

Prevalence of improved drinking water, piped on-premise was higher amongst control households in which men perceived strong leadership in their village (aPR=2.25, 95% CI: 1.12, 4.53). Prevalence of an on-site improved sanitation facility was higher amongst intervention households in which women had lower perceptions of social disorder (aPR=0.90, 95% CI: 0.84, 0.97). This indicates that lower social disorder may have facilitated the construction of improved sanitation facilities in Gram Vikas MANTRA intervention communities. Prevalence of an on-site, enclosed bathing room was associated with several CE factor scores provided by different individuals. The outcome was positively associated with social response factor scores provided by women from intervention villages (aPR=1.12, 95% CI: 1.02, 1.23), but negatively associated with agency (aPR=0.92, 95% CI: 0.86, 0.99) and social networks (aPR=0.96, 95% CI: 0.91, 1.00) factor scores provided by men from intervention villages and social networks factor scores (aPR=0.73, 95% CI: 0.64, 0.83) provided by women from control villages. The negative association with social networks factors scores may be a result of fewer opportunities these women have to move outside of the household and socialise with their peers.

Table 2a. Multivariate analyses, behavioural outcomes of interest - women respondents

	WASH facilities coverage*						WASH facilities utilisation*			
	Access to improved drinking water, piped on premise		Access to improved on-site sanitation facility		Access to enclosed, on-site bathing room		All HH members usually use improved, on-site sanitation facility		Exclusive use of enclosed, on-site bathing room	
	aPR (95% CI)						aPR (95% CI)			
	I (n=485)	C (n=499)	I (n=484)	C (n=574)	I (n=542)	C (n=424)	I (n=431)	C (n=541)	I (n=379)	C (n=70)
Respondent, household, village characteristics										
Respondent-level characteristic										
Social response factor scores (weighted)					1.12* (1.02, 1.23)	1.02 (0.74, 1.41)				
Social response factor score*Years lived in village						1.02* (1.00, 1.04)				
Social disorder factor scores (weighted)			0.90* (0.84, 0.97)							0.83 (0.67, 1.03)
Social disorder*social networks factor scores										1.13*** (1.06, 1.21)
Agency factor scores (weighted)							1.17* (1.00, 1.37)	0.69* (0.49, 0.97)		
Agency factor score*Respondent educational attainment										
*Agency*No formal education*							1.14 (0.77, 1.68)			
*Agency*At least some primary (up to year 4)*							0.85 (0.72, 1.00)			
*Agency*At least some secondary (up to year 10)*							2.18* (1.01, 4.68)			
*Agency*Higher than secondary*										
Village leadership factor scores (weighted)								3.86* (1.67, 8.97)		
Village leadership factor score*Village area (hectares)								0.88* (0.81, 0.96)		
Social networks factor scores (weighted)						0.73*** (0.64, 0.83)				0.77*** (0.66, 0.89)
Respondent holds a leadership role in the village	1.37*** (1.14, 1.64)									
Number of years respondent has lived in village		1.02* (1.00, 1.03)	1.00* (1.00, 1.004)	1.02* (1.01, 1.03)		0.96 (0.91, 1.01)				
Respondent's highest educational attainment										
No formal education		Ref		Ref		Ref	Ref			Ref
At least some primary (up to year 4)		0.86 (0.27, 2.81)		0.86 (0.50, 1.48)		1.10 (0.48, 2.51)	0.78 (0.25, 2.38)			1.06 (0.70, 1.62)
At least some secondary (up to year 10)		2.32* (1.31, 4.12)		2.64*** (1.52, 4.59)		3.07*** (1.59, 5.92)	1.81* (1.19, 2.74)			1.25 (0.89, 1.77)
Higher than secondary		10.01*** (2.95, 33.95)		6.82*** (3.41, 13.60)		9.31*** (4.85, 17.90)	0.11 (0.01, 1.37)			1.50*** (1.07, 2.11)

Table 2a. Multivariate analyses, behavioural outcomes of interest - women respondents

	WASH facilities coverage*						WASH facilities utilisation*			
	Access to improved drinking water, piped on premise		Access to improved on-site sanitation facility		Access to enclosed, on-site bathing room		All HH members usually use improved, on-site sanitation facility		Exclusive use of enclosed, on-site bathing room	
	aPR (95% CI)						aPR (95% CI)			
	I (n=485)	C (n=499)	I (n=484)	C (n=574)	I (n=542)	C (n=424)	I (n=431)	C (n=541)	I (n=379)	C (n=70)
Household-level characteristic										
Number of household members		1.11* (1.00, 1.23)	1.02* (1.01, 1.03)	1.07* (1.02, 1.12)	1.03*** (1.01, 1.05)	1.09* (1.01, 1.18)				1.02* (1.00, 1.04)
Wealth quintile										
Highest (i.e., richest)					0.99 (0.81, 1.22)					
Fourth					1.21* (1.01, 1.46)					
Middle					1.14 (0.92, 1.43)					
Second					1.15 (0.92, 1.42)					
Lowest (i.e., poorest)					Ref					
Caste/tribe of the head-of-household										
Scheduled caste	Ref	Ref	Ref			Ref				
Scheduled tribe	1.03 (0.56, 1.90)	+ ***	1.26* (1.01, 1.58)			+ ***				
Other backward caste	1.49* (1.04, 2.15)	1.51 (0.57, 3.98)	1.44*** (1.20, 1.72)			1.24 (0.59, 2.63)				
Other caste	1.56* (1.07, 2.27)	2.55 (0.83, 7.91)	1.41*** (1.17, 1.69)			1.65 (0.72, 3.80)				
Head-of-household's occupation										
Professional						Ref			Ref	
Sales/service worker						0.51* (0.31, 0.83)			0.95 (0.89, 1.02)	
Production worker						0.30*** (0.18, 0.48)			0.93 (0.85, 1.01)	
Agricultural worker						0.14*** (0.07, 0.27)			0.91* (0.84, 0.98)	
Unemployed - 6						0.22*** (0.11, 0.45)			0.94 (0.84, 1.05)	
Village-level characteristic										
Area of village (hectares)					1.05*** (1.02, 1.07)	1.13*** (1.07, 1.21)		1.63*** (1.23, 2.17)		
Behavioural and nutritional outcomes of interest										
WASH facilities utilisation outcomes										
Presence of a pan that is broken, blocked, or full of leaves/debris							0.19*** (0.07, 0.51)			
Presence of working water source in/near bathing room									1.76*** (1.29, 2.42)	

Notes. * $p < 0.05$; *** $p \leq 0.001$; + parameter perfectly predicted

Table 2b. Multivariate analyses, behavioural outcomes of interest - men respondents

	WASH facilities coverage*						WASH facilities utilisation*	
	Access to improved drinking water source, piped on premise		Access to improved on-site sanitation facility		Access to enclosed, on-site bathing room		All household members usually use improved, on-site sanitation facility	
	aPR (95% CI)						aPR (95% CI)	
	I (n=229)	C (n=296)	I (n=212)	C (n=293)	I (n=199)	C (n=250)	I (n=212)	C (n=280)
Respondent, household, village characteristics								
Respondent-level characteristic								
Agency factor scores (weighted)					0.92* (0.86, 0.99)			
Village leadership factor scores (weighted)		2.25* (1.12, 4.53)						
Social networks factor scores (weighted)					0.96 (0.91, 1.00)			
Social networks factor scores*Age								
18-34 years					Ref			
35-59 years					0.97 (0.89, 1.06)			
60+ years					1.09* (1.03, 1.15)			
Respondent holds a leadership role in the village		7.97* (1.20, 52.83)			1.15* (1.04, 1.27)			
Age								
18-34 years					Ref			
35-59 years					1.18 (0.83, 1.69)			
60+ years					0.84 (0.69, 1.00)			
Respondent's highest educational attainment								
No formal education				Ref	Ref			
At least some primary (up to year 4)				0.90 (0.35, 2.29)	1.20 (0.95, 1.50)			
At least some secondary (up to year 10)				1.85* (1.04, 3.29)	1.24* (1.01, 1.52)			
Higher than secondary				3.23*** (1.72, 6.08)	1.33* (1.06, 1.68)			
Household-level characteristic								
Number of household members - median (IQR)	1.05* (1.02, 1.08)	1.19*** (1.10, 1.29)	1.03* (1.00, 1.05)					
Wealth quintile								
Highest (i.e., richest)		0.36 (0.07, 1.75)	0.85* (0.73, 0.98)					

Table 2b. Multivariate analyses, behavioural outcomes of interest - men respondents

	WASH facilities coverage*						WASH facilities utilisation*	
	Access to improved drinking water source, piped on premise		Access to improved on-site sanitation facility		Access to enclosed, on-site bathing room		All household members usually use improved, on-site sanitation facility	
	aPR (95% CI)						aPR (95% CI)	
	I (n=229)	C (n=296)	I (n=212)	C (n=293)	I (n=199)	C (n=250)	I (n=212)	C (n=280)
Fourth		0.35 (0.08, 1.46)	0.85 (0.71, 1.01)					
Middle		0.26* (0.07, 0.94)	0.87 (0.74, 1.03)					
Second		0.98 (0.40, 2.37)	0.95 (0.83, 1.08)					
Lowest (i.e., poorest)		Ref	Ref					
Caste/tribe of the head-of-household								
Scheduled caste	Ref	Ref	Ref	Ref	Ref	Ref		Ref
Scheduled tribe	0.42* (0.18, 0.94)	+ ***	0.89 (0.64, 1.23)	0.44 (0.06, 3.05)	1.22 (0.95, 1.57)	+ ***		+ ***
Other backward caste	1.18 (0.89, 1.56)	2.20 (0.42, 11.48)	1.21 (0.99, 1.49)	4.54* (1.71, 12.03)	1.29* (1.03, 1.61)	1.53 (0.41, 5.74)		3.26* (1.21, 8.82)
Other caste	1.27 (0.94, 1.72)	1.58 (0.33, 7.68)	1.22* (1.01, 1.48)	5.87*** (1.97, 17.49)	1.37* (1.09, 1.72)	6.06* (1.46, 25.10)		5.01* (1.61, 15.63)
Head-of-household's occupation								
Professional - 1					Ref	Ref	Ref	
Sales/service worker - 2					1.21 (0.96, 1.51)	0.64 (0.36, 1.15)	0.69* (0.50, 0.97)	
Production worker - 4					0.77 (0.50, 1.19)	0.07* (0.01, 0.48)	0.53* (0.34, 0.82)	
Agricultural worker - 5					1.34* (1.05, 1.72)	0.35*** (0.19, 0.66)	0.65* (0.49, 0.88)	
Unemployed - 6					1.26 (0.98, 1.62)	0.35 (0.12, 1.01)	0.69* (0.48, 0.99)	
Village-level characteristic								
Area of village (hectares)				1.11*** (1.05, 1.17)		1.15*** (1.07, 1.25)		
Paved road access to the village			0.71* (0.55, 0.92)					

Notes. * $p < 0.05$; *** $p \leq 0.001$; † parameter perfectly predicted

WASH facilities utilisation outcomes

Prevalence of household-level utilisation of an improved, on-site sanitation facility for defecation was higher amongst intervention households in which women had higher perceived agency (aPR=1.17, 95% CI: 1.00, 1.37). Prevalence of household-level sanitation facilities utilisation in control villages was higher amongst households in which women scored village leadership indicators high (aPR=3.86, 95% CI: 1.67, 8.97), but lower amongst households in which women had lower perceptions of agency (aPR=0.69, 95% CI: 0.49, 0.97). Prevalence of women's utilisation of an on-site, enclosed bathing room was negatively associated with social networks factor scores provided by women from control villages (aPR=0.77, 95% CI: 0.66, 0.89). Social disorder factor scores provided by these women were also associated with use of an on-site bathing room, but this association was conditional on the women's social networks factor score.

Nutritional outcomes amongst children aged less than five years

Findings from our stratified multivariate analyses indicate that some CE factor scores provided by adults from study household were independently associated with nutritional outcomes amongst children aged less than five years who resided in the household. Below, we summarise results by nutritional outcome, with further detail provided in Tables S1a and S1b.

HAZ was associated with agency factor scores provided by men from control villages (β=0.22, 95% CI: 0.03, 0.40), and social disorder factor scores provided by men from intervention villages (β=0.29, 95% CI: 0.03, 0.54). This nutritional outcome was also negatively associated with social response factor scores provided by women from intervention villages (β=-0.25, 95% CI: -0.40, -0.09).

WHZ was negatively associated with social response perceptions of women in intervention communities, but this was conditional on the respondent's age. WHZ was also positively associated with village leadership factor scores (β_1=0.21, 95% CI: 0.01, 0.41), and negatively associated with response factor scores (β_2=-0.31, 95% CI: -0.59, -0.03) provided by men in intervention villages. This nutritional outcome was also independently and positively associated with agency factors scores (β_1=0.22, 95% CI: 0.04, 0.39), and negatively associated with social attachment factor scores (β_2=-0.51, 95% CI: -0.97, -0.05) provided by men in control villages.

WAZ was negatively associated with social response factor scores amongst women in intervention communities (β= -0.21, 95% CI: -0.38, -0.04). WAZ was also independently associated with agency

factor scores provided by men from control villages (β= 0.27, 95% CI: 0.10, 0.45). This finding suggests that in villages where no intervention occurred, children from households in which men scored agency factors relatively high had better weight-for-age z-scores.

DISCUSSION

This study examined associations between CE factors and improved WASH behaviours and downstream nutritional outcomes. Three CE factors were associated with village participation in MANTRA, a WASH programme that focuses on collective action, social inclusion, and social equity. Agency and village leadership perceptions were higher, while social disorder was perceived to be lower amongst MANTRA participants than their counterparts from control villages. We also observed that CE factors were independently associated with WASH facilities coverage and utilisation behaviours, and nutritional outcomes in both intervention and control villages. Observing associations between CE factors and specific MANTRA programme targets in intervention villages (i.e., WASH facilities coverage and utilisation) suggests that CE is an important factor in the uptake of improved WASH behaviours. Results generated from control villages indicated that certain CE factors, such as village leadership and agency, were associated with individual and household-level uptake of improved WASH behaviours and nutritional outcomes, respectively, even in the absence of community-based WASH initiatives. Based on these findings, we conclude that assessing CE-related factors amongst villages, groups, or collectives prior to intervention can help target programmes predicated on collective action, and harness these behavioural control perceptions to further facilitate intervention uptake amongst those ready for change.

Our findings suggest that CE-related factors may have the potential for changing and sustaining behaviour, and that CE factors may be behavioural antecedents of improved WASH practices, regardless of directed intervention. Given the limitations of our study design, we are not able to draw causal inferences from these results. In other words, we cannot rule out that the intervention villages were different from control villages on unmeasured criteria that may have been confounders. That said, findings from this study do provide justification for further examination into possible causal associations between CE factors, behavioural outcomes, and downstream health and development impacts.

Results from our study align with other evidence that suggests social constructs are important behavioural factors related to the effectiveness of WASH interventions. For example, previous studies of social capital have revealed that constituent factors of social capital, such as differences in leadership and community structures, mediate collective action related to sanitation use and water access [27, 28]. While our CE scale incorporates indicators that tap to cognitive and structural social capital, it moves beyond measuring capital, and includes important indicators that allow for the measure of behavioural control perceptions at the individual (i.e., self-efficacy) and collective (i.e., collective agency) levels. These factors play important roles in collective goal setting, group performance including the amount of effort and resolve the group will expend, and ultimate goal attainment [12, 13, 29].

Our findings also corroborate existing evidence regarding the role social conflict and exclusion may play in inhibiting uptake of improved WASH behaviours [28, 30]. While these findings concur with others suggesting social fragmentation is a barrier to improved sanitation, they do not reinforce the presumption that differences in caste, class, and religion preclude heterogeneous villages or collectives from being able to perceive of themselves as cohesive units that can cooperate to achieve collective goals. That said, we did observe persistent caste disparities related to prevalence of improved, on-site sanitation facilities, with lower prevalence observed amongst scheduled caste (i.e., Dalit) and scheduled tribe (i.e., Adivasi) households compared with households belonging to other castes (e.g., Other Backward Caste, other castes).

Results from our examination of CE revealed the potential for unintended negative consequences of improved WASH practices on social networks. Moving outside the household compound to fetch water and bathe provides opportunities for socialisation amongst women. Sometimes, these activities provide the only opportunities for women to move outside of their household compounds. Installing sanitation and hygiene facilities and water sources on the premises of the household compound improves living conditions, but may negatively affect perceptions regarding one's social networks. Given no pre- or peri-intervention measures of social networks perceptions were captured, we are unable to ascertain whether these individuals had low perceptions of their social networks prior to the installation of their on-site bathing rooms. The MANTRA intervention incorporates intervention activities to increase empowerment and social mobilisation of women (e.g., establishment of self-help groups), which may explain why our results did not indicate lower

social networks perceptions amongst women from intervention communities. These findings have implications for programming. They imply that social inclusion initiatives and alternative socialisation opportunities, particularly those targeted toward women, need to be incorporated into mainstream WASH programming. These intervention components can ensure that as people, particularly women, move to exclusively utilise on-premise WASH facilities, they maintain opportunities to socialise and develop and maintain their social networks. Similar suggestions were made by others who have studied the role of social capital in the uptake of WASH interventions [28].

Our nutritional outcome data align with other findings emerging from India. Average HAZ amongst children aged less than five years from MANTRA intervention villages were similar to average HAZ amongst Indian children in 2015-2016, as revealed by an analysis of the fourth round of National Family Health Survey (NFHS), India's Demographic and Health Survey (DHS) [31]. However, these height-for-age z-sores were significantly better than those of children aged less than five years from control villages. Average HAZ amongst children from control villages was similar to average HAZ amongst children of the same age from disadvantaged states [31]. Our findings indicated that CE factors such as village leadership, agency, and social response were associated with nutritional outcomes. However, our results also revealed some findings that were counter-intuitive (e.g., negative associations between social response factor scores and nutritional outcomes, positive association between social disorder factor scores and HAZ, negative association between social attachment and WHZ). These counterintuitive findings may have arisen as a result of factors such as other infectious diseases (e.g., malaria and measles), care practices, and maternal undernutrition known to influence child nutritional outcomes [32, 33], but were not measured during the study (i.e., were not included in our models).

Our study has several limitations. The CE scale we developed and employed reflected the first iteration of CE scales developed by our research group. This early version was less developed than subsequent iterations that underwent additional formative development, and built on lessons learnt through the development and administration of this India-specific tool. While factor model fit was adequate, the social networks factor was less conceptually saturated than later iterations of our CE scale [19], and certain motivational aspects related to CE were not fully assessed. It may be possible that true associations did not emerge within these India-specific multivariate models due to these shortcomings. Given we were conducting an external evaluation of an intervention for

which early stages of implementation commenced more than ten years prior to data collection, no pre-intervention CE or related measures were available. As a result of our post-only design, we were not able to investigate causal relationships between CE and intervention effectiveness or directionality thereof. However, given the substantial length of time between intervention implementation and data collection, we were able to assess sustainability of outcomes. Our findings suggest that the impacts produced by the MANTRA intervention were enduring up to three years post-programme completion. Finally, while we measured indicators of exclusive breast feeding and dietary diversity, per WHO/UNICEF guidelines (though we applied these for children up to five years of age), we did not measure indicators of maternal nutrition or care practices, which evidence suggests may influence nutritional outcomes amongst children [32, 33].

CONCLUSIONS

Our findings indicate that CE factors are associated with improved WASH practices and nutritional outcomes amongst children aged less than five years. This study demonstrated the ability of our CE scale to pinpoint specific CE-related factors that either contribute toward or prevent the uptake of improved WASH behaviours. CE-related factors can and should be considered when identifying villages, groups, or collectives amongst which community-based interventions should be targeted, as those indicating relatively high CE factor scores are likely more ready for and motivated to change. CE-related considerations should also be incorporated into the design of community-based intervention content and implementation approaches.

The results of this study yielded two key hypotheses related to diagnostic examinations of community-based interventions that should be tested in future experimental studies. First, we hypothesise that CE may be an important interpersonal, or community-level behavioural antecedent of collective action, and in turn, effective community-based interventions. Second, we hypothesise that there is a bi-directional relationship between collective efficacy and intervention effectiveness. In other words, we hypothesise that innate properties of CE factors exist amongst collectives, and may play a role in the uptake of community-based interventions, but CE may also be developed and enhanced through the interventions.

ACKNOWLEDGEMENTS

We acknowledge, with gratitude, the support and cooperation of Gram Vikas, the study communities engaged with MANTRA, and the respondents who so generously gave their time to

participate in our study. We are also grateful to our team of enumerators and data managers. This work was supported by Bill & Melinda Gates Foundation (OPP1008048 and OOP1125067).

Table S1a. Multivariate analyses, nutritional outcomes of interest - women respondents

	Nutritional outcomes amongst children aged less than five years					
	Height-for-age z-score (haz)		Weight-for-height z-score (whz)		Weight-for-age z-score (waz)	
	β coefficient (95% CI)					
	I (n=401)	C (n=411)	I (n=371)	C (n=361)	I (n=435)	C (n=433)
Respondent, household, village characteristics						
Respondent-level characteristic						
Social response factor scores (weighted)	-0.25* (-0.40, -0.09)		-0.07 (-0.29, 0.16)		-0.21* (-0.38, -0.04)	
Social response factor score*Respondent age						
18-34 years			Ref			
35-59 years			-0.13 (-0.66, 0.40)			
60+ years			-0.14 (-0.47, 0.20)			
			-0.91* (-1.68, -0.13)			
18-34 years						
Age						
18-34 years			Ref			
35-59 years			0.05 (-0.27, 0.36)			
60+ years			0.52* (0.01, 1.04)			
Respondent's highest educational attainment						
No formal education	Ref		Ref			Ref
At least some primary (up to year 4)	0.50*** (0.24, 0.77)		0.40 (-0.86, 1.66)			0.18 (-0.26, 0.63)
At least some secondary (up to year 10)	0.25* (0.02, 0.49)		0.48 (-0.54, 1.51)			0.23* (0.02, 0.45)
Higher than secondary	0.16 (-0.46, 0.78)		1.68 (-0.61, 3.98)			-0.09 (-0.62, 0.44)
Household-level characteristic						
Caste/tribe of the head-of-household						
Scheduled caste		Ref	Ref	Ref		Ref
Scheduled tribe		-0.21 (-0.65, 0.24)	-0.56* (-1.01, -0.11)	-0.25 (-0.69, 0.19)		-0.14 (-0.62, 0.34)
Other backward caste		0.42* (0.10, 0.73)	0.001 (-0.40, 0.40)	0.38*** (0.17, 0.59)		0.39* (0.11, 0.67)

Other caste		0.52* (0.18, 0.86)	0.11 (-0.29, 0.50)	0.39* (0.10, 0.69)		0.46* (0.09, 0.82)
Head-of-household's occupation						
Professional				Ref		
Sales/service worker				0.63* (0.05, 1.20)		
Production worker				0.70* (0.15, 1.25)		
Agricultural worker				0.83*** (0.27, 1.38)		
Unemployed				0.94*** (0.33, 1.55)		
Village-level characteristic						
Area of village (hectare)	0.10*** (0.05, 0.15)				0.11*** (0.06, 0.16)	0.04* (0.01, 0.06)
Behavioural and nutritional outcomes of interest						
WASH facilities coverage and utilisation outcomes						
Improved drinking water source PIPED, on-premise		0.41* (0.10, 0.72)				0.62*** (0.28, 0.96)
On-site improved household sanitation facility				0.51* (0.17, 0.86)		
WASH facilities utilisation outcomes						
ALL HH members aged 5 yrs+ usually defecating in an on-site improved sanitation facility	0.48*** (0.25, 0.71)					
Female respondent's use of an on-site bathing room						
Always use					Ref	
Sometimes use					-0.58* (-1.06, -0.11)	
Never use					-0.23 (-0.68, 0.21)	

Notes. * p < 0.05; *** p ≤ 0.001

Table S1b. Multivariate analyses, behavioural and nutritional outcomes of interest - men respondents

	Nutritional outcomes					
	Height-for-age z-score (haz)		Weight-for-height z-score (whz)		Weight-for-age z-score (waz)	
	β coefficient (95% CI)					
	I (n=204)	C (n=251)	I (n=204)	C (n=305)	I (n=195)	C (n=272)
Respondent, household, village characteristics						
Respondent-level characteristic						
Social response factor scores (weighted)			-0.31* (-0.59, -0.03)			
Social disorder factor scores (weighted)	0.29* (0.03, 0.54)					
Agency factor scores (weighted)		0.22* (0.03, 0.40)		0.22* (0.04, 0.39)		0.27* (0.10, 0.45)
Village leadership factor scores (weighted)			0.21* (0.01, 0.41)			
Social attachment factor scores (weighted)				-0.51* (-0.97, -0.05)		
No formal education	Ref				Ref	
At least some primary (up to year 4)	0.88* (0.15, 1.61)				0.43 (-0.12, 0.97)	
At least some secondary (up to year 10)	0.39 (-0.08, 0.86)				0.18 (-0.23, 0.59)	
Higher than secondary	0.90*** (0.39, 1.40)				0.64* (0.02, 1.27)	
Household-level characteristic						
Wealth quintile						
Highest (i.e., richest)						-0.32 (-0.74, 0.09)
Fourth						-0.43* (-0.83, -0.04)
Middle						-0.26 (-0.67, 0.16)
Second						-0.13 (-0.57, 0.32)
Lowest (i.e., poorest)						Ref
Caste/tribe of the head-of-household						
Scheduled caste		Ref			Ref	Ref
Scheduled tribe		-0.30 (-0.72, 0.11)			-0.23 (-0.69, 0.23)	-0.54* (-0.95, -0.13)

Other backward caste		0.36 (-0.002, 0.74)	0.50* (0.04, 0.97)	0.50* (0.09, 0.91)
Other caste		0.63* (0.05, 1.21)	0.32 (-0.20, 0.83)	0.67* (0.15, 1.18)
Village-level characteristic				
Area of village (hectares)	0.11* (0.04, 0.18)			
Child age			-0.16*** (-0.25, -0.06)	
Behavioural and nutritional outcomes of interest				
WASH facilities coverage outcomes				
Improved drinking water source PIPED, on-premise			0.54* (0.17, 0.91)	
On-site improved household sanitation facility		0.64* (0.15, 1.13)	0.66* (0.14, 1.18)	
WASH facilities utilisation outcomes				
ALL HH members aged 5 yrs+ usually defecating in an on-site improved sanitation facility		-0.56* (-1.02, -0.11)		

Notes. * $p < 0.05$; *** $p \leq 0.001$

SUPPLEMENTAL MATERIAL

Collective efficacy and the adoption and influence of a community-based water supply and sanitation programme: a cross-sectional study in Odisha, India

Maryann G. Delea, Gloria D. Sclar, Corey L. Nagel, Sheela S. Sinharoy, Heather Reese, Thomas F. Clasen

ADDITIONAL BACKGROUND

Additional details regarding the Gram Vikas MANTRA programme

Gram Vikas will not initiate village enrollment into MANTRA before a village demonstrates its unanimous commitment to the programme. Details regarding the MANTRA programme are presented below, and highlight how factors related to collective behaviour are deeply engrained in the programme's design. MANTRA is implemented in three phases, which are carried out over approximately three years. The programme commences with the Motivational Phase, then continues through to Implementation and Completion Phases.

During its Motivational Phase, Gram Vikas identifies villages with the potential of being fully engaged with the MANTRA programme. In addition to the programme requirements outlined in the main text of this paper, during this phase, Gram Vikas requires that potential programme villages demonstrate their commitment to the objectives of the MANTRA programme prior to fully engaging with the village, and moving it to the Implementation Phase of MANTRA. A Village Executive Committee comprised of five men and five women must also be elected by the village, and representation must be proportional to the village's caste and class distribution. The village must also establish a Village Corpus Fund, which every household pays into every month until the fund is complete, at which point it is deposited into the bank, the interest of which can be used to expand the village water distribution system and toilets and bathing rooms for new families. Once a village has demonstrated this level of commitment, Gram Vikas then fully engages the village in the MANTRA programme, as it initiates its Implementation Phase. During this phase, MANTRA staff provide technical assistance and supervise the construction of the village's water tank and distribution system and every household's sanitation and hygiene block. These Field Officers build capacity of the Village Executive Committee to apply for reimbursement from the government for the construction of the village water distribution

system, which is fully subsidised. MANTRA staff also assist with the application of government subsidies for the installation of improved household sanitation facilities.

In addition to these basic programming activities, MANTRA aims to improve women's empowerment and social inclusion, and social equity across castes through other programme initiatives such as self-help groups.

ADDITIONAL METHODS

Summary of the five-step CE scale development and validation process

We used a multiple-step, sequential exploratory and confirmatory approach [17] to develop, refine, and validate a theoretically-grounded and evidence-based CE measurement scale [18]. This process relied heavily on a factor analytic approach. Factor analysis is a psychometric method that allows for the measurement of latent constructs that cannot be directly observed or measured. In a factor analytic framework, latent constructs are measured through the analysis of manifest variables (e.g., survey items), indicators that represent certain aspects of the latent construct [34]. The analyses of these data can elucidate the underlying structure of the construct and its constituent sub-constructs.

Factor analysis comprises a suite of analytical methods, including exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA is a descriptive analytical method used to determine the number of common factors in a measurement model, and identify which measured variables are indicators of the latent construct (i.e., identify factor structure) [24]. CFA is an evaluation approach that allows for direct testing and validation of hypothesised factor structures to assess their appropriateness as measurement models (i.e., determine construct validity, falsify hypothesised models). These types of analytical methods utilise matrix algebra to generate statistics used to reveal a construct's underlying factor structure [35].

Our multiple-step scale development and validation approach entailed a literature review to examine conceptualisations of CE and other related social constructs (e.g., social capital, social cohesion, self-efficacy). We performed an applied thematic content analysis to re-organise emergent sub-constructs into key CE domains, dimensions/factors, and facets to generate a hypothesised CE framework that was used to develop a draft survey. The resulting framework represented a seven-factor conceptualisation, with items tapping to social order, social response, social capital, social equity, common values, community attachment, and agency.

These factors represent aspects of three domains: informal social control, social cohesion, and behavioural control.

After generating an item pool by extracting relevant prompts from existing survey instruments [18], coding those items against our framework, removing repetitive items, and designing new ones, as needed, we contextually authenticated the tool via a formative research process. We re-structured all survey items such that they worked well with a five-point, Likert type response format [18]. Once our draft tool was developed, it was translated into Odia, and back-translated into English to ensure the quality of the translations. During January 2016 (summer season in Odisha, India), we conducted formative work on the tool. This consisted of cognitive validation of the tool to assess face validity. We performed a series of in-depth cognitive interviews that employed 'think aloud' and probing techniques to gauge participants' understanding of the meaning of survey prompts and appropriateness of response selection [36]. Findings from these in-depth cognitive interviews informed survey refinements. Subsequent to pilot testing and survey refinement, the study team administered the refined, 36-item CE survey amongst all available and consenting households enrolled in the larger matched cohort study.

Summary of psychometric analyses

Initial data cleaning and descriptive analyses were performed in Stata (version 15.0 StataCorp, College Station, TX, USA). We performed subsequent descriptive and all factor analyses (EFA, preliminary CFA, single-group CFA of EFA-derived factor solutions, multiple-group CFA and Multiple Indicators Multiple Causes [MIMIC] modelling) in Mplus software (version 8 Muthén & Muthén, Los Angeles, CA, USA).

To prepare our data for analyses, we partitioned our CE dataset in two through the use of a random-number seed to identify two separate random split-halves. We designated one random split-half sample for scale development via EFA; the remaining random hold-out sample was reserved for scale validation via CFA of the EFA-derived factor solutions. Univariate analyses performed in Stata and verified results in Mplus examined respondent/household characteristics and item distributions for all CE items.

Preliminary confirmatory factor analysis of the hypothesised CE framework

We decided *a priori* to first test our hypothesised CE framework via a preliminary CFA [37]. Poor model fit statistics for this preliminary CFA would signal that the hypothesised CE framework may need modification in order to produce an appropriate CE measurement framework. In the event model fit statistics indicated poor fitness, we decided *a priori* that we would perform EFA

to determine alternative CE factor structures derived from our own empirical data, and conduct CFA again to test and validate the resulting EFA-derived factor structures [24].

For preliminary CFA, we utilised a robust weighted least-squares with mean and variance adjustment (WLSMV) estimation method based on assessments of polychoric correlation matrices [38]. A sandwich estimator was applied to adjust for non-independence of observations within clusters. Because it would have been justifiable to conclude our analyses with CFA if the complex preliminary CFA indicated good model fit, we performed these analyses on the full sample. We examined goodness-of-fit indices, assessing both absolute fit (e.g., χ^2:df ratio, root mean square error of approximation [RMSEA]) and incremental, or relative fit (e.g., comparative fit index [CFI], Tucker-Lewis index [TLI]). Standard thresholds of acceptable and good model fitness were employed (i.e., χ^2:df ratio < 3.0; RSMEA of ≤ 0.10 acceptable fit, ≤ 0.05-0.06 good fit; CFI & TLI ≥ 0.90 acceptable fit, ≥ 0.95 good fit) [24, 39, 40]. Factor loadings less than 0.3 were considered non-salient (i.e., not statistically meaningful) [41]. *Post hoc* refinements included the deletion of items with non-salient (factor loadings < 0.3) and/or non-significant (two-tailed $p > 0.05$) factor loadings. We also dropped all factors with less than two items with salient and significant factor loadings, as these factors may have insufficient component saturation, meaning the factor may not have been fully conceptually explained by the emergent items, which could compromise factor interpretation [37].

Exploratory factor analysis

We performed complex EFA on one split-half of data. As with preliminary CFA, we utilised a robust WLSMV estimation method based on assessments of polychoric correlation matrices for EFA, and applied a sandwich estimator to adjust for non-independence [38, 40, 42]. An oblique rotation was indicated due to hypothesised item correlation, and Promax was selected *a priori* as the specific oblique rotation method for these analyses.

Decision rules related to factor retention were based on a combination of: 1) mathematically-based and heuristic descriptive guides (i.e., Kaiser-Guttman rule [eigenvalue > 1.0] [43], scree-plot); 2) goodness-of-fit; and 3) other substantive justification, such as results from cognitive interviews, and theoretical and empirical evidence [24, 37, 44]. As with preliminary CFA, we employed a holistic approach to evaluate goodness-of-fit indices for EFA. The same thresholds used for preliminary CFA were used for EFA, but we also included an assessment of root mean square residual (RMSR); values below 0.08 indicate reasonable model fit [39].

Oblique rotations produce pattern coefficients that do not fully characterise the relationship between an item and a given factor [44]. Therefore, in order to appropriately interpret EFA results, we evaluated both the factor pattern and factor structure matrices [37]. Structure and pattern coefficients with an absolute value greater than 0.3 were considered salient. Items with factor loadings less than this threshold poorly measured the latent factors, and were eliminated in a step-wise manner [41]. We iteratively re-analysed measurement models subsequent to item reduction [44]. In order to be retained, factors needed to demonstrate adequate component saturation and sufficient evidence that they were at least adequately measured (i.e., at least two items with factor loadings greater than 0.3, and no or limited item cross-loadings) [45]. Only complex variables (i.e., those with salient factor loadings on more than one factor [cross-loadings]) with strong substantive justification for their cross-loadings were retained. Models that represented the most readily interpretable (i.e., the simplest solution, per Thurstone criteria [46]) and theoretically justifiable solutions were selected for the refined, gender-specific factor solutions [44].

With regard to the interpretation of EFA results, factor loadings indicate the pattern of item-factor relationships, and are often referred to as pattern coefficients [24]. Factor loadings represent completely standardised estimates of regression slopes for predicting the indicators from the latent variable [24]. While some methodologists caution against the use of thresholds, common guidelines for the interpretation of factor loadings can be used to facilitate interpretation of results (e.g., factor loadings > 0.71 excellent, > 0.63 very good, > 0.55 good, > 0.45 fair, and > 0.32 adequate) [41].

Confirmatory factor analysis of the EFA-derived factor solution

During CFA of EFA-derived factor solutions, we used the split-half hold-out sample to validate the EFA-indicated measurement model. The underlying structure used to operationalise the latent factors were those indicated in the factor solution produced via EFA. We identified the scale of every latent factor through the use of marker indicator items, which we identified as the item that demonstrated the highest factor loading on its respective factor, per EFA results [24]. As with our preliminary CFA and EFA, we performed these CFAs using WLSMV with a sandwich estimator to adjust for non-independence. Through *post hoc* model refinements, we eliminated items with non-salient (i.e., factor loading < 0.3) and/or non-significant factor loadings.

We used the same process for holistically examining goodness-of-fit and carrying out *post hoc* model refinements for the CFA of EFA-derived models as those employed during the preliminary CFA. After examining fit statistics, we assessed residuals and modification indices for indications of localised areas of strain (i.e., misfit) in the measurement models. Modification indices greater than 3.84 indicated opportunities for further model refinement and fit improvement, through the estimation of additional parameters, if justified [24, 47].

Multiple-group CFA and MIMIC modelling for assessment of differential item functioning

We performed Multiple Indicator Multiple Causes (MIMIC) modelling to test the validity of our parsimonious CE measurement model in the presence of other relevant covariates, and assess differential item functioning (DIF) [48]. DIF, or measurement *non*-invariance occurs when people from different groups (e.g., men, women) with similar levels of the latent construct have different probabilities of responding to an item in a certain way [49]. Our structural equation MIMIC models consisted of a measurement model component reflected by the refined parsimonious CE model, and a structural model component that specified the direct effects of gender and leadership covariates on latent factor variables and relevant item indicators. Significant direct effects would indicate DIF between these groups.

The same validation sub-sample used for CFA was used for these analyses. As we constructed our MIMIC model, we first established a baseline model that introduced gender and leadership status covariates, but assumed no direct effects of the covariates on any individual CE items. Then, we employed a step-wise, forward selection approach to assess direct effects between these covariates and relevant item indicators. We examined the modification indices, and identified the item indicator with the highest significant, meaningful, and substantively justifiable modification index. We added a direct path between the identified item indicator and relevant covariate. We employed the DIFFTEST option in Mplus to assess whether the additional direct path improved model fit. Given we had a relatively large sample size, it was likely that DIFFTEST statistics would be significant [49], so we evaluated and compared other model fit indices as well.

Collective efficacy factor score calculation

We used a coarse CE factor score calculation methods (i.e., non-refined, un-sophisticated procedures) to generate weighted average CE factor scores [26]. Higher factor scores represented higher levels of perceived behavioural control over the respective CE factors. We generated CE factor scores for each respondent by summing his/her responses across all items

in each factor (i.e., 1=completely disagree, 2=partially disagree, 3=neither agree nor disagree [neutral], 4=partially agree, 5=completely agree), and dividing that sum by the number of items tapping to the factor. This approach, however, assumes all items have the same level of influence, or measurement proximity to their respective latent factor. We have demonstrated that this is not the case. We then calculated weighted average CE factor scores, for which a weight that was equivalent to the item's factor loading was applied to each item score prior to the generation of the average (weighted) factor score [26].

Multilevel modelling of CE and key outcomes of interest

We performed a series of univariate, bivariate, and multivariate analyses in Stata to examine whether there was evidence of independent associations between collective efficacy, as measured through our CE scale factor scores and: 1) intervention status, 2) uptake of improved WASH behaviours, and 3) nutritional status of children aged less than five years.

We conducted univariate analyses to describe various aspects of the study sample (e.g., respondent, household, and study cluster characteristics; CE factor scores; prevalence of behavioural outcomes and undernutrition). We also utilised descriptive analyses to examine unit and item non-response rates. We performed bivariate analyses to explore crude associations between potential covariates and behavioural and nutritional outcomes. We performed multivariate analyses to determine whether there was evidence of independent associations between CE factor indicators and behavioural and downstream nutritional outcomes, while adjusting for other relevant explanatory variables.

We stratified our examinations of CE factor scores by gender given we observed differentials in CE perceptions between women and men during preliminary analyses. We stratified by study arm given the complex relationship between CE and intervention exposure. We assessed results within the intervention arm to investigate whether there was evidence of independent associations between CE factors and behavioural and nutritional outcomes. Given CE is an innate characteristic operating amongst any collective or group of people, we also investigated whether there was evidence of associations between CE factors and uptake of improved WASH behaviours in control villages. Exploring whether there was evidence of these relationships in control villages provided counterfactual assessments while also allowing us to examine innate properties of CE where no community-based WASH interventions were operating.

As indicated in the main text of this paper, multilevel modelling occurred in three phases. We employed a two-level random intercepts model using a Poisson estimator to produce

prevalence ratios for phase I and II to accommodate the hierarchical data structure, and adjust for village and household/respondent-level covariates. For these models, village-level variables represented level two covariates. Given CE data were only collected from one household member, both household and respondent-related variables represented level one covariates. The conditional distribution of the response given the random effects is assumed to be Poisson. Therefore, robust variance estimation was applied to adjust for violations against the assumption that the data fit a Poisson distribution.

Existing evidence suggested that there may be important differences in behavioural control perceptions (i.e., CE factor scores) between men and women, and those with leadership status vs. non-leaders. Therefore, during phase I, we examined whether there were important differences between CE factor scores across study arms and between genders and those with leadership status vs. those without. We then examined whether there was evidence of independent associations between CE factor scores and intervention status (i.e., exposure to the Gram Vikas MANTRA intervention) using maximum likelihood estimation via mixed effects linear regression with robust standard errors. This examination into differentials in CE factor scores produced subject-specific estimates.

Phase III analyses comprised of modelling of nutritional outcomes. Observations from all children aged less than five years of age were included in these analyses. Therefore, the phase III model framework accommodated additional hierarchical nesting, with village-level variables representing level three covariates, household and respondent-level variables representing level two factors, and child-level variables representing level one covariates.

ADDITIONAL RESULTS

Univariate statistics

The India CE survey included 36 items for factoring (Table S2). The top five items to which respondents most frequently selected “completely agree” aligned for men and women, though there were some differences with regard to the proportions of those responses between sexes (Table S7). These items reflected those related to community attachment and vicarious affective feelings: “I feel proud to be part of this village” (96% of men, 94% of women); “Being a member of this village is part of my identity or how I identify myself” (95% of men, 94% of women); “I live in closeness with the village and its people” (95% of men, 93% of women); “I feel accepted as a member of this village” (94% of men, 93% of women); “I feel happy for my neighbour if they get some new work” (92% of men, 87% of women).

Three of the top five items to which respondents most frequently selected "completely disagree" aligned between men and women, however the proportions of those responses differed between sexes, as did the remaining two items. Overlapping items reflected those related to personal associational membership in community structures, intolerance, and availability of supporting networks for individual-level problem-solving: "I am an active member of at least one community organisation, such as a self-help group" (64% of men, 61% of women); "Differences between people, such as wealth and caste, often cause problems in this village" (45% of men, 37% of women); and "If you suddenly had to go away for a day or two, you could count on your neighbours to take care of your children" (33% of men, 49% of women). The remaining two items that men most frequently selected "completely disagree" reflected those related to social disorder and social inequity: "In this village, often crimes and violence occur" (41% of men); and "Sometimes people need to bribe village leaders in order to get paperwork or other things done" (38% of men), respectively. The remaining two items that women most frequently selected "completely disagree" reflected those related to personal involvement or participation in activities organised by community structures and perceived individual-level behavioural control over contribution to collective or group goal attainment: "I participate in voting and community meetings or other community activities like that in the village" (61% of women), "I have the capacity to build this village into a developed place" (48% of women).

Preliminary CFA results

Fit statistics of the initial baseline (i.e., models with no *post hoc* adjustments) and refined models (i.e., models with *post hoc* adjustments) identified via preliminary CFA of the initial hypothesised CE framework are displayed in Table S3. The non-refined, initial baseline model indicated moderately acceptable absolute fit, but poor incremental or relative fit (χ^2:df=3.984; RMSEA=0.041, 90% CI: 0.040, 0.043; CFI=0.826; TLI=0.809). These findings suggest that while the initial hypothesised CE framework was plausible, another model may provide a better fit.

We moved forward with *post hoc* model refinements to eliminate non-salient and non-significant factor loadings as well as any factors with insufficient component saturation. This resulted in the elimination of seven items. Five items were eliminated for non-salient factor loadings (OWNWELF=-0.092, RESTRSER=0.202, CRIMECON=-0.218, BRIBELDR=-0.226, SAFEATHO=0.298), and two items were eliminated because they had both a non-salient and non-significant loading on their designated factors (CHEATS=-0.011, p=0.838; DIFPROBS=0.033,

p=0.446). After eliminating items that were non-significant and non-salient, one factor (social disorder) remained with only one item, which is not sufficient for component saturation. That factor was therefore eliminated along with the remaining item, which otherwise demonstrated a salient and significant loading (HARMONY=0.704, p<0.001). The standardised estimates of the remaining factor loadings in the model were acceptable (Table S3). The refined preliminary CFA model of the hypothesised CE framework did not demonstrate adequate absolute or incremental model fit (χ^2:df=5.386; RMSEA=0.050, 90% CI: 0.048, 0.053; CFI=0.859; TLI=0.841). These results suggest that the initial hypothesised CE framework did not represent a plausible structure of the mechanisms through which the CE process operates amongst men and women in the rural Odisha context. We therefore moved on to perform EFA and CFA of the EFA-derived factor solution.

EFA and CFA of EFA-derived model results

Scale development and validation samples

Partitioning the data resulted in a split-half EFA sample of 867 observations, and a split-half CFA sample of 867 observations. The participant to item ratio was slightly over 24:1, which was more than sufficient according to standard guidance [37, 41, 50].

Factor extraction

Six factors were extracted during final EFA. The last factor extracted had an eigenvalue of 1.207; the first factor not retained had an eigenvalue of 1.039. While the first factor not retained had a value above the renowned 1.0 threshold (i.e., Kaiser-Guttman rule [eigenvalue > 1.0], Kaiser 1960), the retention of that factor was not warranted by strong substantive or statistical justification. Including that factor merely because its eigenvalue was slightly greater than 1.0 would go against our pre-analysis plan, which indicated that factor extraction decision rules would be based on heuristic and model fit criteria as well as other important theoretical and empirical considerations (e.g., results from cognitive interviews, pilot testing of the CE instrument and other prior theoretical and empirical evidence) [24, 37, 44]. The decision not to include such factors is supported by literature that indicates the Kaiser-Guttman rule tends to result in overfactoring or underfactoring because sampling error may influence eigenvalues [24, 51, 52].

EFA-derived factor solutions

Factors and items indicated by the EFA-derived factor solutions are presented in Table S4. Complete EFA results reflect coefficients from both rotated (Promax) pattern and structure

matrices along with initial and refined CFA results. While not all factor loadings demonstrated in Table S4 are in the range of excellent to very good, we hypothesise that some of those items are conceptually more distal (i.e., marginally less important) to the measurement of their respective latent factors. Further details regarding the measurement model is presented below.

EFA results revealed a six-factor CE measurement model with good model fit (χ^2:df=1.476; RMSEA=0.023, RMSR=0.041). During the EFA analyses, we eliminated seven items, in a step-wise manner, via subsequent model runs: four items were eliminated because they had no salient loadings on any factor (i.e., the following were "bad variables": ACTCBGP, HAPPYNEI, SIMHOPES, RESTRSER); three items (HLPCRPDZ, LOSTCOW, COMPRSLV) were eliminated because they were complex variables (i.e., they loaded to more than one factor), yet had little substantive justification for supporting those cross-loadings; and one item (OWNWELF) was eliminated because although its pattern coefficient (i.e., factor loading) was salient, its structure coefficient was not salient. This item reduction process resulted in a 28-item, six factor CE measurement model.

This EFA-derived CE measurement model reflected items that tapped to the following six factors, which can be described as reflecting the following dimensions of CE: social response, social disorder, agency, village leadership, social networks, and social attachment (belonging/place identity). These six factors can be conceived as corresponding to four domains: behavioural control, informal social control, structural and cognitive social capital, and social cohesion.

Factors one and two pertained to the informal social control domain, though factor one also tapped to concepts related to cognitive social capital (e.g., trust in village members). Factor one, labelled "social response" included four items that tapped to villagers' willingness to intervene when delinquent behaviour is observed and their propensity to address sub-village-level issues (i.e., engage in conflict-resolution), a sense of harmony or lack thereof in the village, and perceived trust in village members. Social response factor loadings ranged from 0.056-0.389 Factor two, labelled "social disorder" included four items that tap to concepts related to perceived presence of incivility or delinquent behaviour and deceitful individuals, intolerance, and corruption in the village. Factor loadings for this factor ranged from 0.645-0.500

Factor three, labelled "agency" corresponded to the behavioural control domain. On EFA, the factor contained four items that tap to individual and collective behavioural control, including perceived individual-level behavioural control regarding one's own ability to contribute to

collective goal attainment. The factor also included an item that tapped to the responsiveness of one's supporting network for individual-level problem solving specific to child caregiving responsibilities. Factor loadings for this factor ranged from 0.956-0.393.

Factor four, labelled "village leadership", corresponded to the structural social capital domain, though also tapped to aspects of cognitive social capital (e.g., trust, reciprocity) and social cohesion (e.g., social justice and equity). The factor included seven items that tap to various aspects of formal and informal village leadership, including trust in and level of support, activity of leaders; presence of informal leaders; equal distribution of resources, and decision-making that is pleasing for most villagers; similar collective morals; and individual contribution to collective goals, such as the development of the village. Factor loadings for this factor ranged from 0.966-0.301.

Factor five, labelled "social networks", pertained to the cognitive social capital domain. This factor contained two items that tapped to the availability and reciprocity of support networks for individual-level problem solving. Constituent items had factor loadings ranging from 0.873-0.667.

Factor six, labelled "social attachment", represented the social cohesion domain. This factor contained four items that tapped to the strength of social bonds and social acceptance within the village, place identity and sense of belonging to the village, and pride in being a member of the village. Social attachment factor loadings ranged from 0.987-0.820.

CFA results and post hoc model refinements of the EFA-derived CE measurement model

CFA was conducted on the 28 items tapping to six factors, as indicated by the EFA-derived factor solution. The initial, baseline run of CFA with no *post hoc* refinements demonstrated good model fit according to both absolute and incremental fit statistics (χ^2:df=2.127; RMSEA=0.036, 90% CI: 0.032, 0.040; CFI=0.921; TLI=0.911). *Post hoc* model refinements consisted of the elimination of SAFEATHO, the only item that with a non-salient loading on CFA (factor loading=0.235). The factor loadings for the resulting 27-item, six factor solution were acceptable (Table S4). With this one modification, the refined CFA model validated the majority of the EFA-derived factor structure, and demonstrated good absolute and incremental model fit (χ^2:df=2.215; RMSEA=0.037, 90% CI: 0.034, 0.041; CFI=0.923; TLI=0.912). Over half of the items (56%, 15 of 27) in the refined 27-item, six item factor solution demonstrated very good to excellent factor loadings (i.e., loadings > 0.630) on a single factor, with another 19% (5 of 27) of the factors demonstrating good factor loadings (i.e., loadings between 0.550-0.629). These

results suggest that the refined CE scale with *post hoc* adjustments demonstrated good construct validity for measuring CE in the rural Odisha context.

Comparison of hypothesised vs. empirically-derived factor solutions

Fit statistics from the preliminary CFA of our initial hypothesised CE framework and the CFA of the EFA-derived factor solution suggest that slight revisions that were substantively justified resulted in a CE measurement model that demonstrated good construct validity in the rural Odisha context (Tables S2 and S3).

MIMIC model results

The baseline MIMIC model with latent variables regressed on respondent gender and leadership covariates, but no direct effects between item indicators demonstrated good model fit (χ^2:df=2.123; RMSEA=0.036, 90% CI: 0.032, 0.040; CFI=0.916; TLI=0.903) (Table S5 and Baseline MIMIC Model 2 in Table S6). Upon examination of the modification indices from the baseline MIMIC model with latent factors regressed on respondent gender and leadership status, three items demonstrated modification indices above 3.84. The PARTCBGP indicator had the highest, albeit relative low modification index (31.465) on gender. This finding indicated that there was DIF between men and women. We believed there was justification for men and women having different levels of participation in village-based groups. We therefore added a direct path between PARTCBGP and gender, and re-ran the model.

On the model iteration with the direct path between PARTCBGP and gender, both unstandardized B and standardised β were salient (i.e., 0.781 and 0.695, respectively). These findings indicated that men scored PARTCBGP, on average, 0.7 units higher than women. This refined model fit the data well (χ^2:df=2.040; RMSEA=0.035, 90% CI: 0.031, 0.038; CFI=0.922; TLI=0.910), and DIFFTEST statistics indicated that model fit improved with the inclusion of this direct effects parameter. When examining the modification indices for this model, we noted that BORMONEY had a modification index above 3.85 on leadership status, and SELFEFF had a modification index above 3.85 on gender. We therefore fit a direct path between BORMONEY and leadership status, and both unstandardized B and standardised β were salient (i.e., -0.747, and -0.744, respectively). This indicated that individuals with leadership status scored BORMONEY -0.7 units lower than individuals with leadership status. We also moved forward with re-running the model after fitting a direct path between SELFEFF and gender, however, both unstandardized B and standardised β were just above the threshold for salient (i.e., 0.316 and 0.301, respectively), and relatively little was gained from adding this parameter ($\Delta\chi^2$ [df]

15). Therefore, the final MIMIC model reflects the model with latent variables regressed on gender and leadership status, and direct paths between PARTCBGP and gender, and BORMONEY and leadership (i.e., model 2b in Table S6).

Table S2. Administered collective efficacy tool, India

Domain	Factor	Item	Survey item (i.e., prompt)	Hypothesised facets tapped
Informal social control	Social disorder	HARMONY	Most of the people in this village stay in peace and harmony with each other.	Sense of **harmony** within the village
		CHEATS	*In this village, you have to be careful otherwise someone may cheat you.*	*Perceived presence of **deceitful people***
		CRIMECON	*In this village, often crimes and violence occur.*	*Perceived presence of **incivility, delinquent acts***
		SAFEATHO*	When I am at home alone, I feel safe from crimes and violence.	Feeling of **safety** while at home
	Social response	INTERCRI	If the people of this village see crime like activities then they will stop it.	Willingness to **intervene**
		COMPRSLV	When there is a village problem, people come together and share their opinions on how it should be solved.	Group **problem-solving, conflict-resolution**
		SLVDISPU	If there is a big fight between two persons, other people from the village will help in solving the fight.	Villager's **propensity to address sub-village**-level **issues; conflict-resolution**
		HLPCRPDZ	If there were a problem that affected the entire village, for instance crop disease, people in this village would help each other.	Community's **propensity to address village-wide issues; conflict-resolution**
		HAPPYNEI*	I feel happy for my neighbour if they get some new work.	**Vicarious affective feelings** - happiness
Social cohesion	Social capital	COMTRUST	People in this village can be trusted.	**Perceived trust** in village members
		CLOSE	I live in closeness with the village and its people.	**Strength of social bonds** within village
		LOSTCOW	If a neighbour loses a cow or buffalo, someone in the village would return it to them.	**Perceived responsiveness of social networks**, expectations that help will be given to/received by others when in need, **individuals cooperating to support one another** for either one-sided or mutual gain§
		BORMONEY	If you suddenly need some money, nearly to run a family expenditure for a week, then you could borrow from someone in your village who is not a money lender.	
		NEICAREG	If you suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	
		UNOFLDRS	There are people in this village who are unofficial leaders, in other words, people who care about their neighbours and speak out for them.	Perceived presence of **individuals demonstrating attributes of leadership**
		ACTLDR2‡	Village leaders do many things for this village.	**Supportiveness** of the leaders of **exogenous** village structures
		TRUSTLDR	Village leaders can be trusted.	Perceived **social trust in village leaders**
		HAVEFRND*	In this village, I have such friends with whom I can share my problems.	**Availability of support networks** for individual-level **problem-solving**
		COME4HLP*	My friends or my neighbours sometimes come to me with their problems for help.	**Reciprocity** of individual-level **problem-solving**
	Social equity	COMMGDEC	When the village committee makes decisions they are pleasing and good for most of the households.	**Social equity** prioritised during **village decision-making** processes
		BRIBELDR	*Sometimes people need to bribe village leaders in order to get paperwork or things done.*	***Corruption** amongst village leaders*
		DISTCRIS	During a crisis situation, such as crop failure, government services are distributed equally to all households in need.	**Equal distribution** of exogenous resources during crises
		CONTRDEV	The people of this village will contribute time OR money for the village development.	Common **moral principles** & **codes of behaviour**

		OWNWELF	*Here people only think about their own family's development or welfare, and they do not think about village's development.*	***Commitment to village** development, **goal attainment***
		RESTRSER	*Some households are restricted from the government services available in the village.*	***Social injustice, restrictions** to resources*
	Common values	SIMHOPES	Most people in this village have similar hopes about their future.	**Common hopes** for village **goal attainment**
		SIMBLIEF	Most people in this village have similar beliefs on what is right and what is wrong.	Collective **morals**
		DIFPROBS	Differences between people, such as wealth and caste, often cause problems in this village.	**Tolerance**
	Social attachment	ACCEPT	I feel accepted as a member of this village.	**Social acceptance** within the village
		PROUD*	I feel proud to be part of this village.	**Pride in being a member** of the village
		IDENTITY*	Being a member of this village is part of my identity or how I identify myself.	**Place identit**y, sense of **belonging**
Behavioural control	Agency	SELFEFF*	I have the capacity to control my future.	Perceived **individual-level behavioural control** over **personal goal attainment**
		SEDEV*	I have the capacity to build this village into a developed place.	Perceived **individual-level behavioural control** over contribution to village **goal attainment**
		COLLEFF	People in this village have the ability to come together and make a positive change.	Perceived village**-level behavioural control**; **capacity** and **autonomy** control beliefs
		ACTCBGP*†	I am an active member of at least one community organization such as a self-help group.	**Personal** associational **membership**/participation, **endogenous** village structures
		PARTCBGP*†	I participate in voting and community meetings or other community activities like that in the village.	**Personal** involvement/**participation** in activities organised by **endogenous** village structures

Notes. Indicated sub-constructs reflect those conceptualised via our initial hypothesised collective efficacy framework. Items presented in italicised text reflect hypothesised inverse relationship between the items and CE. Given the various conceptualisations of these latent constructs, substantive justification existed for the re-conceptualisation articulated in our EFA-derived factor structures. * self-referent item prompts about the respondent's own sense of self, agency, autonomy, and level of engagement within his/her village; all other items reflect group-referent items prompts about ecological aspects of the respondent's village; † items that refer to endogenous village structures (e.g., village-initiated/organised community associations) - local endogenous structures used as examples, but should be adapted to the given local context; ‡ items that refer to exogenous village structures (e.g., government or NGO-initiated/organised village associations) - local exogenous structures used as examples, but should be adapted to the given context; § measured through different scenarios reflecting different levels of urgency/need

Table S3. Factor loadings from preliminary CFA of initial hypothesised CE framework, India

Factors and associated items	Item	Initial preliminary CFA (N = 1,734)	Refined† preliminary CFA (N = 1,734)
Factor 1: Agency (average factor loading on final CFA = 0.626)			
People in this village have the ability to come together and make a positive change.	COLLEFF	0.783*	0.780*
I have the capacity to build this village into a developed place.	SEDEV	0.735*	0.742*
I have the capacity to control my future.	SELFEFF	0.612*	0.614*
I participate in voting and community meetings or other community activities like that in the village.	PARTCBGP	0.585*	0.578*
I am an active member of at least one community organization such as a self-help group.	ACTCBGP	0.413*	0.416*
Factor 2: Common values (average factor loading on final CFA = 0.608)			
Most people in this village have similar hopes about their future.	SIMHOPES	0.644*	0.643*
Most people in this village have similar beliefs on what is right and what is wrong.	SIMBLIEF	0.580*	0.572*
Differences between people, such as wealth and caste, often cause problems in this village.	DIFPROBS	0.033	-
Factor 3: Social capital (average factor loading on final CFA = 0.626)			
I live in closeness with the village and its people.	CLOSE	0.814*	0.822*
Village leaders can be trusted.	TRUSTLDR	0.793*	0.798*
Village leaders do many things for this village.	ACTLDR2	0.755*	0.759*
In this village, I have such friends with whom I can share my problems.	HAVEFRND	0.655*	0.659*
My friends or my neighbours sometimes come to me with their problems for help.	COME4HLP	0.643*	0.646*
People in this village can be trusted.	COMTRUST	0.629*	0.611*
There are people in this village who are unofficial leaders, in other words, people who care about their neighbours and speak out for them.	UNOFLDRS	0.553*	0.550*
If you suddenly had to go away for a day or two, you could count on your neighbours to take care of your children.	NEICAREG	0.524*	0.526*
If you suddenly need some money, nearly to run a family expenditure for a week, then you could borrow from someone in your village who is not a money lender.	BORMONEY	0.451*	0.452*
If a neighbour loses a cow or buffalo, someone in the village would return it to them.	LOSTCOW	0.448*	0.437*
Factor 4: Social equity (average factor loading on final CFA = 0.602)			
When the village committee makes decisions they are pleasing and good for most of the households.	COMMGDEC	0.667*	0.657*
The people of this village will contribute time OR money for the village development.	CONTRDEV	0.654*	0.642*
During a crisis situation, such as crop failure, government services are distributed equally to all households in need.	DISTCRIS	0.517*	0.508*
Sometimes people need to bribe village leaders in order to get paperwork or other things done.	BRIBELDR	-0.224*	-

Some households are restricted from the government services available in the village.	RESTRSER	0.200*	-
Here people only think about their own family's development or welfare, and they do not think about village's development.	OWNWELF	-0.091*	-
Factor 5: Social attachment (average factor loading on final CFA = 0.847)			
I feel proud to be part of this village.	PROUD	0.967*	0.968*
I feel accepted as a member of this village.	ACCEPT	0.826*	0.828*
Being a member of this village is part of my identity or how I identify myself.	IDENTITY	0.750*	0.746*
Factor 6: Social response (average factor loading on final CFA = 0.654)			
When there is a village problem, people come together and share their opinions on how it should be solved.	COMPRSLV	0.770*	0.778*
If there is a big fight between two persons, other people from the village will help in solving the fight.	SLVDISPU	0.704*	0.698*
If there were a problem that affected the entire village, for instance crop disease, people in this village would help each other.	HLPCRPDZ	0.682*	0.685*
If the people of this village see crime like activities then they will stop it.	INTERCRI	0.566*	0.565*
I feel happy for my neighbour if they get some new work.	HAPPYNEI	0.551*	0.542*
Factor 7: Social disorder			
Most of the people in this village stay in peace and harmony with each other.	HARMONY	0.737*	-
When I am at home alone, I feel safe from crimes and violence.	SAFEATHO	0.297*	-
In this village, often crimes and violence occur.	CRIMECON	-0.228*	-
In this village, you have to be careful otherwise someone may cheat you.	CHEATS	-0.011	-
Model fit statistics			
χ^2:df		3.984	5.386
RMSEA (90% CI)		0.041 (0.040-0.043)	0.050 (0.048-0.053)
CFI		0.826	0.859
TLI		0.809	0.841

Notes. ***Matrix:*** Polychoric correlations; ***Estimation method:*** WLSMV with sandwich estimator to adjust for non-independence of observations within 90 village clusters; ***Extraction:*** Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; ***Rotation:*** Promax. * $p \leq 0.05$; † Refined CFA reflects post hoc model adjustments, such as item reduction due to non-salient (loadings < 0.300) or non-significant (two-tailed $p > 0.05$) factor loadings. RMSEA=Root mean square error of approximation; RMSR=Root mean square residual; CFI=Comparative fit index; TLI=Tucker-Lewis index

Table S4. Factor loadings for random split-half samples for EFA and CFA of EFA-derived factor solutions, India

Factors and associated items	Item	Final EFA - Factor pattern coefficient (n_1 = 867)	Final EFA - Factor structure coefficient (n_1 = 867)	Baseline CFA (n_2 = 867)	Refined CFA† (n_2 = 867)
Factor 1: Social response (average factor loading = 0.593 ; average structure coefficient = 0.627 ; average factor loading on refined CFA = 0.691)					
If the people of this village see crime like activities then they will stop it.	INTERCRI	0.856	0.762	0.655*	0.651*
If there is a big fight between two persons, other people from the village will help in solving the fight.	SLVDISPU	0.754	0.777	0.728*	0.730*
Most of the people in this village stay in peace and harmony with each other.	HARMONY	0.518	0.630	0.617*	0.616*
When I am at home alone, I feel safe from crimes and violence.	SAFEATHO	0.450	0.385	0.235*	-
People in this village can be trusted.	COMTRUST	0.389	0.582	0.764*	0.766*
Factor 2: Social disorder (average factor loading = 0.585 ; average structure coefficient = 0.584 ; average factor loading on refined CFA = 0.480)					
In this village, often crimes and violence occur.	CRIMECON	0.645	0.648	0.754*	0.758*
Sometimes people need to bribe village leaders in order to get paperwork or other things done.	BRIBELDR	0.629	0.636	0.329*	0.332*
Differences between people, such as wealth and caste, often cause problems in this village.	DIFPROBS	0.564	0.547	0.394*	0.390*
In this village, you have to be careful otherwise someone may cheat you.	CHEATS	0.500	0.504	0.448*	0.441*
Factor 3: Agency (average factor loading = 0.641 ; average structure coefficient = 0.660 ; average factor loading on refined CFA = 0.641)					
I have the capacity to build this village into a developed place.	SEDEV	0.956	0.835	0.698*	0.699*
I have the capacity to control my future.	SELFEFF	0.783	0.670	0.610*	0.609*
People in this village have the ability to come together and make a positive change.	COLLEFF	0.568	0.654	0.760*	0.759*
I participate in voting and community meetings or other community activities like that in the village.	PARTCBGP	0.506	0.570	0.495*	0.496*
I am an active member of at least one community organization such as a self-help group.	NEICAREG	0.393	0.569	0.642*	0.642*
Factor 4: Village leadership (average factor loading = 0.558 ; average structure coefficient = 0.599 ; average factor loading on refined CFA = 0.606)					
Village leaders can be trusted.	TRUSTLDR	0.966	0.883	0.865*	0.865*
Village leaders do many things for this village.	ACTLDR2	0.916	0.836	0.807*	0.808*
The people of this village will contribute time OR money for the village development.	CONTRDEV	0.520	0.590	0.595*	0.595*
During a crisis situation, such as crop failure, government services are distributed equally to all households in need.	DISTCRIS	0.514	0.531	0.519*	0.519*
When the village committee makes decisions they are pleasing and good for most of the households.	COMMGDEC	0.454	0.571	0.663*	0.661*
There are people in this village who are unofficial leaders, in other words, people who care about their neighbours and speak out for them.	UNOFLDRS	0.404	0.530	0.567*	0.565*
Most people in this village have similar beliefs on what is right and what is wrong.	SIMBLIEF	0.385	0.439	0.404*	0.405*

If you suddenly need some money, nearly to run a family expenditure for a week, then you could borrow from someone in your village who is not a money lender.	BORMONEY	0.301	0.410	0.433*	0.431*
Factor 5: Social networks (average factor loading = 0.770 ; average structure coefficient = 0.842 ; average factor loading on refined CFA = 0.893)					
In this village, I have such friends with whom I can share my problems.	HAVEFRND	0.873	0.922	0.905*	0.905*
My friends or my neighbours sometimes come to me with their problems for help.	COME4HLP	0.667	0.762	0.881*	0.881*
Factor 6: Social attachment (average factor loading = 0.882 ; average structure coefficient = 0.887 ; average factor loading on refined CFA = 0.868)					
I live in closeness with the village and its people.	CLOSE	0.987	0.963	0.963*	0.963*
I feel proud to be part of this village.	PROUD	0.873	0.893	0.917*	0.917*
I feel accepted as a member of this village.	ACCEPT	0.849	0.853	0.841*	0.841*
Being a member of this village is part of my identity or how I identify myself.	IDENTITY	0.820	0.837	0.752*	0.750*
Model fit statistics					
χ^2:df			1.476	2.127	2.215
RMSEA			0.023	0.036	0.037
90% CI			N/A	(0.032-0.040)	(0.034, 0.041)
RMSR			0.041	N/A	N/A
CFI			N/A	0.921	0.923
TLI			N/A	0.911	0.912

Notes. ***Matrix:*** Polychoric correlations; ***Estimation method:*** WLSMV with sandwich estimator to adjust for non-independence of observations within 90 village clusters; ***Extraction:*** Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; ***Rotation:*** Promax. * $p \leq 0.05$; † Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.300) or non-significant (two-tailed $p > 0.05$) factor loadings. RMSEA=Root mean square error of approximation; RMSR=Root mean square residual; CFI=Comparative fit index; TLI=Tucker-Lewis index

Table S5. Factor loadings and fit indices for baseline and final MIMIC, and final CFA (with MIMIC refinement) models, India

Factors and associated items	Item	Baseline MIMIC model (n_2=867)	Final MIMIC model† (n_2=867)	Final CFA model‡ (n_2 = 867)
Factor 1: Social response (average final MIMIC model factor loading = 0.695 ; average final CFA model with MIMIC model deletions = 0.691)				
If the people of this village see crime like activities then they will stop it.	INTERCRI	0.657*	0.658*	0.658*
If there is a big fight between two persons, other people from the village will help in solving the fight.	SLVDISPU	0.733*	0.733*	0.730*
Most of the people in this village stay in peace and harmony with each other.	HARMONY	0.627*	0.627*	0.614*
People in this village can be trusted.	COMTRUST	0.761*	0.760*	0.763*
Factor 2: Social disorder (average final MIMIC model factor loading = 0.484 ; average final CFA model with MIMIC model deletions = 0.481)				
In this village, often crimes and violence occur.	CRIMECON	0.736*	0.736*	0.754*
Sometimes people need to bribe village leaders in order to get paperwork or other things done.	BRIBELDR	0.322*	0.322*	0.333*
Differences between people, such as wealth and caste, often cause problems in this village.	DIFPROBS	0.415*	0.415*	0.387*
In this village, you have to be careful otherwise someone may cheat you.	CHEATS	0.464*	0.464*	0.449*
Factor 3: Agency (average final MIMIC model factor loading = 0.618 ; average final CFA model with MIMIC model deletions = 0.677)				
I have the capacity to build this village into a developed place.	SEDEV	0.656*	0.684*	0.693*
I have the capacity to control my future.	SELFEFF	0.604*	0.597*	0.607*
People in this village have the ability to come together and make a positive change.	COLLEFF	0.760*	0.765*	0.776*
I participate in voting and community meetings or other community activities like that in the village.	PARTCBGP	0.500*	0.402*	-
I am an active member of at least one community organization such as a self-help group.	NEICAREG	0.645*	0.642*	0.630*
Factor 4: Village leadership (average final MIMIC model factor loading = 0.609 ; average final CFA model with MIMIC model deletions = 0.633)				
Village leaders can be trusted.	TRUSTLDR	0.870*	0.870*	0.872*
Village leaders do many things for this village.	ACTLDR2	0.819*	0.819*	0.816*
The people of this village will contribute time OR money for the village development.	CONTRDEV	0.601*	0.601*	0.590*
During a crisis situation, such as crop failure, government services are distributed equally to all households in need.	DISTCRIS	0.523*	0.523*	0.517*
When the village committee makes decisions they are pleasing and good for most of the households.	COMMGDEC	0.655*	0.656*	0.660*
There are people in this village who are unofficial leaders, in other words, people who care about their neighbours and speak out for them.	UNOFLDRS	0.554*	0.554*	0.562*
Most people in this village have similar beliefs on what is right and what is wrong.	SIMBLIEF	0.402*	0.403*	0.417*
If you suddenly need some money, nearly to run a family expenditure for a week, then you could borrow from someone in your village who is not a money lender.	BORMONEY	0.444*	0.448*	-

Factor 5: Social networks (average final MIMIC model factor loading = 0.895; average final CFA model with MIMIC model deletions = 0.893)				
In this village, I have such friends with whom I can share my problems.	HAVEFRND	0.899*	0.899*	0.905*
My friends or my neighbours sometimes come to me with their problems for help.	COME4HLP	0.890*	0.890*	0.881*
Factor 6: Social attachment (average final MIMIC model factor loading = 0.873 ; average final CFA model with MIMIC model deletions = 0.881)				
I live in closeness with the village and its people.	CLOSE	0.961*	0.961*	0.964*
I feel proud to be part of this village.	PROUD	0.918*	0.918*	0.917*
I feel accepted as a member of this village.	ACCEPT	0.855*	0.855*	0.893*
Being a member of this village is part of my identity or how I identify myself.	IDENTITY	0.757*	0.757*	0.751*
Model fit statistics				
χ^2:df		2.123	2.032	2.379
RMSEA (90% CI)		0.036	0.034	0.040
90% CI		(0.032-0.040)	(0.031-0.038)	(0.036-0.044)
CFI		0.916	0.923	0.924
TLI		0.903	0.911	0.913

Notes. ***Matrix:*** Polychoric correlations; ***Estimation method:*** WLSMV with sandwich estimator to adjust for non-independence of observations within 90 village clusters; ***Extraction:*** Combination of Kaiser-Guttman rule (i.e., eigenvalue > 1.0), scree test, goodness-of-fit indices, and substantive justification grounded in theoretical and empirical evidence; ***Rotation:*** Promax. * $p \leq 0.05$; † Refined CFA reflects *post hoc* model adjustments, such as item reduction due to non-salient (loadings < 0.300) or non-significant (two-tailed $p > 0.05$) factor loadings. RMSEA=Root mean square error of approximation; RMSR=Root mean square residual; CFI=Comparative fit index; TLI=Tucker-Lewis index. †Final MIMIC model reflects refined, parsimonious CE measurement model with latent variables regressed on gender & leadership status plus the inclusion of a direct path between PARTCBGP and gender, and BORMONEY and leadership status; ‡ Final CFA model reflects refined, parsimonious CE measurement model with PARTCBGP and BORMONEY deleted due to DIF.

Table S6. Competing MIMIC models: Fit statistics, unstandardised B estimates, standard errors, and standardised β estimates, India

Model	n	χ^2 (df)	Δ χ^2 (df)	RMSEA (90% CI)	CFI	TFI	B	S.E.	β
MIMIC models with gender covariate only									
Baseline MIMIC model 1 i.e., CE measurement model with ALL latent factors regressed on gender, no direct effects	867	2.165	N/A	0.037 (0.033-0.040)	0.919	0.907	-	-	-
Refined Model 1a i.e., ALL latent factors regressed on gender) + direct path between PARTCBGP item indicator and gender	867	2.077	56*	0.035 (0.032-0.039)	0.926	0.914	0.768*	0.099	0.685*
Refined Model 1b i.e., ALL latent factors regressed on gender + direct path between PARTCBGP, gender AND between SELFEFF, gender	867	2.067	14*	0.035 (0.031-0.039)	0.926	0.915	0.313*	0.083	0.300*
MIMIC models with gender and leadership status covariates									
More saturated Baseline MIMIC Model 2 i.e., model with ALL latent factors regressed on gender AND leadership, no direct effects between covariates & items	867	2.123	N/A	0.036 (0.032-0.040)	0.916	0.903	-	-	-
Refined Model 2a i.e., direct path between PARTCBGP and gender	867	2.040	57*	0.035 (0.031-0.038)	0.922	0.910	0.781*	0.099	0.695*
Refined Model 2b i.e., Model 2a with direct path between BORMONEY and leadership status	**867**	**2.032**	**15***	**0.034** **(0.031-0.038)**	**0.923**	**0.911**	**-0.747***	**0.195**	**-0.744***
Refined Model 2c i.e., Model 2b with direct path between SELFEFF and gender	867	2.023	14*	0.034 (0.031-0.038)	0.924	0.912	0.316*	0.086	0.301*
MIMIC models with gender, leadership status, and years respondent lived in village covariates									
More saturated Baseline MIMIC Model 3 i.e., model with ALL latent factors regressed on gender, leadership, AND years respondent lived in village	866	2.082	N/A	0.035 (0.032-0.039)	0.914	0.900	-	-	-

Notes. ***Matrix:*** Polychoric correlations; ***Estimation method:*** WLSMV with sandwich estimator to adjust for non-independence of observations within village clusters; Δ χ2 assessed via DIFFTEST; B: unstandardised estimate; S.E. = standard error; β: standardised estimate; * $p \leq 0.001$, though χ^2 statistics are sensitive to sample size, therefore, DIFFTEST statistics are likely to be significant with large samples. Baseline MIMIC model regressed latent variables on gender and leadership status; additional direct effect paths between gender and leadership covariates incorporated in subsequent models via step-wise forward selection based on the magnitude of the item indicator's modification index.

Table S7. Univariate descriptive statistics: Frequency of responses, by CE respondent gender

CE item	Item response	Aggregate N = 1,734		Women n_W = 1,123		Men n_M = 611	
HARMONY	***Most of the people in this village stay in peace and harmony with each other.***						
	Completely disagree	86	4.96%	59	5.25%	27	4.42%
	Mildly/partially disagree	15	0.87%	14	1.25%	1	0.16%
	Neither agree nor disagree	49	2.83%	37	3.29%	12	1.96%
	Mildly/partially agree	284	16.38%	197	17.54%	87	14.24%
	Completely agree	1300	74.97%	816	72.66%	484	79.21%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
CHEATS	***In this village, you have to be careful otherwise someone may cheat you.***						
	Completely disagree	247	14.24%	158	14.07%	89	14.57%
	Mildly/partially disagree	30	1.73%	21	1.87%	9	1.47%
	Neither agree nor disagree	30	1.73%	20	1.78%	10	1.64%
	Mildly/partially agree	204	11.76%	140	12.47%	64	10.47%
	Completely agree	1223	70.53%	784	69.81%	439	71.85%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
CRIMECON	***In this village, often crimes and violence occur.***						
	Completely disagree	603	34.78%	352	31.34%	251	41.08%
	Mildly/partially disagree	42	2.42%	26	2.32%	16	2.62%
	Neither agree nor disagree	70	4.04%	45	4.01%	25	4.09%
	Mildly/partially agree	393	22.66%	264	23.51%	129	21.11%
	Completely agree	626	36.10%	436	38.82%	190	31.10%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
SAFEATHO	***When I am at home alone, I feel safe from crimes and violence.***						
	Completely disagree	202	11.65%	154	13.71%	48	7.86%
	Mildly/partially disagree	14	0.81%	10	0.89%	4	0.65%
	Neither agree nor disagree	34	1.96%	25	2.23%	9	1.47%
	Mildly/partially agree	131	7.55%	93	8.28%	38	6.22%
	Completely agree	1353	78.03%	841	74.89%	512	83.80%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
INTERCRI	***If the people of this village see crime like activities then they will stop it.***						
	Completely disagree	214	12.34%	161	14.34%	53	8.67%
	Mildly/partially disagree	23	1.33%	17	1.51%	6	0.98%
	Neither agree nor disagree	40	2.31%	30	2.67%	10	1.64%
	Mildly/partially agree	180	10.38%	120	10.69%	60	9.82%
	Completely agree	1277	73.64%	795	70.79%	482	78.89%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%

SLVDISPU	***If there is a big fight between two persons, other people from the village will help in solving the fight.***						
	Completely disagree	146	8.42%	116	10.33%	30	4.91%
	Mildly/partially disagree	14	0.81%	10	0.89%	4	0.65%
	Neither agree nor disagree	17	0.98%	13	1.16%	4	0.65%
	Mildly/partially agree	110	6.34%	76	6.77%	34	5.56%
	Completely agree	1447	83.45%	908	80.85%	539	88.22%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
HLPCRPDZ	***If there were a problem that affected the entire village, for instance crop disease, people would help each other.***						
	Completely disagree	261	15.05%	193	17.19%	68	11.13%
	Mildly/partially disagree	30	1.73%	24	2.14%	6	0.98%
	Neither agree nor disagree	58	3.34%	51	4.54%	7	1.15%
	Mildly/partially agree	212	12.23%	139	12.38%	73	11.95%
	Completely agree	1173	67.65%	716	63.76%	457	74.80%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
CONTRDEV	***The people of this village will contribute time OR money for the village development.***						
	Completely disagree	286	16.49%	175	15.58%	111	18.17%
	Mildly/partially disagree	29	1.67%	24	2.14%	5	0.82%
	Neither agree nor disagree	104	6.00%	82	7.30%	22	3.60%
	Mildly/partially agree	233	13.44%	154	13.71%	79	12.93%
	Completely agree	1082	62.40%	688	61.26%	394	64.48%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
DIFPROBS	***Differences between people, such as wealth and caste, often cause problems in this village.***						
	Completely disagree	687	39.62%	414	36.87%	273	44.68%
	Mildly/partially disagree	39	2.25%	33	2.94%	6	0.98%
	Neither agree nor disagree	84	4.84%	69	6.14%	15	2.45%
	Mildly/partially agree	295	17.01%	186	16.56%	109	17.84%
	Completely agree	629	36.27%	421	37.49%	208	34.04%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
HAPPYNEI	***I feel happy for my neighbour if they get some new work.***						
	Completely disagree	62	3.58%	47	4.19%	15	2.45%
	Mildly/partially disagree	10	0.58%	9	0.80%	1	0.16%
	Neither agree nor disagree	24	1.38%	18	1.60%	6	0.98%
	Mildly/partially agree	94	5.42%	69	6.14%	25	4.09%
	Completely agree	1544	89.04%	980	87.27%	564	92.31%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%

COMTRUST	***People in this village can be trusted.***						
	Completely disagree	307	17.70%	243	21.64%	64	10.47%
	Mildly/partially disagree	36	2.08%	27	2.40%	9	1.47%
	Neither agree nor disagree	66	3.81%	52	4.63%	14	2.29%
	Mildly/partially agree	413	23.82%	274	24.40%	139	22.75%
	Completely agree	912	52.60%	527	46.93%	385	63.01%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
OWNWELF	***Here people only think about their own family's development or welfare, and they do not think about village's development.***						
	Completely disagree	369	21.28%	235	20.93%	134	21.93%
	Mildly/partially disagree	65	3.75%	41	3.65%	24	3.93%
	Neither agree nor disagree	106	6.11%	76	6.77%	30	4.91%
	Mildly/partially agree	318	18.34%	197	17.54%	121	19.80%
	Completely agree	876	50.52%	574	51.11%	302	49.43%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
LOSTCOW	***If a neighbour loses a cow or buffalo, someone in the village would return it to them.***						
	Completely disagree	105	6.06%	80	7.12%	25	4.09%
	Mildly/partially disagree	11	0.63%	9	0.80%	2	0.33%
	Neither agree nor disagree	43	2.48%	34	3.03%	9	1.47%
	Mildly/partially agree	73	4.21%	53	4.72%	20	3.27%
	Completely agree	1502	86.62%	947	84.33%	555	90.83%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%
BORMONEY	***If you suddenly need some money, nearly to run a family expenditure for a week, then you could borrow from someone in your village who is not a money lender.***						
	Completely disagree	260	14.99%	179	15.94%	81	13.26%
	Mildly/partially disagree	5	0.29%	4	0.36%	1	0.16%
	Neither agree nor disagree	13	0.75%	9	0.80%	4	0.65%
	Mildly/partially agree	135	7.79%	82	7.30%	53	8.67%
	Completely agree	1321	76.18%	849	75.60%	472	77.25%
	TOTAL	1734	100.00%	1123	100.00%	611	100.00%

5.5 Chapter references

1. Glanz, K. and D.B. Bishop, *The Role of Behavioral Science Theory in Development and Implementation of Public Health Interventions.* Annual Review of Public Health, 2010. **31**(1): p. 399-418.
2. Davis, R., R. Campbell, Z. Hildon, L. Hobbs, and S. Michie, *Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review.* Health Psychology Review, 2015. **9**(3): p. 323-344.
3. Routray, P., W.-P. Schmidt, S. Boisson, T. Clasen, and M.W. Jenkins, *Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study.* BMC Public Health, 2015. **15**(1): p. 880.
4. Sahoo, K.C., K.R.S. Hulland, B.A. Caruso, R. Swain, M.C. Freeman, P. Panigrahi, and R. Dreibelbis, *Sanitation-related psychosocial stress: A grounded theory study of women across the life-course in Odisha, India.* Social Science & Medicine, 2015. **139**: p. 80-89.
5. O'Reilly, K. and E. Louis, *The toilet tripod: Understanding successful sanitation in rural India.* Health & Place, 2014. **29**: p. 43-51.
6. Dreibelbis, R., P.J. Winch, E. Leontsini, K.R. Hulland, P.K. Ram, L. Unicomb, and S.P. Luby, *The Integrated Behavioural Model for Water, Sanitation, and Hygiene: a systematic review of behavioural models and a framework for designing and evaluating behaviour change interventions in infrastructure-restricted settings.* BMC Public Health, 2013. **13**(1): p. 1015.
7. Fuller, J.A. and J.N. Eisenberg, *Herd Protection from Drinking Water, Sanitation, and Hygiene Interventions.* Am J Trop Med Hyg, 2016. **95**(5): p. 1201-1210.
8. Garn, J.V., S. Boisson, R. Willis, A. Bakhtiari, T. al-Khatib, K. Amer, W. Batcho, P. Courtright, M. Dejene, A. Goepogui, K. Kalua, B. Kebede, C.K. Macleod, K.I.M. Madeleine, M.S.A. Mbofana, C. Mpyet, J. Ndjemba, N. Olobio, A.L. Pavluck, O. Sokana, K. Southisombath, F. Taleo, A.W. Solomon, and M.C. Freeman, *Sanitation and water supply coverage thresholds associated with active trachoma: Modeling cross-sectional data from 13 countries.* PLOS Neglected Tropical Diseases, 2018. **12**(1): p. e0006110.
9. Oswald, W.E., A.E. Stewart, M.R. Kramer, T. Endeshaw, M. Zerihun, B. Melak, E. Sata, D. Gessese, T. Teferi, Z. Tadesse, B. Guadie, J.D. King, P.M. Emerson, E.K. Callahan, D. Flanders, C.L. Moe, and T.F. Clasen, *Active trachoma and community use of sanitation, Ethiopia.* Bulletin of the World Health Organization, 2017. **95**(4): p. 250-260.
10. Bandura, A., *Self-efficacy mechanism in human agency.* American Psychologist, 1982. **37**(2): p. 122-147.
11. Bandura, A., *Exercise of Human Agency through Collective Efficacy.* Current Directions in Psychological Science, 2000. **9**(3): p. 75-78.
12. Bray, S.R., *Collective Efficacy, Group Goals, and Group Performance of a Muscular Endurance Task.* Small Group Research, 2004. **35**(2): p. 230-238.
13. Durham, C.C., D. Knight, and E.A. Locke, *Effects of Leader Role, Team-Set Goal Difficulty, Efficacy, and Tactics on Team Effectiveness.* Organizational Behavior and Human Decision Processes, 1997. **72**(2): p. 203-231.
14. Hodges, L. and A.V. Carron, *Collective efficacy and group performance.* International Journal of Sport Psychology, 1992. **23**(1): p. 48-59.
15. Reese, H., P. Routray, B. Torondel, G. Sclar, M.G. Delea, S.S. Sinharoy, L. Zambrano, B. Caruso, S.R. Mishra, H.H. Chang, and T. Clasen, *Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India.* BMJ Open, 2017. **7**(3).
16. Reese, H., P. Routray, B. Torondel, S.S. Sinharoy, H.H. Chang, and T.F. Clasen, *Effect of a combined household-level piped water and sanitation intervention on child diarrhea, acute respiratory infection, soil-transmitted helminth infection, and nutritional status:*

a matched cohort study in rural Odisha, India. Bulletin of the World Health Organization, 2018. **Submitted**.
17. Creswell, J.W. and V.L.P. Clark, *Designing and Conducting Mixed Methods Research*. 2007: SAGE Publications.
18. Spector, P., *Summated Rating Scale Construction*. 1992: Newbury Park, California.
19. Delea, M., G. Sclar, M. Woreta, R. Haardörfer, C. Nagel, B. Caruso, R. Dreibelbis, A. Gobezayehu, T. Clasen, and M. Freeman, *Collective Efficacy: Development and Validation of a Measurement Scale for Use in Public Health and Development Programmes.* International Journal of Environmental Research and Public Health, 2018. **15**(10): p. 2139.
20. Cogill, B., *Anthropometric indicators measurement guide*. 2003, Academy for Educational Development (AED): Washington, DC.
21. de Onis, M., A.W. Onyango, J. Van den Broeck, C.W. Chumlea, and R. Martorell, *Measurement and Standardization Protocols for Anthropometry Used in the Construction of a New International Growth Reference.* Food and Nutrition Bulletin, 2004. **25**(1_suppl1): p. S27-S36.
22. World Health Organization., *Catalouge of health indicators: a selection of important health indicators recommended by WHO programmes*, D.o.H.S.a.T. Assessment, Editor. 1996, World Health Organization: Geneva.
23. WHO/UNICEF, *Indicators for assessing infant and young child feeding practices*, in *Part 2: Measurement*. 2010, World Health Organization, UNICEF: Geneva, Switzerland.
24. Brown, T.A., *Confirmatory Factor Analysis for Applied Research*. 2006: Guilford Publications.
25. Woods, C.M., T.F. Oltmanns, and E. Turkheimer, *Illustration of MIMIC-Model DIF Testing with the Schedule for Nonadaptive and Adaptive Personality.* J Psychopathol Behav Assess Journal of Psychopathology and Behavioral Assessment, 2009. **31**(4): p. 320-330.
26. DiStefano, C., M. Zhu, and D. Minddrila, *Understanding and Using Factor Scores: Considerations for the Applied Researcher.* Practical Assessment, Research & Evaluation, 2009. **14**(20): p. 1-11.
27. Dickin, S., E. Bisung, and K. Savadogo, *Sanitation and the commons: The role of collective action in sanitation use.* Geoforum, 2017. **86**: p. 118-126.
28. Bisung, E., S.J. Elliott, C.J. Schuster-Wallace, D.M. Karanja, and A. Bernard, *Social capital, collective action and access to water in rural Kenya.* Social Science & Medicine, 2014. **119**: p. 147-154.
29. Bandura, A. and E.A. Locke, *Negative self-efficacy and goal effects revisited.* J Appl Psychol, 2003. **88**(1): p. 87-99.
30. Hathi, P., D. Spears, and D. Coffey, *Can collective action strategies motivate behaviour change to reduce open defecation in rural India?* Waterlines, 2016. **35**(2): p. 118-135.
31. Coffey, D. and D. Spears, *Child Height in India: Facts and Interpretations from the NFHS-4, 2015-16.* Economic & Political Weekly, 2018. **LIII**(30): p. 8794.
32. Black, R.E., L.H. Allen, Z.A. Bhutta, L.E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, and J. Rivera, *Maternal and child undernutrition: global and regional exposures and health consequences.* The Lancet, 2008. **371**(9608): p. 243-260.
33. Black, R.E., C.G. Victora, S.P. Walker, Z.A. Bhutta, P. Christian, M. de Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, R. Martorell, and R. Uauy, *Maternal and child undernutrition and overweight in low-income and middle-income countries.* The Lancet, 2013. **382**(9890): p. 427-451.
34. Pett, M.A., N.R. Lackey, and J.J. Sullivan, *Making sense of factor analysis the use of factor analysis for instrument development in health care research*. 2003, Thousand Oaks, Calif.; London: SAGE.

35. Yong, A. and S. Pearce, *A beginner's guide to factor analysis: Focusing on exploratory factor analysis.* Tutorials in Quantitative Methods for Psychology, 2013. **9**(2): p. 72-94.
36. Beatty, P.C. and G.B. Willis, *Research Synthesis: The Practice of Cognitive Interviewing.* Public Opinion Quarterly, 2007. **71**(2): p. 287-311.
37. Henson, R.K. and J.K. Roberts, *Use of Exploratory Factor Analysis in Published Research:Common Errors and Some Comment on Improved Practice.* Educational and Psychological Measurement, 2006. **66**(3): p. 393-416.
38. Baglin, J., *Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR.* Practical Assessment, Research & Evaluation, 2014. **19**(5).
39. Hu, L.t. and P.M. Bentler, *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.* Structural Equation Modeling: A Multidisciplinary Journal, 1999. **6**(1): p. 1-55.
40. Muthen, L.K. and B. Muthen, *Mplus : statistical analysis with latent variables : user's guide*. 2015.
41. Tabachnick, B.G. and L.S. Fidell, *Using multivariate statistics*. 2007, Boston: Pearson/Allyn & Bacon.
42. Yang, Y. and Y. Xia, *On the number of factors to retain in exploratory factor analysis for ordered categorical data.* Behavior Research Methods, 2015. **47**(3): p. 756-772.
43. Kaiser, H.F., *The Application of Electronic Computers to Factor Analysis.* Educational and Psychological Measurement, 1960. **20**(1): p. 141-151.
44. Hancock, G.R. and R.O. Mueller, *The reviewer's guide to quantitative methods in the social sciences*. 2010, New York: Routledge.
45. Costello, A.B. and J.W. Osborne, *Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis.* Practical assessment, research & evaluation, 2005. **10**(7): p. 1-9.
46. Thurstone, L.L., *Multiple-factor analysis*. 1947, Chicago; London: University of Chicago Press.
47. Schreiber, J.B., A. Nora, F.K. Stage, E.A. Barlow, and J. King, *Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review.* The Journal of Educational Research, 2006. **99**(6): p. 323-338.
48. Holland, P.W. and H. Wainer, *Differential item functioning*. 1993, Hillsdale: Lawrence Erlbaum Associates.
49. Bentler, P.M. and D.G. Bonett, *Significance tests and goodness of fit in the analysis of covariance structures.* Psychological Bulletin, 1980. **88**(3): p. 588-606.
50. Stevens, S.S., *On the Theory of Scales of Measurement.* Science, 1946. **103**(2684): p. 677-80.
51. Comrey, A.L. and H.B. Lee, *A first course in factor analysis*. Second Edition ed. 1992, Hillsdale, NJ: Lawrence Erlbaum.
52. Gorsuch, R.L., *Factor analysis.* 2015.

EPILOGUE

Chapter 6: Discussion and reflections

6.1 Chapter overview

Prior chapters of this thesis have provided an overview of common social constructs and examined how the unique attributes of each can and should be considered for incorporation in the design, targeting, and evaluation of community-based interventions predicated on collective action. Information was provided on how these constructs represent interpersonal factors of collective behaviour and behaviour change, and how collective efficacy remains under-utilised in the WASH sector, and public health and development practice more broadly.

In order to demonstrate the utility of collective efficacy in informing intervention content and implementation approaches, targeting of intervention activities, and diagnostic evaluation of community-based interventions, work from two separate WASH evaluations was presented. Under the auspices of these studies, context-specific CE scales were developed, and construct validity was examined. One CE scale was then employed to demonstrate how the CE scales may be used to examine associations between CE-related factors and the uptake of improved WASH facilities coverage and utilisation, and downstream nutritional outcomes amongst children aged less than five years. This chapter will synthesise findings from the thesis and reflect on their contribution to the existing knowledge base, the limitations of this body of work, and what could have been done to improve the research.

6.2 Discussion of key research findings

This thesis sought to examine characteristics of the various social constructs commonly identified as theoretically-grounded and evidence-based predictors of cooperative behavior and collective action. Findings from this thesis can be used to inform future considerations related to the incorporation of these constructs in the design, targeting, and evaluation of community-based interventions. Enhancing the incorporation of social constructs in general, and collective efficacy in

particular may serve to elucidate and address persistent shortcomings in behaviour change interventions, and may help explain why health and development gains related to some interventions have been lower than expected [1-3].

6.2.1 Social constructs conceptualisations, theoretical underpinnings, and attributes

Findings from Research Paper 1 highlight the unique contributions of various social constructs. While each of the constructs examined provides its own important contributions, these findings suggest that efficacy-related constructs allow for examinations of factors known to predict cooperative behaviour above and beyond those provided by social capital and social cohesion constructs alone. This is important and relevant to note given the volume of work that has focused exclusively on examinations of social capital as the interpersonal factor influencing the uptake and effectiveness of community-based interventions, particularly in the WASH sector [4-6]. In assessments of social capital, proponents of the construct indicate that their findings suggest that social capital may be "necessary but insufficient" for development programmes, in general [7], but especially "for improving access to water and sanitation in marginalised communities" [5].

Efficacy assessments provide information related to perceived capability and autonomy to pursue common goals/tasks, motivational factors related to goal/task setting, performance during pursuit thereof, and the amount of effort and persistence that will be exerted in pursuit of the goal/task. This information demonstrates the construct's utility, and provides justification for its inclusion in community-based interventions predicated on collective action.

6.2.2 Collective efficacy scale development and validation: Findings from two contexts

This thesis presents findings from two separate investigations of CE in Ethiopia and India. These studies reflect an iterative approach to the development of context-specific CE tools. As a result of the desire to contextualise the scales, each study generated distinct CE factor solutions. The Ethiopia study was designed to allow a more thorough examination of gender-specific mechanisms through which CE operates amongst men and women in rural Ethiopia. This was deemed necessary and important given the results of the CE work would be used to inform the design of the content and implementation strategies employed in the intervention we were designing to be evaluated via the RCT.

Findings from the exploratory analyses performed on both Ethiopian and Indian datasets suggest that CE is a complex, multi-dimensional social construct. Both studies produced a six-factor CE measurement model. While the total number of factors and items reflected in these scales are similar, the underlying structures (number and nature of items) differed, as did the number of sub-construct domains to which the factors and related items tapped. The parsimonious CE scale developed in Ethiopia represented a 26-item, six-factor measurement model that tapped to four CE sub-construct domains. The CE scale developed in India represented a 25-item, six-factor measurement model that tapped to five CE sub-construct domains.

6.2.2.1 Which collective efficacy factors seem to transcend contexts

Several domains, factors, and related items appear to transcend contexts. While I will reflect on commonalities arising from the two investigations carried out in the context of this thesis, the small number of studies (n=2) should be noted. Future work on collective efficacy being carried out elsewhere can provide further evidence that can expand available evidence.

The parsimonious Ethiopia and India CE measurement models shared several CE domains – both structures had factors that tapped to social cohesion, informal social control, and cognitive and socio-structural social capital. In addition to these domains, the India CE structure had one factor – agency – that tapped to the behavioural control domain. Four of the six factors in both Ethiopia and India models were also shared. CE-factors related to social response, social networks, social attachment, and leadership appear to transcend at least the rural Amhara, Ethiopia and rural Odisha, India contexts. It is worth noting, however, that the nature of these factors differed slightly in each context (e.g., the social networks factor in Ethiopia also reflected aspects of personal agency, thereby indicating that personal agency was derived from perceptions regarding one's social networks). Seven items (i.e., facets) emerged from both CE measurement models as well. These include: SLVDISPU, COMTRUST (i.e., social response items); UNOFLDRS, ACTLDR2 (i.e., leadership items); ACCEPT, IDENTITY (i.e., social attachment items); and COME4HLP (i.e., social networks item).

The considerable proportion of overlap in construct domains (80% of the India CE domains, 100% of the Ethiopia CE domains) and factors (66% of factors in both CE measurement models) suggests that the CE survey is measuring the same construct in Ethiopia and India. These findings also indicate that the higher level sub-constructs are consistent.

6.2.2.2 How do these findings corroborate existing evidence

The findings presented in this thesis corroborate findings from other investigations of CE, though our measurement models suggest more comprehensive alternatives compared to previous theoretical, empirical, and psychometric examinations. Other psychometric examinations of CE support our inclusion of belonging (i.e., social attachment), association (i.e., associational participation), and social disorder [8, 9]. Our findings suggest that social cohesion and informal social control domains proposed by authors of these prior examinations are important for the measure of CE, but do not necessarily manifest as factors themselves [9]. Similarly, aspects of activism and informedness, two factors suggested by one study [8] may not have presented as factors in our models, but these concepts were represented by specific items included in the Ethiopia and India CE measurement models.

With regard to specific factors emerging from the CE measurement models and associations between CE and uptake of improved WASH practices, leadership emerged as a particularly important CE-related factor. This finding is supported by existing literature, which suggests that leaders can serve as catalysts to initiate new practices in the community. Other studies suggest that involving community opinion leaders can make an intervention more effective, and community leadership may play a role in ensuring the sustainability of collective management of sanitation [6].

It was interesting to note that reciprocity of one's networks with regard to caregiving was included in agency-related factors in both Ethiopia and India CE measurement models. While this finding makes sense anecdotally, it is also supported by other evidence, as indicated in the "household production of health" [10], and the idea that time allocation factors limit an individual's available time to practice some behaviours [11]. This may be particularly true of behaviours that contribute toward a collective goal as opposed to an individual goal.

In the India CE measurement model and application thereof, social disorder proved to be an important CE-related factor. This was not the case in Ethiopia. However, it is worth noting the history of caste-based intolerance and political economy of India when considering these findings. Social disorder has been noted to be a factor related to CE and social capital in other examinations of social constructs. In a study investigating violent crime in urban Chicago, Illinois, USA [9], and water access in rural Usoma Kenya [5], social cohesion and social disorder were implicated as important factors related to collective efficacy and collective action, respectively.

Findings from this thesis align with the findings from a predictive model of communication for water treatment and safe water storage and a related literature review of the social, cultural, and behavior correlates [12]. In that work, community leadership for collective action, social cohesion, community action and resources (e.g., social ogranisations), goal-setting, and collective efficacy and control are indicated as community factors that may play a significant role in the adoption of new behaviours. The authors of the work reflect on findings from their literature review regarding how contextual factors such as gender roles and gender equity in mobility, decision-making, and access to resources and services may influence household behaviours [12].

6.3 Contribution of key thesis findings to the advancement of public health and development practice

In many sectors, including but not limited to WASH, community-based programmes that target higher-order groups (e.g., households, villages, health centres, government ministries) often inadequately address factors of collective behaviour in their intervention design and implementation strategies [13]. Efficacy-related factors are likely important to consider and incorporate into community-based programmes, as they represent proximal factors of intentions and behaviour, according to the Theory of Triadic Influence [14]. The studies carried out under the auspices of this thesis research contribute to existing knowledge regarding the underlying structure of CE. Elucidating factors related to CE and examining their constituent sub-constructs is important to consider for the design of intervention content and development of implementation strategies. This work also contributes to the development of a metric that can be used in community-based health and development programmes, to inform intervention design, identify communities ready for programmatic targeting, and diagnose factors related to intervention effectiveness. Materials (i.e., survey tools, CE scales) produced as a result of this body of work contribute to the pool of knowledge and tools available to examine and address interpersonal factors that may affect the uptake and effectiveness of community-based interventions that target collective behaviours.

6.3.1 Informing intervention content

Failing to acknowledge and address factors related to CE as part of community-based intervention approaches may be problematic for the uptake of such interventions. The collective efficacy scales and larger surveys presented in this thesis can be adapted and deployed to help inform intelligent

intervention design (i.e., the design of evidence-based interventions addressing context-specific barriers to the sustainable adoption of improved behaviours through the employment of appropriate intervention techniques). The data generated from these instruments can be used to inform various aspects of intervention content, from its goals or motto to the motives and intervention techniques that the intervention will leverage to bring about and sustain change. Such tools fill a gap in terms of the number and nature of instruments available for quantitative assessments of proximal interpersonal behavioural factors.

In order to maximise the collective power of a group and enhance innate perceptions regarding its ability and autonomy to come together to pursue a communal goal, motivational aspects should be incorporated into a community-based programme from its outset. A programme's goal or motto, and accompanying messaging should leverage motives important to the group, and serve to enhance collective efficacy perceptions (e.g., *Together we can make develop our community*, *Together we can make a change*). The collective efficacy scales, and larger surveys presented in this thesis can be adapted and deployed to obtain pre-intervention assessments of collective efficacy and other related social constructs. Resulting data can help determine the magnitude of innate CE a community or group maintains, and pinpoint specific factors contributing to low collective efficacy perceptions. These data can be leveraged to conduct formal behavioural diagnostic examinations and intervention mapping. Such design measures can help ensure interventions are evidence-based, and employ techniques specifically devised to address context-specific barriers to improved practices and collective behaviours. The specific intervention techniques incorporated in the intervention should reflect those that have demonstrated success in addressing attributes related to the interpersonal behavioural factors targeted by the programme, as outlined in existing literature [15]. For instance, to address beliefs about efficacy judgements, intervention techniques such as verbal persuasion, mastery experiences, and vicarious reinforcement can be used. Similarly, goal setting, action planning, performance monitoring, problem solving/coping planning, and behavioural contracts can be incorporated in an intervention to address motivation and behavioural regulation [16, 17].

6.3.2 Enhancing implementation approaches

Findings from the gender-specific investigations of CE in India, and more thorough examinations in Ethiopia indicate that there are differentials in behavioural control, efficacy, and agency perceptions amongst men and women. Factors contributing to these differentials are well cited in

the literature [18, 19], and align with complementary findings emerging from our data (e.g., lower perceptions with regard to availability and demonstrated reciprocity of social networks; social bonds, their own social acceptance and sense of belonging; participation in community structures compared to men). The CE surveys and scales presented in this thesis can be employed by others in a similar manner to examine whether there are disparities in CE perceptions amongst certain sub-groups that should be leveraged to enhance implementation approaches. These disparities should be acknowledged and addressed not only in the design of intervention content, but also the implementation strategies employed by community-based interventions.

In situations in which the social inclusion of women in formal and informal community structures is limited, programmes can consider the incorporation of specific intervention activities that could create opportunities for women to become more engaged (e.g., establishment or enhancement of women's groups, with representation within broader community associations). Such interventions can present opportunities for women to socialise with peers, move outside of the home, and become exposed to more information, ideas, and social network contacts. These opportunities can minimise negative implications on women's social network perceptions when their sanitation and hygiene practices are restricted to the household setting.

Perceptions regarding one's social networks as a facilitator of efficacy perceptions was an important finding of this thesis research. The role of social support networks should be considered and mainstreamed within in the design and implementation strategies for community-based programmes predicated on cooperative behaviour and collective action. Such sentiments have been articulated in examinations of social capital, where it was noted that enhancement of these networks is often a missing component of community-based programmes, yet is necessary to facilitate collective action capital [5].

6.3.3 Improving targeting of participant communities, change agents, and programme participants

The scales and surveys generated by this thesis are also well suited for improving targeting of participant communities, change agents, and programme participants. The CE instruments can be adapted and incorporated into rapid community appraisals during the pre-intervention planning phase of a programme or project. Data generated from these assessments can be used to identify communities ready for change (i.e., those with high in innate CE relative to other communities – in

other words, those with innately strong positive perceptions about their ability and autonomy to catalyse and sustain change in their communities).

Results from this thesis research suggest that gender-specific disparities in behavioural control and agency perceptions exist. The CE scales presented herein can be used to determine whether these findings hold in other contexts. If such disparities exist, they should be considered when identifying appropriate intervention delivery modalities (i.e., change agents) and mechanisms (i.e., the structures and techniques through which interventions are delivered). In other words, young women with external loci of control (i.e., those who do not believe they have control over their lived experiences, their behaviours cannot change their lots in life) may not be the most appropriate agents to incite change amongst their peers, adolescent boys, and adult men. Similarly, individual and household-level counselling visits may reflect more suitable intervention delivery options than those presented by community-based associations or organisations in which women are not well represented.

While the findings presented in this thesis highlight behavioural control disparities between men and women, and, in some cases, leaders and non-leaders, there may be other marginalised sub-groups that should be considered when targeting programme participants. For instance, in India, some castes and other categories of people (e.g., those practising minority religions) may be marginalised. According to social control theory, this may mean that they feel "uninvolved with, uncommitted to, or alienated from conventional society" [20]. In turn, they may not internalise conventional values or standards for conventional behaviour. This is an issue that needs to be addressed in order to improve uptake of WASH facilities coverage and utilisation amongst these groups.

6.3.4 Diagnosing intervention effectiveness

The CE scales generated by this thesis offer new tools for the examination of interpersonal behavioural factors operating at group and community levels. These tools can be used to conduct diagnostic investigations into the role CE factors play in the uptake of community-based interventions and their impacts on health and development. They also facilitate the generation of evidence related to factors falling along the causal chain, which may explain why biologically plausible health gains are not being achieved by WASH interventions, as expected [1-3]. When

experimental study designs are used, these CE tools can be used to examine causal associations between CE and community-based interventions.

6.4 Strengths and limitations of the thesis research

6.4.1 Strengths

The scale development approach employed during these CE studies was a strength in that it reflected a mixed-methodological process that leveraged findings from social, behaviour, and behaviour change theory and empirical testing thereof, and included cognitive validations of both instruments. I actively reflected on qualitative data from cognitive interviews (India and Ethiopia) and focus group discussions (Ethiopia), and other theoretical and empirical evidence to make evidence-based, substantively justified modelling decisions. That said, our own qualitative evidence aside, there is a dearth of existing theoretical and empirical evidence on collective efficacy in low literacy and resource poor settings such as those in which this thesis research was conducted. The resulting CE measurement models derived through this approach therefore reflect rigorously derived models that were grounded in theory and evidence-based.

6.4.2 Limitations

This thesis research had several procedural and analytical limitations. Perhaps the biggest limitation of this body of work was the limited inclusion of indicators related to certain motivational components of CE in the India and Ethiopia surveys, and in turn, CE scales. Items that would allow for an examination of facets such as perceptions related to autonomy and controllability [21], past performance attainment, vicarious experiences, goal setting/pursuit and efficacy expectations [22, 23] were not included to the extent they could have been. This was due, in large part, to field realities (i.e., the need to strike a balance between survey content and length to ensure data quality and stay within budget). From the outset of this work, an indirect approach that considered CE perceptions in the context of related interpersonal factors, such as social capital, cohesion, and control was taken to measure CE appraisals via multi-dimensional factor structures. This came at the cost of more direct measures of specific efficacy-related facets.

The size of our women sub-sample in Ethiopia may have been a limitation. While gender-specific split-half sub-samples met common sample size guidelines, our sample size-dependent model fit

statistics (e.g., χ^2 p-value) indicated the sample sizes may have only been borderline sufficient for a factor structure as complex as those related to our CE measurement models.

With regard to component saturation, one of the items in the India CE measurement model, social networks, only contained two items. This was considered sufficient given the nature of the items and the magnitude of their factor loadings (each was high, and greater than 0.800). Other structural equation models examining similar social constructs generated by experts in the field of factor analytics have included factors with a similar number of items [24], so justification exists for this decision, particularly when the items demonstrate such high factor scores.

More generally, when compared to the Ethiopia model, and gold-standard guidelines for factor loadings (independently, and on average across all items in tapping to the factor), the India CE measurement model did not perform as well (i.e., factor loadings, while salient, were lower; absolute and comparative model fit statistics, though acceptable, were lower). This may have been an artefact of the shorter, less comprehensive CE survey that was administered in India (36 survey prompts compared to 50 survey prompts administered in Ethiopia). Had these extra items been included in the India survey, the resulting measurement model may have been different. That said, India provided a learning opportunity, and lesson learnt from that study were applied to the Ethiopia CE study. The study team was also provided more time and a larger budget to conduct the CE formative work, which included information generated from focus group discussions (not included in the India work). It is not uncommon, however, in scale development and refinement for subsequent iterations of a scale to build off of previous iterations.

The study design permitted in the examination of the relationship between CE and the uptake and influence of the Gram Vikas programme was a limitation. Given no pre-intervention CE (or related construct) measures were obtained prior to programme implementation, as a result, the post-only cross-sectional study design did not allow for investigation of causal relationships between CE and intervention effectiveness or directionality thereof. In other words, I was unable to determine innate CE perceptions prior to the intervention, or ascertain whether CE perceptions were enhanced via MANTRA. Future investigations using experimental study designs should be carried out in order to examine causal association and facilitate path analyses. This work is planned at the conclusion of the *Andilaye* trial (Ethiopia) in summer of 2019.

6.5 Reflections on what could have been done to improve the research

Several modifications could have improved this research. First, the administration of a more comprehensive tool, especially in the early (India) investigation would have afforded insights into other possible survey items and potential CE factors. A tool that included prompts probing further into specific motivational concepts (e.g., past group performance attainment, goal setting, vicarious community-based WASH programme successes in nearby villages, confidence in the group's ability to achieve its goals) may have enhanced these CE investigations, and could have influenced the results of the psychometric assessments in both India and Ethiopia. However, given this investigation into CE was only one of numerous areas of inquiry for our larger experimental studies, we were also limited with regard to survey length (i.e., timing of administration). As a result, and because we did not want to lean too heavily on one existing theory over others, the conceptual saturation of our CE surveys may have been limited to a certain extent.

While collective efficacy was the focus of this thesis, a concerted effort was made to consider various levels of motivational concepts (e.g., behavioural control, efficacy, agency). This effort was reflected in the hypothesised CE frameworks that considered behaviour control as a CE domain, with dimensions and facets related to individual and collective agency and efficacy perceptions. However, I could have done a better job of ensuring that CE survey prompts explored the nuances between these concepts. The primary barriers to moving forward in this manner were linguistic/translation limitations and comprehension amongst respondents with little to no formal education who were not used to responding to surveys using Likert-type responses. For example, several survey prompts that were developed for the India and Ethiopia studies were dropped because either: 1) there was no local word to reflect the nuances between statements that tapped to similar, yet theoretically different concepts, 2) some questions that were appropriately translated were not well understood by respondents during cognitive interventions and survey piloting. Further formative investigations into survey prompts that examine various concepts related to these motivational factors should be prioritised in future CE investigations.

Some of the WASH indicators employed in the evaluation of the Gram Vikas MANTRA intervention did not reflect recent learning in the sector. For example, while sanitation practices were captured on all individuals within study households, the major indicator used to assess these practices

inquired about usual place of defecation. Findings from validations of respondent-reported measures suggest that not only is this metric not precise, it is also less accurate, compared with sensor-recorded latrine use, than metrics that inquire about defecation practices during the 48-hours prior to survey administration [25]. It is possible that the limitations of these metrics could have influenced the results of the findings presented in Research Paper 3. Using WASH metrics that are more precise, which evidence suggests are more accurate, could have improved this research.

6.6 Chapter references

1. Clasen, T., S. Boisson, P. Routray, B. Torondel, M. Bell, O. Cumming, J. Ensink, M. Freeman, M. Jenkins, M. Odagiri, S. Ray, A. Sinha, M. Suar, and W.-P. Schmidt, *Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial.* The Lancet Global Health, 2014. **2**(11): p. e645-e653.
2. Luby, S.P., M. Rahman, B.F. Arnold, L. Unicomb, S. Ashraf, P.J. Winch, C.P. Stewart, F. Begum, F. Hussain, J. Benjamin-Chung, E. Leontsini, A.M. Naser, S.M. Parvez, A.E. Hubbard, A. Lin, F.A. Nizame, K. Jannat, A. Ercumen, P.K. Ram, K.K. Das, J. Abedin, T.F. Clasen, K.G. Dewey, L.C. Fernald, C. Null, T. Ahmed, and J.M. Colford, Jr., *Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial.* The Lancet Global Health, 2018. **6**(3): p. e302-e315.
3. Null, C., C.P. Stewart, A.J. Pickering, H.N. Dentz, B.F. Arnold, C.D. Arnold, J. Benjamin-Chung, T. Clasen, K.G. Dewey, L.C.H. Fernald, A.E. Hubbard, P. Kariger, A. Lin, S.P. Luby, A. Mertens, S.M. Njenga, G. Nyambane, P.K. Ram, and J.M. Colford, Jr., *Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial.* The Lancet Global Health, 2018. **6**(3): p. e316-e329.
4. Cameron, L.A., S. Olivia, and M. Shah, *Initial Conditions Matter: Social Capital and Participatory Development.* IZA Discussion Papers, 2015. **No. 9563**.
5. Bisung, E., S.J. Elliott, C.J. Schuster-Wallace, D.M. Karanja, and A. Bernard, *Social capital, collective action and access to water in rural Kenya.* Social Science & Medicine, 2014. **119**: p. 147-154.
6. Dickin, S., E. Bisung, and K. Savadogo, *Sanitation and the commons: The role of collective action in sanitation use.* Geoforum, 2017. **86**: p. 118-126.
7. Ostrom, E., *Social capital : a fad or a fundamental concept?*, in *Social capital : a multifaceted perspective.*, P. Dasgupta and I. Serageldin, Editors. 2000, World Bank Publications: Washington, D.C. p. 172-214.
8. Carroll, J.M., M.B. Rosson, and J. Zhou, *Collective efficacy as a measure of community*, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2005, ACM: Portland, Oregon, USA. p. 1-10.
9. Sampson, R.J., S.W. Raudenbush, and F. Earls, *Neighborhoods and violent crime: a multilevel study of collective efficacy.* Science, 1997. **277**(5328): p. 918-24.
10. Schumann, D.A. and W.H. Mosley, *The household production of health: Introduction.* Social Science & Medicine, 1994. **38**(2): p. 201-204.
11. Heckman, J., *Introduction to A Theory of the Allocation of Time by Gary Becker.* The Economic Journal 2015. **125**(583): p. 403-409.
12. Figueroa, M.E. and D.L. Kincaid, *Social, Cultural and Behavioral Correlates of Household Water Treatment and Storage*, in *Health Communication Insights*. 2010, Johns Hopkins Bloomberg School of Public Health, Center for Communication Programs: Baltimore.
13. Delea, M.G., H. Solomon, A.W. Solomon, and M.C. Freeman, *Interventions to maximize facial cleanliness and achieve environmental improvement for trachoma elimination: A review of the grey literature.* PLOS Neglected Tropical Diseases, 2018. **12**(1): p. e0006178.
14. Flay, B. and J. Petraitis, *The Theory of Triadic Influence: A New Theory of Health Behavior With Implications for Preventive Interventions.* Advances in Medical Sociology, 1994. **4**: p. 19-44.

15. Michie, S., M.M. van Stralen, and R. West, *The behaviour change wheel: A new method for characterising and designing behaviour change interventions.* Implementation Science, 2011. **6**(1): p. 42.
16. Michie, S. and M.A. West, *Managing people and performance: an evidence based framework applied to health service organizations.* International Journal of Management Reviews, 2004. **5-6**(2): p. 91-111.
17. Michie, S., M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M.P. Eccles, J. Cane, and C.E. Wood, *The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions.* Ann Behav Med, 2013. **46**(1): p. 81-95.
18. Mabsout, R. and I. van Staveren, *Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia.* World Development, 2010. **38**(5): p. 783-796.
19. Folbre, N., *Who pays for the kids? : gender and the structure of constraint*. 1994, London; New York: Routledge.
20. Petraitis, J., B.R. Flay, and T.Q. Miller, *Reviewing theories of adolescent substance use: Organizing pieces in the puzzle.* Psychological Bulletin, 1995. **117**(1): p. 67-86.
21. Ajzen, I., *Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior.* Journal of Applied Social Psychology, 2002. **32**(4): p. 665-683.
22. Bandura, A., *Self-efficacy: Toward a unifying theory of behavioral change.* Psychological Review, 1977. **84**(2): p. 191-215.
23. Bandura, A., *Social foundations of thought and action : a social cognitive theory*. 1986, Englewood Cliffs, N.J.: Prentice-Hall.
24. Steinmetz, H., E. Davidov, and P. Schmidt, *Three Approaches to Estimate Latent Interaction Effects: Intention and Perceived Behavioral Control in the Theory of Planned Behavior.* Methodological Innovations Online, 2011. **6**(1): p. 95-110.
25. Sinha, A., C.L. Nagel, E. Thomas, W.P. Schmidt, B. Torondel, S. Boisson, and T.F. Clasen, *Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors.* The American journal of tropical medicine and hygiene, 2016. **95**(3): p. 720-727.

Chapter 7: Conclusions and implications

7.1 Conclusions and recommendations for research and practice

The overarching goal of this doctoral research was to produce evidence-based recommendations to strengthen community-based programmes. As a result, the findings reflected in this thesis highlight implications for the design, targeting, and evaluation of community-based programme interventions, particularly those that are predicated on collective action. Recommendations for utilisation of the CE survey and scales produced by this thesis are presented in the discussion sections of Research Papers 2 and 3, and Supplemental Material provided in Research Paper 2. Further reflection regarding the implications of this thesis research on public health and development practice are presented in Chapter 6.

In summary:

- Assessments of collective efficacy can provide important information related to motivational factors influencing communal goal setting as well as the amount of effort and resolve that will be expended in pursuit of those goals.
- Organisations and researchers implementing community-based interventions, particularly those dependent on collective action, can adapt and employ the CE scales provided herein to:
    - Conduct rapid community appraisals or situation assessments during the programme planning phase. Doing so will allow for the capture of data for decision-making regarding the targeting of participant communities. Communities with high collective efficacy perceptions relative to other communities are more likely to set and pursue communal goals in line with public health and development programmes, including the adoption of improved behaviours;
    - Inform intervention content, implementation techniques and strategies that align with and address the specific barriers to and facilitators of cooperative behaviour and collective action in those particular communities; and
    - Perform diagnostic evaluations related to the effectiveness of community-based interventions, including articulations of the pathways of change.
- While this thesis focused on WASH interventions, the CE instruments developed herein were specifically designed such that they were *not* domain-specific, so they would be applicable for integrated programmes. In other words, these tools are applicable for

targeting, designing, and evaluating any community-based intervention predicated on collective action.

    - If domain, programme, or goal-specific investigations are desirable, the CE instruments can be tailored for those purposes.
    - It is worth noting that there is a need to evaluate face validity of CE prompts via cognitive interviews and locally adapt scale items prior to the deployment of the CE tool. Assessing face validity of any item in a survey instrument is essential for the generation of high quality data, and represents good research practice, so this step should not pose undue burden to implementers interested in performing rigorous programme targeting, design, or evaluation.
- Although more complicated factor score calculation procedures were carried out to produce weighted factors scores in Research Papers 2 and 3, more simplistic factor score averages were generated in Research Paper 2 as well. This was done in order to demonstrate the ability of the CE scales to be used in that manner, and to present non-weighted results that can be used in comparative analyses.

Future research into collective efficacy should:

- Investigate and pinpoint appropriate item indicators to assess motivational concepts that were not thoroughly explored in the India and Ethiopia work.
- Include casual modelling and path analyses to determine causal attribution, test remaining hypotheses, and identify change pathways leading to collective action.
    - Related analyses are planned for *Andilaye* trial data. The ex-ante study design will allow for causal attribution and hypothesis testing regarding the proposed bi-directional relationship between collective efficacy and intervention effectiveness.
    - Similar investigations across the WASH, public health, and development sectors are recommended in order to provide additional data that can be used to ascertain which factors transcend contexts, and for further refinement of CE scales.

7.2 Final remarks

Collective efficacy is an underutilised social construct and interpersonal factor of cooperative behaviour and collective action. Findings from this thesis indicate that CE is a complex, multi-dimensional construct. While the full 50-item CE survey and reduced Ethiopia (parsimonious

and gender-specific) and India CE scales are readily available for local adaptation and contextualisation, cognitive validation and vetting of survey items is essential to ensure high quality data. As with any rigorous evaluation, quantitative results produced through the administration of the CE scales should be accompanied by complementary qualitative examinations that serve to aid in the interpretation of quantitative results and the substantive justification of model decisions.

It is recognised that CE perceptions are generated from inferential processes, and that the factors affecting CE may differ substantially across contexts, as related sub-constructs are largely informed by context-specific political economies and social schemas. As such, adaptation of CE surveys for context-specific examinations of CE are recommended, when possible.

BMJ Open

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Heather Reese,[1] Parimita Routray,[2] Belen Torondel,[2] Gloria Sclar,[1] Maryann G Delea,[1,2] Sheela S Sinharoy,[3] Laura Zambrano,[1] Bethany Caruso,[1] Samir R Mishra,[4] Howard H Chang,[5] Thomas Clasen[1,2]

To cite: Reese H, Routray P, Torondel B, *et al*. Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India. *BMJ Open* 2017;**7**:e012719. doi:10.1136/bmjopen-2016-012719

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2016-012719).

Received 19 May 2016
Revised 25 October 2016
Accepted 9 January 2017

For numbered affiliations see end of article.

Correspondence to
Heather Reese;
heather.e.reese@emory.edu

ABSTRACT

Introduction: Government efforts to address massive shortfalls in rural water and sanitation in India have centred on construction of community water sources and toilets for selected households. However, deficiencies with water quality and quantity at the household level and community coverage and actual use of toilets have led Gram Vikas, a local non-governmental organization in Odisha, India, to develop an approach that provides household-level piped water connections contingent on full community-level toilet coverage.

Methods: This matched cohort study was designed to assess the effectiveness of a combined piped water and sanitation intervention. Households with children <5 years in 45 randomly selected intervention villages and 45 matched control villages will be followed over 17 months. The primary outcome is prevalence of diarrhoeal diseases; secondary health outcomes include soil-transmitted helminth infection, nutritional status, seroconversion to enteric pathogens, urogenital infections and environmental enteric dysfunction. In addition, intervention effects on sanitation and water coverage, access and use, environmental fecal contamination, women's empowerment, as well as collective efficacy, and intervention cost and cost-effectiveness will be assessed.

Ethics and dissemination: The study protocol has been reviewed and approved by the ethics boards of the London School of Hygiene and Tropical Medicine, UK and KIIT University, Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation to stakeholders, government officials, implementers and researchers.

Trial registration number: NCT02441699.

Strengths and limitations of this study

- The study assesses a combined household-level piped water and sanitation intervention that requires complete community-level compliance.
- The intervention was not randomly allocated; but, controls are selected through a restriction process to limit possible partial exposure to the intervention through spill over, and matched to intervention villages using preintervention data.
- The study uses a holistic definition of health to assess intervention impacts on physical, mental and social well-being, including more novel outcomes such as seroconversion to enteric pathogens, environmental enteric dysfunction and sanitation insecurity. It also assesses intervention coverage, cost-effectiveness and collective efficacy.
- The time lapse between intervention completion and the beginning of the evaluation process prevents baseline comparison or assessment of immediate intervention impacts. However, it allows for a biologically plausible length of time for die-off of even the most persistent pathogens in the environment and provides time for children to have been born into this environment.

INTRODUCTION

Of the one billion people practicing open defecation worldwide, over half of them live in India.[1] While international and national pressure on improving sanitation conditions in India has led to over 350 000 people gaining access to improved toilets since 1990, it has barely kept up with population growth.[1 2] Recent studies show that even in areas with access to household-level improved sanitation, use of these toilets is low.[3–5] This may be due in part to a mismatch between the culturally acceptable pour-flush toilets and the level of water access. Coverage of improved water sources, usually community-level pumps or taps, is

relatively high even in rural areas in India, but it may not be sufficient for flushing purposes on top of other daily water needs.[1 6]

Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary, meta-analyses have found that individual or combined WASH interventions decrease diarrhoeal disease prevalence by up to 48%.[7–11] While combined interventions would be expected to have a greater influence on multiple exposure pathways and thus a greater combined impact on health, there is limited evidence of additive benefits.[12] This may be due to poor uptake, inconsistent use or an incomplete understanding of relevant pathways.[8–10] In India, combining water and sanitation interventions may be more critical than just interrupting multiple transmission pathways for enteric infection; evidence suggests that household-level water access is integral to the use of improved sanitation in this context.[13]

While the intent of improved sanitation facilities is to separate human feces from human contact, most of the focus is on constructing household toilets to increase improved sanitation coverage—the primary metric used in monitoring progress towards international targets. However, studies in India have further shown that toilet construction does not translate into toilet use in this context.[5 14–16] Moreover, with the interdependence between members of households and households within communities, safe water and sanitation is a community-level issue. There is growing emphasis on assessing health risk from poor water and sanitation conditions not simply due to individual or even household-level risk factors, but also from conditions in the community environment.[17] There is evidence that even households without toilets, and households which do not filter drinking water, showed decreased health risk if they live in communities with high levels of coverage and use.[18–20]

Moreover, the effectiveness of community interventions may be higher in communities with positive perceptions of their collective ability to come together to improve their conditions. Collective efficacy (CE), a latent construct comprised the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal, may explain some variance in intervention effectiveness across communities receiving WASH interventions.[21]

A main risk of poor WASH conditions is enteric infection, caused by a diverse array of bacteria, viruses, protozoa and parasites, including soil-transmitted helminths. These infections may cause diarrhoea, the second leading cause of mortality for children <5 years worldwide and in India, a leading cause of mortality regardless of age.[22 23] There is also growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat enteric infections contributing to chronic malnutrition, environmental enteric dysfunction (EED), poor cognitive outcomes and poor vaccine uptake.[24–29] Poor WASH conditions are also linked to increased risk of respiratory infection, the leading cause of mortality for children <5 years worldwide.[22 30 31] Poor water and sanitation access can also affect the social, physical and mental well-being of women, acting through pathways ranging from unsafe menstrual hygiene management practices and increased risk of violence.[32–34]

Description of the intervention

Over the past few decades, there has been a global commitment to determine water and sanitation interventions with demonstrated effectiveness, not just efficacy.[35] Gram Vikas, a non-governmental organisation based in Odisha, India (http://www.gramvikas.org/), has responded by implementing its MANTRA (Movement and Action Network for Transformation of Rural Areas) water and sanitation programme in more than 1000 villages since 2002.[36] This approach includes household-level piped water connections and community-level mobilisation for culturally appropriate household toilets. A previous interrupted time series analysis of the MANTRA intervention reported it to be protective against diarrhoeal diseases.[37] However, in addition to limitations of design, this study relied on outcome data collected and reported by Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts on environmental fecal contamination.

The MANTRA water and sanitation intervention is rolled out in a three-phase process over an average of 3 years. During the first, or Motivational, phase (~8–12 months), representatives of Gram Vikas visit the identified village several times to assess village interest and progress towards a set of Gram Vikas requirements, including: (1) the commitment of every household to participate, (2) creation of a village corpus fund from contributions from every household and (3) development of village guidelines for maintenance and use of facilities.

Once this set of requirements is achieved, the village progresses into the second, or Operational, phase of the intervention (~17–35 months). Each household constructs a pour-flush toilet with two soak-pits and a separate bathing room. The households hire a local, skilled mason and provide their own unskilled labour and locally available materials to complete the superstructure. Gram Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a water tank, community meeting space and piped water distribution system connected to every household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is constructed through a similar collaborative process.

All households must construct a toilet and bathing room for the village to progress into the final, or completed, phase of the intervention, in which the water system is turned on. Notably, this three-phase process only allows each household access to piped water once every household in the village has a toilet and bathing

room. This model contrasts with most previous water and sanitation interventions, including those implemented under India's Total Sanitation Campaign and other government programmes, which do not require community-level sanitation compliance and do not provide a piped water supply at the household level.[38]

Study aims

The primary objective of this study is to evaluate the effectiveness of the combined household-level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India. Towards that objective, this study aims to:

1. Assess the effectiveness of the intervention in improving water and sanitation infrastructure coverage, access, and use, and to assess fecal sludge management practices in intervention communities.
2. Assess the effectiveness of the intervention in reducing environmental fecal contamination.
3. Assess the effectiveness of the intervention in improving health. This includes reported diarrhoeal disease in children <5 years (primary outcome), acute respiratory infection, infection with soil-transmitted helminthes, nutritional status, EED, seroconversion for selected enteric pathogens and urogenital diseases associated with menstrual hygiene management practices. Mental and social well-being will be explored through assessment of sanitation insecurity and women's empowerment.
4. Assess the cost and cost-effectiveness of the intervention.
5. Develop and assess a theoretically grounded, empirically informed CE scale and determine the effect of CE on intervention effectiveness.

METHODS

Setting

The study is located in Ganjam and Gajapati districts in eastern Odisha, India (figure 1). These two contiguous districts were a single district until 1992. Over 44% of the population in these districts is recognised by the Government of India as being below the poverty line

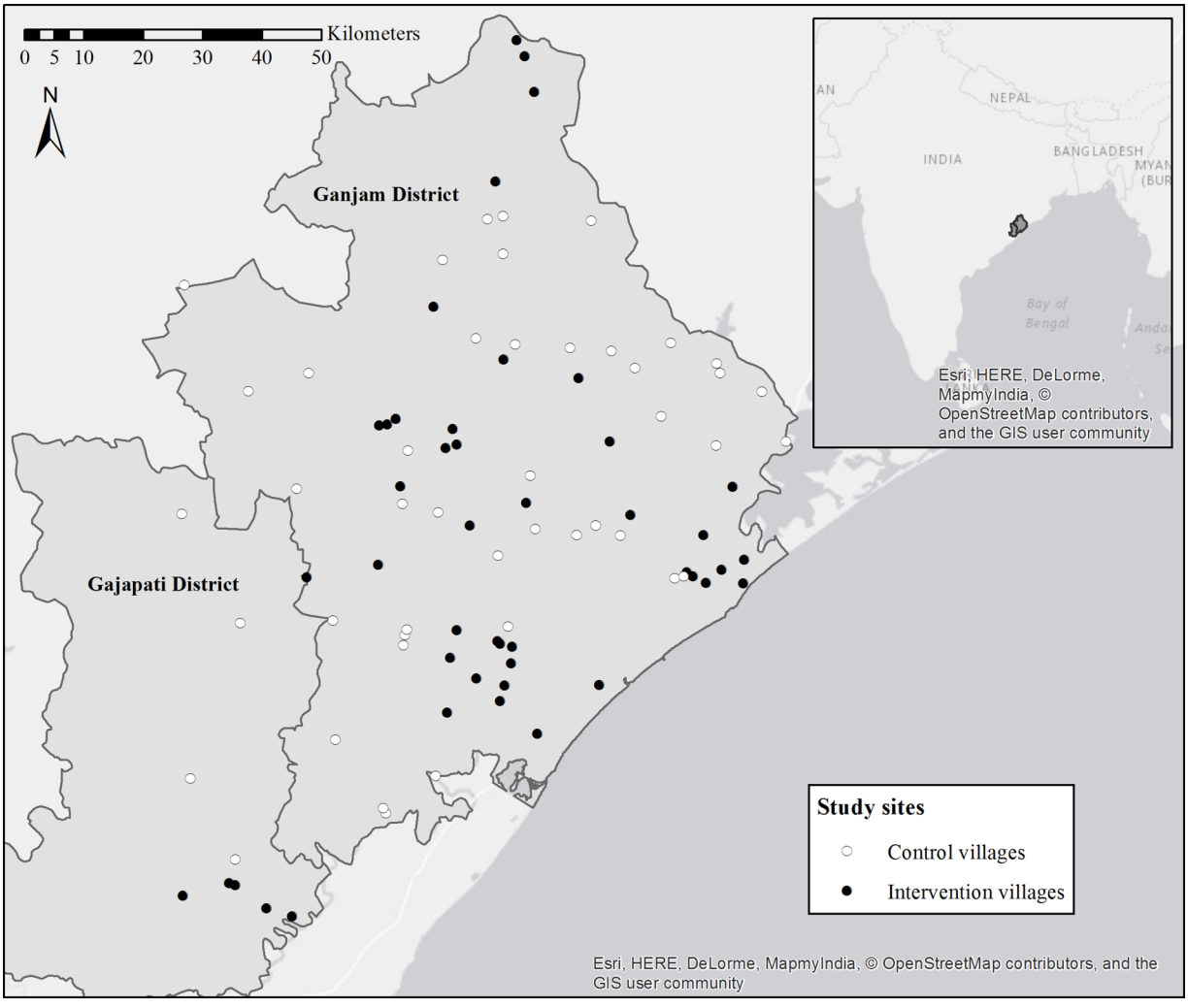


Figure 1 Study sites in Ganjam and Gajapati districts, Odisha, India, with intervention villages in black and control villages in white. Inset shows location of districts in India.

(BPL).[39] As of 2008, a majority of households in both districts had access to an improved, likely community-level, drinking water source, with over 23% of households in Ganjam having access to any sanitation facility, compared to only 8% of households in Gajapati.[39] The area is primarily rural and agrarian, and the climate is characterised by a monsoon season from June to September, with an average rainfall of ~1400 mm/year.

Study design

This study uses a matched cohort design to assess the effectiveness of a completed intervention with data collected across four study rounds from June 2015 to October 2016 (figure 2). Data were collected in all study rounds for diarrhoea, acute respiratory infection, nutritional status and stored and source water outcomes to assess seasonality. Data were collected in rounds 2 and 4 for EED, seroconversion, and hand-rinses, and cross-sectionally in one or more rounds for the remaining outcomes. As described below, control villages were matched to randomly selected intervention villages through a multistep restriction, genetic matching and exclusion process using the following eligibility criteria.

Eligibility criteria for villages

Restriction

Intervention villages were randomly selected from a list of Gram Vikas villages in Ganjam and Gajapati districts provided by the non-governmental organization (NGO), after restriction to villages with a Motivation phase start date between 2002 and 2006 and a Construction phase start date no earlier than 2003. Since the intervention process takes on average 3 years, the criteria for the Motivation start date helped to identify those villages with ongoing interventions at the same time. In addition, this allowed the use of the Government of India Census 2001 and the BPL Survey 2002 data to characterise baseline characteristics used in the matching process in intervention and control villages.

Eligible control villages include all villages without a Gram Vikas intervention within the study districts which: (1) are not within the same Gram Panchayat (a political subdivision with some administrative responsibility for water and sanitation comprised several villages) as a Gram Vikas village, or bordering a Gram Vikas village, and (2) had not received a Motivation visit from the Gram Vikas NGO. These criteria serve to limit the possibility of previous partial exposure to the intervention through spill over from adjacent villages or direct contact with the NGO. These criteria also increase the strength of the comparison provided by the control villages, that is, it increases the likelihood that if they had received a motivation visit from Gram Vikas, the control villages would have been equally as likely as the intervention villages to demand the intervention.

In addition, to be eligible for inclusion intervention and control villages must (1) appear in the Government of India Census 2001 and the BPL Survey 2002, (2) have a population of at least 20 households and (3) be within ~3 hours travel from the study office in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

Matching

After restriction, genetic matching was used to match potential control villages to the randomly selected intervention villages without replacement.[5 40 41] Villages were exact matched on district to limit any political or

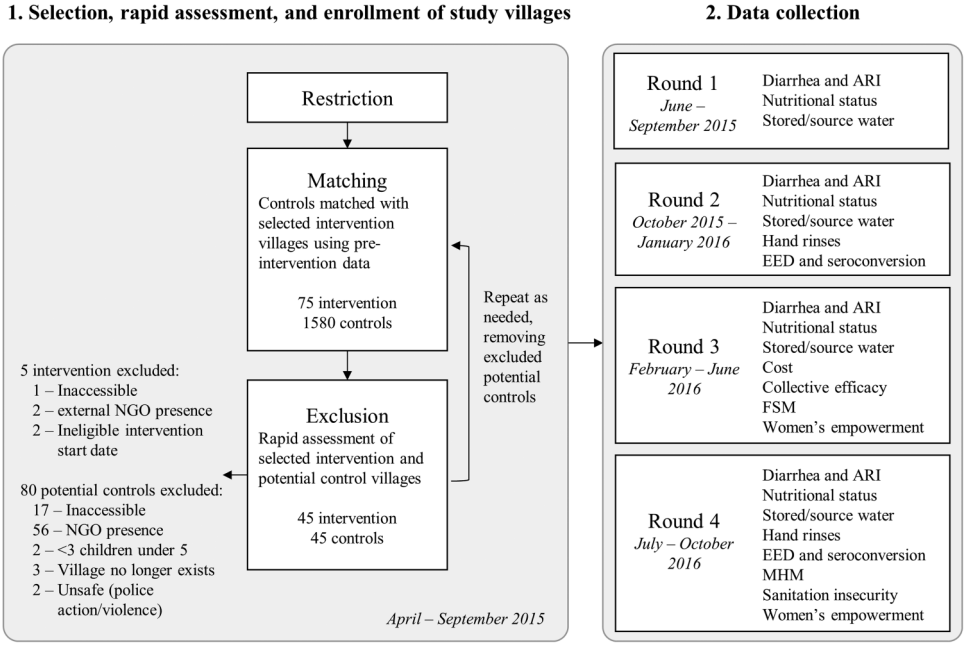


Figure 2 Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

large-scale geographic variation between district populations and were also matched on preintervention demographic, socioeconomic, sanitation and water access characteristics listed in table 1.[5] These village-level matching variables were selected due to their theorised association with the primary outcome, diarrhoeal diseases, as well as data availability.

Exclusion

The field team visited matched potential control villages and intervention villages to assess suitability for the study through a rapid assessment interview with village leadership and to ensure accessibility. Villages were excluded if they are not within 3 hours travel of the field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or if there was a current or planned sanitation or water intervention by an organisation external to the village in the next 12 months as determined through the rapid assessment interview with village leadership. In addition, villages were excluded if there were fewer than three households with children <5 years old. As villages were removed from the pool of prospective control villages, the matching process was repeated for all intervention villages and remaining eligible control villages, and balance measures were assessed. The matching and exclusion processes were repeated as necessary.

After the iterative matching and exclusion process was complete, covariate balance was assessed for all matching variables for the final set of intervention and control villages through examination of balance measures.[42–44] Matching resulted in an improvement in balance as assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov bootstrap p values and standardised differences. After matching, there were no significant differences between the intervention and control groups (table 1).

Eligibility criteria for households

Households within selected intervention and control villages were eligible if they had at least one child <5 years at the time of enrolment, verified with birth or immunisation card and expected to reside in the village for the duration of the study. If there were more than 40 eligible households within a village, only 40 were randomly selected to be enrolled. Informed consent was obtained from the male and/or female head of the selected households. All children <5 years within each enrolled household were eligible and do not age-out over the course of the study. Households with newborn children were enrolled as they became eligible on an ongoing basis throughout the study, in villages with <40 enrolled households.

Sample size

Sample size was determined through a simulation estimating the log odds of diarrhoeal disease (the primary outcome) through a multilevel random effects model and parameterised with data from a previous study in a neighbouring district in Odisha.[16] Sample size estimates were also checked with G*Power.[45] The simulation assumes a longitudinal 7-day period prevalence for diarrhoea of 8.8% in children <5 years, a heterogeneity variance between villages of 0.07, a heterogeneity variance between households of 0.57 and four study rounds.[16] An effect size of 0.20 was selected for public health importance and based on estimates of effect from systematic reviews of water and sanitation studies.[46] Assuming at least 80% power, 0.05 significance level, 10% for loss to follow-up and at least one child per household, we

Table 1 Preintervention characteristics used in matching, and balance diagnostics before and after matching and exclusion process

Variable	Intervention (n=45)	Control (all eligible) (n=1580)	Std diff (all eligible)	Control (study) (n=45)	Std diff (study)
Number of households	157.9	215.5	0.37	148.1	0.06
Population <6 years (%)	16.2	16.9	0.19	16.3	0.02
Household income score ($\bar{x}$)	2.9	3.1**	0.26	2.9	0.01
Household goods owned ($\bar{x}$)	1.1	1.2*	0.27	1.1	0.02
Pucca house (%)	59.2	61.6	0.09	60.5	0.05
≥2 meals a day (%)	57.7	63.7	0.19	57.8	0.01
Scheduled caste (%)	11.5	18.7**	0.46	11.8	0.01
Scheduled tribe (%)	33.4	19.1*	0.31	29.8	0.08
Female literacy (%)	30.9	29.8	0.07	30.9	0.00
Open defecation (%)	95.6	95.2*	0.04	95.8	0.01
Improved drinking water source† (%)	38.6	42.5	0.10	37.2	0.02
Water source <500 m and 50 m elevation (%)	81.5	72.2	0.31	81.7	0.01

All eligible: all villages that are eligible for the matching process after restriction.
Std diff (absolute standardised difference): a value >0.1 is considered meaningful imbalance.[42]
Kolmogorov-Smirnov bootstrap p values: *<0.05, **<0.01.
†Ganjam villages only; no data available for Gajapati villages.

estimate a sample size of 45 villages per study arm and 26 households per village. This estimate was the most conservative compared with sample size estimates for secondary outcomes and was therefore used for the broader study population.

Outcome measurement

Outcomes, and individual, household, and community-level risk factors, will be measured through surveys, interviews or through the collection and analysis of environmental, stool or dried blood spot samples. All survey questions will be translated into the primary local language, Odia, and back-translated to confirm wording. Household surveys include household and individual factors and will be verbally administered by trained field workers to the mother or primary caregiver of the youngest child <5 in each household, unless otherwise specified below. Community surveys will be verbally administered to the *sarpanch* (village head) or any other member of village leadership. Survey data will be collected on mobile phones using Open Data Kit.[47] GPS coordinates for households, water sources and other relevant sites will be collected using Garmin eTrex 10 or 20 devices (Garmin, Olathe, Kansas, USA).

Coverage, access and use of sanitation, water and hygiene infrastructure

Coverage, access and use of WASH infrastructure will be assessed in all four rounds. Presence of and access to toilets, water sources and hand-washing stations will be assessed through standard questions from the Demographic and Health Surveys (DHS) and confirmed through spot observations. Spot observations of household toilets and hand-washing stations will be further used to assess indicators of functionality, maintenance, recent use. Reported water and sanitation practices, including child feces disposal practices, will be captured through household survey questions.

Diarrhoeal diseases

The primary outcome for this study is prevalence of diarrhoeal diseases, recorded as both daily point prevalence over the previous 3 days and 7-day period prevalence, for all household members in all four rounds. Although self-reported diarrhoea is a subjective outcome with a well-established risk of bias, three-day recall reduces recall bias.[48 49] Diarrhoeal disease will be measured using the WHO definition of three or more loose stools in a 24-hour period, with or without the presence of blood. Field workers will use a simple calendar as a visual aid to help respondents with recall. Each household member will be asked to recall his or her own disease status, and the mother or primary caregiver will be asked to report disease for children.

Respiratory infection

Prevalence of respiratory infections will be recorded as both daily point prevalence over the previous 3 days and 7-day period prevalence for all household members in all four rounds. Respiratory infection is defined as the presence of cough and/or shortness of breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness (IMCI).[50] The full IMCI case definition for acute lower respiratory infection also includes measurement of respiratory rate and observation of chest indrawing, stridor and other danger signs; these criteria were excluded from our definition as there was concern about the technical support required to produce consistent and accurate data within this context.[50] Our definition provides a broad assessment of respiratory illness burden. Each household member will be asked to recall his or her own disease status, and the mother or primary caregiver will be asked to report disease for children.

Nutritional status

Anthropometric data will be collected for children under the age of five in all four rounds using standard methods as established by WHO.[51 52] Field workers will be trained and standardised in line with WHO protocols to reduce measurement error.[52] Weight will be measured for all children <5 years of age using Seca 385 digital scales, with 20 g increment for weight below 20 kg and a 50 g increment for weight between 20 and 50 kg. Recumbent length will be measured for children <2 years of age using Seca 417 measuring boards with 1 mm increment. Standing height will be measured for children 2–5 years of age using Seca 213 portable stadiometers with 1 mm increment. Height and weight will be used to calculate height-for-age z-scores (HAZ) and weight-for-height z-scores (WHZ) based on WHO reference standards. A random subset of 10% of households will receive back check visits each day to repeat height/length measurements to ensure interobserver reliability.

Soil-transmitted helminth infection

Stool samples will be collected in rounds 2 and 4 from all household members in a randomly selected subset of 500 households and used to assess the presence and intensity of soil-transmitted helminth (STH) infection. Formalin ether concentration and microscopy will be used to quantify worms and ova for hookworms, *Ascaris lumbricoides*, and *Tricuris trichura*.[53] Quality assurance includes independent duplicate assessment of all positive and 10% of negative samples. After stool collection, each participant will be offered a single dose of Albendazole, a broad-spectrum antihelmenthic drug recommended by the Ministry of Health and Family Welfare, Government of India. Stools collected in round 2 will allow for comparison of STH infection prevalence between intervention and control villages, while the stool samples collected ~8 months later in round 4 will provide a measure of re-infection rate.

Environmental enteric dysfunction

Stools from a randomly selected subset of 200 children <2 years old, collected in rounds 2 and 4, will be used to

assess EED through quantification of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), α-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal permeability and TH1 immune activation, respectively, were selected for this study based on evidence of association with EED, subsequent linear growth deficits and household environmental fecal contamination.[24 25 54]

Seroconversion for enteric pathogens

Serological assays that assess antibody production against various enteric pathogens can provide an objective measure of exposure to enteric infections.[55] Enrolling children aged 6–18 months will reduce the potential for interference from maternally acquired antibodies and permit analysis of seroconversion data in a critical window for young children who experience higher diarrhoeal disease morbidity and mortality before 2 years of age.[56–61] Children who are 6–12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as appropriate, and will be visited again during round 4 for a second capillary blood sample. All blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7 days of collection at −20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*, *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *Escherichia coli*, heat-labile enterotoxin (ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholerae* and *Toxoplasma* spp. will be assessed using multiplex immunoassay technology on the Luminex xMAP platform.[62]

Environmental fecal contamination

Field workers will collect samples of household stored drinking water and source water from a random subset of 500 households in all four rounds, and child hand rinses in rounds 2 and 4. All water and hand rinse samples will be stored on ice during transport and analysed within 6 hours of collection using membrane filtration. Three assays will be used: (1) plating on m-Coli Blue 24 (Millipore, Billerica, Massachusetts, USA) for *E. coli* according to EPA Method 10 029, (2) alkaline peptone water enrichment prior to plating on thiosulfate citrate bile salts sucrose agar and slide agglutination serotyping for *V. cholerae* and (3) plating on xylose lysine desoxycholate agar, and slide agglutination serotyping for *Shigella* spp.[63–65] Source and stored water samples will be assayed for *E. coli*, *V. cholerae* and *Shigella* spp., and hand rinse samples will be assayed for *E. coli* and *Shigella* spp. *E. coli* was selected as a standard non-human specific indicator of fecal contamination, though the limitations of this indicator are well-established.[66–68] In order to better characterise human fecal contamination of the household environment, *V. cholerae* and *Shigella* spp. were selected based on prevalence in southern Asia, evidence of public health importance, and field laboratory limitations.[69–71]

Cost and cost-effectiveness

Costs and potential cost savings (ie, averted costs) associated with the intervention will be assessed through an economic costing approach that recognises and quantifies costs and benefits from a societal perspective.[72] Data on programme and point-of-delivery inputs will be collected at household, community and implementer levels in round 3. Field workers will administer community surveys to a village leader, and household surveys to the household decision-maker for toilet installation, in 20 randomly selected households in 20 matched intervention and control villages. Given cost-effectiveness analyses require the effect of the intervention to be measured against a counterfactual, and the intervention of interest is a community-based intervention, cost and effectiveness measures will be summarised at the village level.[73] Surveys will collect data on household-level and community-level inputs related to materials and labour required to construct household toilets and wash rooms, the community water tank and distribution system and household water connections; longer-term water supply and toilet maintenance costs and financing required for this infrastructure as well as perceived benefits, including averted social opportunity costs. Implementer inputs from Gram Vikas will be collected through an enumeration exercise, interviews and examination of the implementer's financial records.

Collective efficacy

CE is a latent construct comprised the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal.[21] A review of the literature and established conceptual frameworks will be performed to define the CE construct. A sequential exploratory mixed qualitative and quantitative design will be used to develop and refine a scale to measure CE and test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to one randomly selected household member aged 18 years or older in each household in round 3.

Women's empowerment

Four dimensions of women's empowerment will be measured in rounds 3 and 4: group participation, leadership, decision-making and freedom of movement. Group participation and leadership will be measured using modules from the Women's Empowerment in Agriculture Index (WEAI), which has been tested in South Asia.[74] Decision-making will be measured using questions from the women's status module of Demographic and Health Surveys. Freedom of movement will be measured using questions from the project-level Women's Empowerment in Agriculture Index (pro-WEAI). These measures will be collected for the primary female caregiver of the youngest child <5 years of age and were selected based on the importance of women's empowerment for child nutrition.[75 76]

Women's empowerment is conceptualised as an outcome and a potential mediator along the pathway between the Gram Vikas intervention and child health outcomes.

Menstrual hygiene management

Menstrual hygiene management practices vary worldwide and depend on personal preference, socioeconomic status, local traditions and beliefs and access to water and sanitation resources.[77] Unhygienic washing practices are common in rural India and among women and girls in lower socioeconomic groups and may increase risk of urogenital infection.[78–80] However, the link between access to water and sanitation, menstrual hygiene management and urogenital infections has been poorly studied. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a subset of 800 households, and will capture self-reported urogenital infection, defined as at least one of the following symptoms: (1) abnormal vaginal discharge (unusual texture and colour/more abundant than normal), (2) burning or itching in the genitalia, (3) burning or itching when urinating or (4) genital sores.[79]

Sanitation insecurity

This study will assess the associations between sanitation access and sanitation insecurity with mental health among women. In a previous research in Odisha, a contextually specific definition and measure for sanitation insecurity was developed, with associations between facets of sanitation insecurity and mental health independent of sanitation facility access.[81] This previously developed measure will be used to determine if levels of sanitation insecurity differ between intervention and control villages and how it may be associated with mental health outcomes, specifically well-being, anxiety, depression and distress. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a random subset of 800 households.

Fecal sludge management

In sanitation systems where sewerage is not feasible, such as the household toilets constructed as part of the MANTRA intervention, safe management of fecal waste is necessary. Although there is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on urban settings.[82 83] Preliminary research in Odisha suggests that FSM in this rural setting is a substantial challenge and may impact household use of toilets. In round 3, household surveys and spot checks of toilets in intervention villages will be used to assess toilet use and FSM practices.

STATISTICAL ANALYSES

The effect of the intervention on infrastructure coverage, access, and use (aim 1), and the effect of the intervention on improving health (aim 3), will be analysed using logistic, linear, log binomial or negative binomial multilevel regression depending on the outcome, to compare intervention versus control villages. Prevalence of FSM practices in intervention communities will be assessed using multilevel regression (aim 1). For all models, the hierarchical structure of the data will be accounted for using random effects. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Mediation of the potential association between intervention and nutritional status outcomes by women's empowerment will be assessed using multilevel structural equation modelling, and statistical approaches to reduce bias will be explored as needed.[84]

The impact of intervention on reducing environmental fecal contamination (aim 2) will be assessed through two methods. First, hierarchical logistic and negative binomial multilevel regression to estimate intervention effects on the relative scale will be used to compare intervention versus control villages. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Second, a stochastic microbial risk framework will be used to assess differential fecal environmental contamination between intervention and control villages.

The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps. Incremental intervention benefits will be ascertained by combining health benefit data, from analysis of health outcome data and established averted cost data, with other averted social opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific DALY, will be calculated by dividing the incremental intervention costs by the incremental intervention benefits.

The CE scale will be analysed using a psychometric approach in which factor analytics are employed to identify an appropriate factor solution and test the reliability and validity of the CE scores. Once a CE factor solution and an empirically derived multilevel data structure have been identified, the association between CE and intervention effectiveness will be analysed using multilevel generalised linear mixed models to estimate relative risks,[85 86] (aim 5). For all outcomes, variables used in the matching process may be considered as covariates, as needed, in addition to individual, household and community-level risk factors. Covariates that are statistically associated with outcomes of interest in bivariate analyses will be considered for inclusion in final multivariable models, following standard stepwise model-building approaches. Secondary analyses may also evaluate models for effect modification as relevant, including exposure-mediator interaction for mediation models and cross-level interaction, by assessing changes in parameter values based on potential effect modifiers. Potential effect modifiers may include breastfeeding for

seroconversion outcomes, and climate factors and population density for environmental fecal contamination and health outcomes. However, this study was not designed to assess effect modification and therefore is not specifically powered for these analyses. For all outcomes, unadjusted models will be presented along with models adjusting for covariates.

DISCUSSION

This matched cohort study is one of the first to evaluate the effect of a rural combined household-level piped water and sanitation intervention, implemented at the community level, on a large scale. The matched design provides a rigorous means for estimating causal effects given that randomisation to the intervention group was not feasible due to the several year implementation process.[5] By focusing on an intervention where the implementation process is complete, it also limits the risk presented by randomised controlled trials, where the intervention has little uptake, an especially important study challenge given the interdependence of exposure and outcomes within communities, and a problem that has characterised previous trials of sanitation interventions in India.[15 16]

A strength of this study is the assessment of health impacts using the holistic WHO definition of health, including not just disease status, and also mental, social and physical well-being.[8 7] Outcomes along the causal chain include standard, but more subjective measures, such as reported diarrhoeal diseases and respiratory infection, as well as more objective measures such as fecal environmental contamination, soil-transmitted helminth infection and anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be differential by intervention status since the study team is not directly linked to Gram Vikas. Even though field workers may be aware of village intervention status, laboratory staff analysing water, hand rinse, stool and blood samples will be blinded. In addition, this study includes the more novel use of seroconversion for enteric pathogens, biomarkers of EED and measures of CE in an evaluation assessment. While there are limitations inherent to observational studies, the matched study design and multivariable modelling analysis plan reduce the potential for confounding. However, there is still the potential for residual unmeasured confounding.

Dissemination

Efforts will be made to communicate the central findings and implications with study communities, the implementing organisation and government officials in India. The results of this study will be submitted for publication in peer-reviewed journals and presented at conferences. The data collected in the study will be publicly available, with personal identifiable data redacted, following the publication of the primary results within 24 months of the final data collection date.

Author affiliations
[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA
[2]London School of Hygiene and Tropical Medicine, London, UK
[3]Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, Georgia, USA
[4]School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India
[5]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

Contributors TC, HR, PR, BT and HHC contributed to study design. HR, LZ and BT developed laboratory protocols. HR, BT, GS, MGD, SSS, LZ and BC developed data collection tools. All authors contributed to editing and revising the manuscript.

Funding This study is supported by a grant from the Bill & Melinda Gates Foundation to the London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University (OOP1125067).

Competing interests None declared.

Ethics approval Ethics Committee of the London School of Hygiene and Tropical Medicine, UK and Institute Ethics Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. UNICEF, WHO. *Progress and sanitation and drinking water: 2015 update and MDG assessment*. Geneva: World Health Organization, 2015.
2. Water and Sanitation Program. *A decade of the total sanitation campaign: rapid assessment of processes and outcomes*. World Bank, 2010.
3. Coffey D, Gupta A, Hathi P, *et al*. Culture and the health transition: understanding sanitation behavior in rural north India. International Growth Centre Working Paper, 2015.
4. Clasen T, Pruss-Ustun A, Mathers CD, *et al*. Estimating the impact of unsafe water, sanitation and hygiene on the global burden of disease: evolving and alternative methods. *Trop Med Int Heal* 2014;19:884–93.
5. Arnold BF, Khush RS, Ramaswamy P, *et al*. Causal inference methods to study nonrandomized, preexisting development interventions. *Proc Natl Acad Sci USA* 2010;107:22605–10.
6. Ercumen A, Arnold BF, Kumpel E, *et al*. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: a matched cohort study in Urban India. *PLOS Med* 2015;12:e1001892.
7. Fewtrell L, Kaufmann RB, Kay D, *et al*. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2005;5:42–52.
8. Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;381:S44.
9. Clasen T, Bostoen K, Schmidt W, *et al*. Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane Database Syst Rev* 2010;(6):CD007180.
10. Clasen T, Schmidt WP, Rabie T, *et al*. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;334:782.
11. Cairncross S, Hunt C, Boisson S, *et al*. Water, sanitation and hygiene for the prevention of diarrhoea. *Int J Epidemiol* 2010;39 (Suppl 1):i193–205.
12. Clasen TF, Alexander KT, Sinclair D, *et al*. Interventions to improve water quality for preventing diarrhoea. *Cochrane Database Syst Rev* 2015;10:CD004794.
13. Routray P, Schmidt WP, Boisson S, *et al*. Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC Public Health* 2015;15:880.

14. Barnard S, Routray P, Majorin F, *et al*. Impact of Indian total sanitation campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. *PLoS One* 2013;8:e71438.
15. Patil SR, Arnold BF, Salvatore AL, *et al*. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med* 2014;11:e1001709.
16. Clasen T, Boisson S, Routray P, *et al*. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *Lancet Glob Health* 2014;2:e645–53.
17. Eisenberg JNS, Trostle J, Sorensen RJD, *et al*. Towards a systems approach to enteric pathogen transmission: from individual independence to community interdependence. *Annu Rev Public Health* 2012;33:239–57.
18. Root GPM. Sanitation, community environments, and childhold diarrhea in Rural Zimbabwe. *J Heal Popul Nutr* 2001;19:73–82.
19. Bateman OM, Smith S. *A comparison of the health effects of water supply and sanitation in Urban and Rural Guatemala*. U.S. Agency for International Development (USAID), 1999.
20. Huq A, Yunus M, Sohel SS, *et al*. Simple sari cloth filtration of water is sustainable and continues to protect villagers from cholera in Matlab, Bangladesh. *MBio* 2010;1. pii: e00034-10.
21. Bandura A. *Self-efficacy: the exercise of control*. Worth Publishers, 1997.
22. UN Inter-agency Group for Child Mortality Estimation. (UN IGME). Levels and trends in child mortality: Report 2015. UNICEF, 2015.
23. World Health Organization. *Estimated total deaths by cause, sex and WHO Member State, 2008*. Geneva, Switzerland: WHO, 2011.
24. Lin A, Arnold BF, Afreen S, *et al*. Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg* 2013;89:130–7.
25. Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med* 2012;18:328–36.
26. Guerrant RL, Oriá RB, Moore SR, *et al*. Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutr Rev* 2008;66:487–505.
27. Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction: an overview. *Food Nutr Bull* 2015;36(Suppl 1): S76–87.
28. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;374:1032–5.
29. Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: an opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2015;12:1–15.
30. Budge PJ, Griffin MR, Edwards KM, *et al*. Impact of home environment interventions on the risk of influenza-associated ARI in Andean children: observations from a prospective household-based cohort study. *PLoS One* 2014;9:e91247.
31. Aiello AE, Coulborn RM, Perez V, *et al*. Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis. *Am J Public Health* 2008;98:1372–81.
32. Caruso BA, Sevilimedu V, Fung ICH, *et al*. Gender disparities in water, sanitation, and global health. *Lancet* 2015;386:650–1.
33. Sahoo KC, Hulland KR, Caruso BA, *et al*. Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med* 2015;139:80–9.
34. Hulland KRS, Chase RP, Caruso BA, *et al*. Sanitation, stress, and life stage: a systematic data collection study among women in Odisha, India. *PLoS One* 2015;10:e0141883.
35. World Health Organization/The United Nations Children's Fund (UNICEF). *Ending preventable child deaths from pneumonia and diarrhoea by 2025: The integrated Global Action Plan for Pneumonia and Diarrhoea (GAPPD)*. 2013.
36. Gram Vikas. Annual Report 2013–2014. Gram Vikas. 2014.
37. Duflo E, Greenstone M, Guiteras R, *et al*. Toilets can work: short and medium run health impacts of addressing complementarities and externalities in water and sanitation. NBER Working Paper, 2015:21521.
38. Rosa G, Clasen T. Estimating the scope of household water treatment in low- and medium-income countries. *Am J Trop Med Hyg* 2010;82:289–300.
39. International Institute for Population Sciences (IIPS). *District level household and facility survey (DLHS-3), 2007–08: India, Orissa*. Mumbai: International Institute for Population Studies (IIPS), 2010.
40. Brady H, Caughey D, Dehejia R, *et al*. *Genetic matching for estimating causal effects*. 2012.
41. Sekhon JS. Multivariate and propensity score matching software. *J Stat Softw* 2011;42. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9 (accessed 30 May 2014).
42. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083–107.
43. Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Serv Outcomes Res Methodol* 2001;2:169–88.
44. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;15:234–49.
45. Faul F, Erdfelder E, Buchner A, *et al*. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;41:1149–60.
46. Wolf J, Prüss-Ustün A, Cumming O, *et al*. Systematic review: assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. *Trop Med Int Heal* 2014;19:928–42.
47. Hartung C, Anokwa Y, Brunette W, *et al*. Open data kit: tools to build information services for developing regions. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017
48. Feikin DR, Audi A, Olack B, *et al*. Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J Epidemiol* 2010;39: 450–8.
49. Arnold BF, Galiani S, Ram PK, *et al*. Optimal Recall Period for Caregiver-reported illness in risk factor and intervention studies: a multicountry study. *Am J Epidemiol* 2013;177:361–70.
50. World Health Organization. Integrated management of childhood illness: chart booklet. Geneva: World Health Organization, 2014.
51. Cogill B. *Anthropometric indicators measurement guide*. Revised edn. Food and Nutrition Technical Assistance Project. Washington (DC): Academy for Educational Development (AED), 2003:92. http://www.developmentgateway.org/download/202582/anthro_2003.pdf
52. de Onis M, Onyango AW, Van den Broeck J, *et al*. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;25:S27–36.
53. Truant AL, Elliott SH, Kelly MT, *et al*. Comparison of formalin-ethyl ether sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques for detection of intestinal parasites. *J Clin Microbiol* 1981;13:882–4.
54. Kosek M, Haque R, Lima A, *et al*. Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am J Trop Med Hyg* 2013;88:390–6.
55. Crump JA, Mendoza CE, Priest JW, *et al*. Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *Am J Trop Med Hyg* 2007;77:136–41.
56. Steinberg EB, Mendoza CE, Glass R, *et al*. Prevalence of infection with waterborne pathogens: a seroepidemiologic study in children 6–36 months old in San Juan Sacatepequez, Guatemala. *Am J Trop Med Hyg* 2004;70:83–8.
57. Brüssow H, Sidoti J, Link H, *et al*. Age-specific prevalence of antibody to enterotoxigenic *Escherichia coli* in Ecuadorian and German children. *J Infect Dis* 1990;162:974–7.
58. Khanna B, Cutler A, Israel NR, *et al*. Use caution with serologic testing for *Helicobacter pylori* infection in children. *J Infect Dis* 1998;178:460–5.
59. Lindkvist P, Asrat D, Nilsson I. Age at acquisition of *Helicobacter pylori* infection: comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;28:181–4.
60. Ungar BL, Gilman RH, Lanata CF, *et al*. Seroepidemiology of Cryptosporidium infection in two Latin American populations. *J Infect Dis* 1988;157:551–6.
61. Vitral CL, Yoshida CF, Lemos ER, *et al*. Age-specific prevalence of antibodies to hepatitis A in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995. Relationship of prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;93:1–5.
62. Lammie PJ, Moss DM, Brook Goodhew E, *et al*. Development of a new platform for neglected tropical disease surveillance. *Int J Parasitol* 2012;42:797–800.
63. Centers for Disease Control and Prevention. *Laboratory methods for the diagnosis of Vibio cholerae*. Atlanta (GA): Centers for Disease Control and Prevention (CDC), 1999.
64. APHA, AWWA, and WEF. Method 9260 E: Shigella. Standard Methods for the Examination of Water and Wastewater. 21st Edition, 2005. http://www.standardmethods.org/.

65. APHA, AWWA, and WEF. Method 10029: Total Coliforms and E. coli Membrane Filtration Method with m-ColiBlue24 Broth. Standard Methods for the Examination of Water and Wastewater. 21st Edition, 2005. http://www.standardmethods.org/
66. Gronewold AD, Borsuk ME, Wolpert RL, *et al*. An assessment of fecal indicator bacteria-based water quality standards. *Environ Sci Technol* 2008;42:4676–82.
67. Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One* 2014;9:e107429.
68. Levy K, Nelson KL, Hubbard A, *et al*. Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *Am J Trop Med Hyg* 2012;86:499–507.
69. Livio S, Strockbine NA, Panchalingam S, *et al*. Shigella isolates from the global enteric multicenter study inform vaccine development. *Clin Infect Dis* 2014;59:933–41.
70. Kotloff KL, Nataro JP, Blackwelder WC, *et al*. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case–control study. *Lancet* 2013;382:209–22.
71. Kanungo S, Sah BK, Lopez AL, *et al*. Cholera in India: an analysis of reports, 1997–2006. *Bull World Health Organ* 2010;88:185–91.
72. Cellini SR, Kee JE. Cost-effectiveness and cost-benefit analysis. In: Wholey JS, Hatry HP, Newcomer KE, eds. *Handbook of practical program evaluation*. San Francisco: John Wiley & Sons, 2010:493–529.
73. Edejer TT, Baltussen R, Adam T, *et al*. *Making choices in health: WHO guide to cost-effectiveness analysis*. Geneva: World Health Organization, 2003.
74. Malapit HJ, Sproule K, Kovarik C, *et al*. *Measuring progress toward empowerment: women's empowerment in agriculture index: baseline report*. IFPRI, 2014:1–60.
75. Black RE, Victora CG, Walker SP, *et al*. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013;382:427–51.
76. Smith LC, Ramakrishnan U, Ndiaye A, *et al*. *The importance of women's status for child nutrition in developing countries*. 2003. http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-countries
77. Sumpter C, Torondel B. A systematic review of the health and social effects of menstrual hygiene management. *PLoS One* 2013;8:e62004.
78. Dasgupta A, Sarkar M. Menstrual hygiene: how hygienic is the adolescent girl? *Indian J Community Med* 2008;33:77–80.
79. Das P, Baker KK, Dutta A, *et al*. Menstrual hygiene practices, WASH access and the risk of urogenital infection in women from Odisha, India. *PLoS One* 2015;10:e0130777.
80. van Eijk AM, Sivakami M, Thakkar MB, *et al*. Menstrual hygiene management among adolescent girls in India: a systematic review and meta-analysis. *BMJ Open* 2016;6:e010290.
81. Caruso B. *Sanitation insecurity: definition, measurement, and associations with women's mental health in rural Orissa, India* [dissertation]. 2015. http://pid.emory.edu/ark:/25593/rfhnt
82. Peal A, Evans B, Blackett I, *et al*. Fecal sludge management (FSM): analytical tools for assessing FSM in cities. *J Water Sanit Hyg Dev* 2014;4:371.
83. Peal A, Evans B, Blackett I, *et al*. Fecal sludge management: a comparative analysis of 12 cities. *J Water Sanit Hyg Dev* 2014;4:563–75.
84. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol* 2013;42:1511–19.
85. McNutt LA, Wu C, Xue X, *et al*. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 2003;157:940–3.
86. Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done? *Occup Environ Med* 1998;55:272–7.
87. World Health Organization. WHO definition of health. http://www.who.int/about/definition/en/print.html (accessed 26 Apr 2016).

BMJ Open

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Heather Reese, Parimita Routray, Belen Torondel, Gloria Sclar, Maryann G Delea, Sheela S Sinharoy, Laura Zambrano, Bethany Caruso, Samir R Mishra, Howard H Chang and Thomas Clasen

BMJ Open 2017 7:
doi: 10.1136/bmjopen-2016-012719