A new tool to measure approaches to supervision from the perspective of community health workers: A prospective, longitudinal, validation study in seven countries.

Frédérique Vallières1*, Philip Hyland1,2, Eilish McAuliffe3, Ilias Mahmud4,5, Olivia Tulloch6, Polly Walker7, and Miriam Taegtmeyer8

1Centre for Global Health, Trinity College Dublin, 7-9 Leinster Street South, Dublin 2, Ireland
2School of Business, National College of Ireland, Mayor Street, IFSC, Dublin 1, Ireland. Philip Hyland: philip.hyland@ncirl.ie
3School of Nursing, Midwifery and Health Systems, University College Dublin, Dublin 4, Ireland. Eilish McAuliffe: eilish.mcauliffe@ucd.ie
4James P Grant School of Public Health, BRAC University, 68 Shahid Tajuddin Ahmed Sharani, Mohakhali, Dhaka 1212, Bangladesh. Ilias Mahmud: imahmudot@gmail.com
5College of Public Health and Health Informatics, Qassim University, Bukayriah, Qassim, the Kingdom of Saudi Arabia
6Options Consultancy Service, St Magnus House, 3 Lower Thames Street, London EC3R 6HD, United Kingdom. Olivia Tulloch: o.tulloch@options.co.uk
7Global Centre for Health, HIV and WASH, World Vision International, Victoria Charities Centre, 11 Belgrave Road, London, SW1V 1RB, United Kingdom. Polly Walker: Polly.Walker@wvi.org
8Department of International Public Health, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, United Kingdom. Miriam Taegtmeyer: Miriam.Taegtmeyer@lstmed.ac.uk

*Corresponding author
E-mail: fvallier@tcd.ie
Abstract

Background
The global scale-up of community health workers (CHWs) depends on supportive management and supervision of this expanding cadre. Existing tools fail to incorporate the perspective of the CHW (i.e. perceived supervision) in terms of supportive experiences with their supervisor. Aligned to the WHO’s strategy on human resources for health, we developed and validated a simple tool to measure perceived supervision across seven low and middle-income countries.

Methods
Phase 1 was carried out with 327 CHWs in Sierra Leone. Twelve questions, informed by the extant literature on health worker supervision, were reduced to six questions using confirmatory factor analysis. Phase 2 employed structural equation modelling with 741 CHWs in six countries (Bangladesh, Ethiopia, Indonesia, Kenya, Malawi, Mozambique), to assess the factorial validity, predictive validity, and internal reliability of the questions at three time-points, over 8-months.

Results
We developed a robust, 6-item measure of perceived supervision (PSS), capturing regular contact, two-way communication, and joint problem-solving elements as being critical from the perspective of CHWs. When assessed across the six countries, over time, the PSS was also found to have good validity and internal reliability. PSS scores at baseline positively and significantly predicted a range of performance-related outcomes at follow-up.

Conclusion
The PSS is the first validated tool that measures supervisory experience from the perspective of CHWs and is applicable across multiple, culturally-distinct global health contexts with a wide range of CHW typologies. Simple, quick to administer, and freely available in eleven languages, the PSS could assist practitioners in the management of community health programmes.
Keywords

Community Health Workers, Supervision, Perceived Supervision Scale, Motivational Outcomes, Scale Validation
Background

The important role of lower-cadre health workers in achieving Universal Health Coverage (UHC) is widely recognised, with community health workers (CHWs) frequently cited as a cost-effective, critical resource for the efficient delivery of primary care in low- and middle-income contexts (LMICs) [1, 2]. Unfortunately, scaling up and sustaining CHWs programme, as envisioned at Alma-Ata, has been challenging, with wide variations in the availability, coordination, support and management of community health worker programmes [3]. Accordingly, the most recent Global strategy on human resources for health: Workforce 2030 [4] published by the World Health Organization (WHO) reiterates the need to harness the potential of community-based health workers. Specifically, the strategy calls for a global effort to integrate CHWs into national health-care systems as a means to improve their working conditions, capacity, and motivation [4].

More recently, the WHO have also called for rigorous scientific research in the area of community health workers to pay more attention to cross-cutting factors, such as management and supervision, that enable community-based health worker performance [5]. Decades of research on CHW initiatives to date have suggested several cross-cutting factors that contribute to the success of CHW programmes [6]. Among these, supportive supervision consistently emerges as a key factor in determining CHW performance, motivation, and retention [7].

In contrast to more ‘traditional’ methods of supervision which are frequently characterised by performance audits, inspections, use of checklists, and controlling and authoritarian attitudes [7-10], supportive supervision favours shared performance goals, mentoring, and two-way communication [11]. Whereas traditional approaches are frequently criticised for their failure to enhance health worker motivation [12-14],
Supportive approaches to supervision have been shown to increase the impact of CHW programmes as well as the productivity, motivation and job satisfaction of CHWs [7, 15-17]. Moreover, CHWs themselves express clear preferences for supportive approaches that are responsive to the realities of the challenges they face in programme implementation [14, 18].

In addition to supportive approaches to supervision, CHW programmes often advocate for regular supervision of CHWs. Research suggests however that regular interaction with one’s supervisor is insufficient. When compared to colleagues who had recently been supervised and felt supported by their supervisor, health workers who had recently been supervised, but did not feel supported, were found to be less productive [15]. This suggests that not only are health worker’s perceptions of the supervisory relationship significant, but that perceptions of the supportive nature of this relationship is likely a more important predictor of work-related outcomes than frequency alone. This view is consistent with well-established theories within the work psychology literature, which state that subjective, cognitive appraisals of supervision are critical factors in the prediction of a range of work performance-related factors (e.g., motivation, commitment, job satisfaction) [19].

While existing tools measure the supervision of CHWs (i.e. the “CHW Assessment and Improvement Matrix” [20]) by assessing the frequency of supervision and training of supervisors, these measures crucially ignore CHW perceptions of the supervisory process and their impact on work-performance-related factors. Moreover, such tools are lengthy, time-intensive, and require substantial programmatic input and resources; all of which are at a premium within human resource for health programming in LMICs. The need exists to develop a feasible, valid, and reliable measure of
perceived supervision that both recognises the experience of supervision from the
perspective of the individual health worker and that allows the CHW voice to be heard.

The current study aimed to develop and psychometrically validate a new, simple
measure of perceived supervision (the *Perceived Supervision Scale* (PSS)) that could
be used across multiple global health contexts. To maximise the utility of the PSS in
LMICs we sought to construct an easily-translatable measure, comprised of a limited
number of items that can be quickly and easily administered and scored; an approach
that should increase the likelihood of cross-cultural validity and subsequent use.

The development and validation of the PSS included two research phases. Phase
1, conducted in Sierra Leone, was exploratory and sought to determine the most
appropriate indicators of perceived supervision from an initial pool of test items. In
other words, we sought to determine which items, when included in a questionnaire,
measured perceived supervision among CHWs. Phase 2, conducted across six LMICs
and over a period of eight months, sought to provide a comprehensive assessment of
the psychometric properties of the PSS. Specifically, this phase assessed the predictive
validity, factorial validity, cross-cultural and temporal stability of the factor structure,
and the internal reliability of the PSS over time and across multiple cultural contexts.
In other words, we sought to determine whether the questionnaire, as developed in the
Sierra Leonean context also measured perceived supervision among CHWs across six
other contexts, and whether measures of perceived supervision using the PSS at
baseline, predicted a number of related human resource for health outcomes 8-months
later. Additionally, we assessed whether the total score on the PSS could be used by
implementers in the management and monitoring of CHW programmes.

**Methods**

*Participants and Procedures*
Phase 1 was conducted in Bonthe District, Sierra Leone among a convenience 
sample of 327 CHWs, representing 98% of the CHWs active in the four chiefdoms of 
Jong, Imperi, Sogbeni, and Kpanda Kemoh. Data collection took place over three weeks 
in May 2012 as part of a longitudinal cohort study of CHWs participating in World 
Vision Ireland’s Access to Infant and Maternal Health (AIM-Health) programme.

Phase 2 recruited a convenience sample of 741 CHWs from an additional six countries 
(Bangladesh, Ethiopia, Kenya, Indonesia, Malawi and Mozambique) all of whom were 
assessed across three time periods (baseline [T0], 4 months [T1], and 8 months [T2]). 

CHWs were recruited in consultation with either national ministries of health 
(Bangladesh, Malawi, Mozambique, Kenya), regional (Ethiopia) or district-level health 
management teams (Indonesia), and based on the presence of a functioning CHW 
programme in these districts. Data collection took place between October 2014 and 
May 2015 as part of the REACHOUT research consortium 
(www.reachoutconsortium.org). Demographic information for all participants is 
reported in Table 1.

**Development of the Initial Tool**

The 12 items of the PSS were initially constructed to capture aspects of 
supervision described in the literature [21] [22]. Items are scored using a 5-point Likert 
scale anchored by “strongly disagree” (1) and “strongly agree” (5). Items were designed 
to capture key components of supervision, as identified from the literature, including 
perceptions of regular contact (My supervisor meets with me regularly) and strong two-
way communication (My supervisor meets with me regularly to discuss problems and
solutions; My supervisor takes into consideration my views and ideas; and My supervisor is a good communicator). These items were first translated in Phase 1 into Krio, Sierra Leone’s lingua franca. During Phase 2, the refined version of the PSS was further translated into seven additional languages (Bangla, Kiswahili, Kamba, Bahasa-Indonesia, Chichewa, Portuguese, and Amharic). Translated forms of the PSS are available for free download at www.perceivedsupervisionscale.com. All versions were piloted, revised, back-translated, and compared to the original English version prior to being administered by trained enumerators. In the case of illiterate CHWs, the PSS was administered with the help of an enumerator. In the case of literate CHWs, the PSS was completed directly by the CHW. In both phases, enumerators were trained to administer the PSS in the local languages and English.

In Phase 2, work-performance related factors were also assessed over time. Adapted from Mbindo et al. [23], the Motivational Outcome Scale is a 12 item, self-report measure of work-performance related constructs: community commitment (2 items, \(\alpha = .64\)), organizational commitment (2 items, \(\alpha = .44\)), job satisfaction (4 items, \(\alpha = .73\)), and work conscientiousness (4 items, \(\alpha = .73\)). Each item was assessed using a 5-point Likert Scale, anchored by “strongly disagree” (1) and “strongly agree” (5). Among the current sample, the scale possessed satisfactory internal reliability.

**Analysis**

During Phase 1, the initial pool of 12 PSS items were assessed using confirmatory factor analysis (CFA) to develop a short, unidimensional measure of perceived supervision (see Supplementary Table 2). CFA is a statistical technique that tests whether items in a questionnaire effectively measure a theoretical construct, or latent construct, that is itself not directly observable (i.e. perceived supervision) [24].
As Phase I was more exploratory in nature, we did not expect all 12 items to measure perceived supervision in a consistent and robust manner. To determine which of these 12 items should be retained as the best measures of perceived supervision, we set an *a priori* criterion for item retention whereby only items with factor loadings $\lambda_{pre} > .55$ (equalling 30% of variance explained by the latent variable) were retained [25]. In addition to consulting factor loadings, we also consulted modification indices produced in Mplus (Version 7.4). Modification indices provided suggestions of additional items that could be removed to improve model fit (i.e. items with covarying residuals) [26].

Phase 2 also used CFA procedures to determine the factorial validity of the PSS. In addition, structural equation modelling (SEM) methods were used to assess whether perceived supervision scores, as measured by the PSS at baseline (Time 0), predicted the four criterion variables of the Motivational Outcomes Scale at endline (Time 2), controlling for sex and educational status. Here, SEM was chosen to assess the predictive validity of the PSS as it allows for all effects in the model to be estimated simultaneously. In other words, SEM methods were used to test whether the administration of the PSS scale at earlier stages of CHW programmes predicted a range of meaningful human resource for health-related outcomes throughout later stages of a CHW programme, whereby job satisfaction, organizational commitment, community commitment, and work conscientiousness were measured as known determinants of CHW programme success. The internal reliability of the PSS was assessed using composite reliability analysis [27], and descriptive statistics were calculated for each country and at each assessment period.

---

1 Depicted as $\lambda_{pre}$ in Supplementary Table 2, factor loadings indicate what proportion of the variance in each item on the questionnaire can be explained by the underlying latent construct.
Analyses were conducted in Mplus 7.4 [28] using the mean and variance-adjusted weighted least squares (WLSMV) estimator. The WLSMV estimator provides accurate parameter estimates, standard errors, and test-statistics when ordinal indicators are used [29]. Missing data was managed using the default pairwise present analysis method. Standard recommendations for assessing the fit of the CFA and SEM models were followed [30] whereby a non-significant chi-square ($\chi^2$) result indicates good model fit; Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) values $> .90$ indicate good fit; Root-Mean-Square Error of Approximation (RMSEA) with 90% confidence interval (RMSEA 90% CI) values $< .08$ reflect good fit; and values $< 1.0$ for the Weighted Root Mean Square Residual (WRMR) indicate good model fit. In other words, models that met these criteria were seen to be a ‘good’ representation of perceived supervision.

Results

Phase 1: Development of the Perceived Supervision Scale

The fit of the unidimensional, 12-item model to the sample data was poor ($\chi^2 = 355.417$, df = 54, $p < .001$; CFI = .757; TLI = .703; RMSEA [90% CI] = .131 [.119-.145]; WRMR = 1.739). Inspection of the model parameters indicated that six items failed to reach the $a$ priori criterion of factor loadings $> .55$ on the Perceived Supervision factor (Supplementary Table 2). The unidimensional model was subsequently re-estimated based on the remaining six items and model fit was acceptable ($\chi^2 = 43.952$, df = 9, $p < .001$; CFI = .961; TLI = .934; RMSEA [90% CI] = .110 [.079-.143]; WRMR = .910). The factor loadings for the six items were all positive, statistically significant, and of a robust magnitude.

Phase 2: Validity of the Perceived Supervision Scale
Table 2 reports the CFA results for the six-item, unidimensional model of the PSS across six nations, and at three assessment periods. In most cases the $\chi^2$ values were statistically significant and the RMSEA values were above the suggested cut-off point of .08. However, rejection of the models based on these indices is not warranted given the tendency for the $\chi^2$ to generate Type 1 errors, and the RMSEA to generate Type 2 errors in models with few degrees of freedom [31]. Contrastingly, the CFI, TLI, and WRMR results provided consistent support for the factorial validity of the PSS. In all 17 assessments, the CFI, TLI, and WRMR results satisfied the criteria for excellent model fit. Overall, the CFA results provide support for the validity of a unidimensional structure of the PSS that is stable over time, and cross-culturally consistent.

**INSERT TABLE 2 HERE**

The standardised factor loadings for the PSS across each nation, at each assessment, are reported in Supplementary Table 3. Factor loadings at T0 were all positive, significant ($p < .001$), and robust, with mean factor loadings ranging from .68 (Indonesia) to .92 (Kenya). Similarly, at T1 all factor loadings were positive, significant ($p < .001$), and robust, with mean factor loadings ranging from .74 (Indonesia) to .83 (Ethiopia). At T2, there was greater variability in the performance of the model parameters. Within the Indonesian sample it was necessary to add a residual covariance between two items with the lowest factor loadings (PSS4 and PSS6: factor loadings < .50) to achieve acceptable model fit. Additionally, within the Ethiopian sample two items possessed weak factor loadings (PSS2 = .11 and PSS4 = .22). Nonetheless, mean factor loadings were generally robust, ranging from .50 (Ethiopia) to .91 (Bangladesh).

Given the stability of the unidimensional structure of the PSS across nations, and time, all PSS data at T0 was merged. Model fit of this consolidated data was satisfactory ($N = 710$; $\chi^2 = 138.936$, df = 9, $p < .001$; CFI = .987; TLI = .979; RMSEA
Predictive Validity of the Perceived Supervision Scale

A PSS latent variable modelled at T0 was used to predict the summed scores of four criterion variables (job satisfaction, organizational commitment, community commitment, and work conscientiousness) measured eight months later (T2), controlling for sex and educational status. The fit of the model to the data was excellent ($\chi^2 = 91.276$, df = 41, $p < .001$; CFI = .991; TLI = .986; RMSEA [90% CI] = .045 [.033-.058]; WRMR = .847). As detailed in Table 3, the model explained between 5.8% and 16.4% of variance in each of the criterion variables, and perceived supervision positively predicted all variables ($\beta$ values ranged from .16 to .30).

Internal reliability and descriptive statistics for the PSS

Composite reliability analyses indicated that the PSS possesses satisfactory internal reliability (Supplementary Table 3), indicating that the six items were internally consistent and serve as accurate measures of perceived supervision. In every national context, and at each assessment period, the reliabilities ranged from .68 to .97. Descriptive statistics for the PSS across all nations, at each assessment period, are presented in Table 4.

---

2 A unidimensional model indicates that the PSS should be scored by summing questions PSS1-PSS6 to produce a total PSS score.
Discussion

The Perceived Supervision Scale is the first validated tool developed for collecting CHW perceptions of their supervision. The tool is brief, robust and can be applied across multiple, culturally-distinct global health contexts with a wide range of CHW typologies. Despite its recognised importance of supervision in CHW programming, supervision is often one of the weakest and most difficult elements of CHW programming to implement consistently [9, 32]. The factor structure of the PSS allows researchers and implementers to calculate a sum score of perceived supervision within CHW programming. Specifically, the total PSS score allows for a greater understanding the nature of a positive supervisory relationship. Furthermore, it grants the ability to managers to detect problematic supervisory interactions, prompt the introduction of stronger training programmes, and where necessary, the reorganisation of supervisory arrangements, contributing to the sustainability of CHW programmes.

The ability for CHW programme managers to monitor the interpersonal supervisory relationships of CHWs could help prevent deleterious work performance outcomes associated with high staff turnover and loss of worker motivation [7, 33]. The development of the PSS therefore represents a valuable contribution to global efforts to address human resource for health shortages and towards achieving UHC. Furthermore, the development of the PSS contributes towards addressing more recent calls for rigorous approaches towards scale development for human resource for health programming [34].

Phase 1 served to derive the most appropriate indicators of perceived supervision. From an initial pool of 12 item statements, developed from the extant literature on CHW supervision, six items were retained. Consistent with previous
literature, the items retained as part of the final PSS, reflect the importance of both supportive and regular aspects of supervision. Interestingly, those items associated with more traditional forms of supervision (i.e. controlling or negative interactions), were least reflective of the nature of perceived supervision among this sample of CHWs. This suggests that CHWs in Sierra Leone perceived the supervision process as a generally positive, supportive, and regular experience. The items retained as part of the supportive supervision factor offer additional insight into what content or skills should be emphasised or included as part of supervision training programmes. More specifically, the items retained in the PSS are consistent with evidence that a supportive supervisor should: meet regularly with CHWs, offer opportunities for knowledge sharing and refresher training [33], recognise and appreciate the work and efforts of a CHW, take into account the views and ideas of CHWs, and communicate effectively with the CHW [11].

As it was possible that the observed findings from Phase 1 reflected the idiosyncratic responses of the Sierra Leonean CHWs, it was imperative to assess the replicability of these findings in alternate contexts. Phase 2 confirmed the PSS’s unidimensional structure across multiple samples of CHWs from different contexts, cadres, cultures, and demographics. Additionally, the factorial validity of the PSS was evidenced across time, with the scale exhibiting stable psychometric properties (reliability and validity) over a period of eight months. Furthermore, the PSS positively predicted a range of work-performance related indicators eight months later including job satisfaction, work conscientiousness, community commitment, and organizational commitment, while controlling for sex and education. These results indicate that CHWs who perceive greater levels of supervision (i.e. supportive) report greater job satisfaction, work conscientiousness and higher levels of both community and
organizational commitment over time. Administering the PSS during early stages of programme implementation, or when used regularly as a monitoring tool, may therefore help managers to adapt supervision approaches before they negatively impact on other organizational factors in the long-term. Although such findings are important, future research should extend upon these findings and assess the effectiveness of the PSS to also predict objective outcomes of CHW performance and community health outcomes.

The current study has several limitations that should be recognised. The selection of the six PSS items was drawn from a sample of CHWs in Sierra Leone, and although the latent structure of these items was confirmed cross-culturally, it is possible that had the scale refinement process been conducted in a different setting, a different set of indicators may have been retained. It is important to note that the PSS is not presented as a comprehensive measure of perceived supervision, but rather a brief measure of the construct that possesses high utility across global health contexts. Second, the country-specific CFA models generated during Phase 2 of the study were carried out using relatively small sample sizes. Although not ideal for latent variable modelling, the small number of indicators in the PSS render this a minor limitation [35]. Third, it is worth noting that a residual covariance was added between two items in one (Indonesia, time 2) of seventeen assessments of model fit. Finally, while the PSS has been validated among CHWs across a range of LMIC contexts, it is necessary to determine the reliability and validity of PSS among more highly skilled cadres of health workers globally.

Conclusion

In comparison to current tools [20] that focus on capturing the frequency and regularity of supervision, the PSS allows for the subjective measurement for supervision as a predictor of future CHW satisfaction, engagement, and commitment.
Simple and quick to administer, and currently available in nine languages, the validated PSS has the potential to contribute towards a more accurate understanding of CHW’s perspectives of supervision, as a critical determinant of successful CHW programmes across a wide range of contexts.
Declarations

Ethics Approval and Consent to Participate

Ethical approval was obtained from the Ethical Review Committee, Trinity College Dublin and Connaught Hospital Freetown (Phase 1), in addition to the Liverpool School of Tropical Medicine Research Ethics Committee and six local ethics committees (Phase 2) (Supplementary Table 1).

Consent for Publication

Not Applicable

Availability of Data and Material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests

Funding

Phase 1 was funded by Irish Aid through World Vision Ireland’s Access to Infant and Maternal (AIM-Health) programme, as well as the Department for International Development and the United Kingdom through their Programmes Partnership Agreement. Phase 2 of this research was nested within REACHOUT, funded by the European Union FP7 grant (number 306090).

Author Contributions

FV conducted the literature review, designed the study and, together with EM, designed the items and coordinated Phase 1 data collection and analysis. PH and FV led the overall data analysis. FV, PH, PW and MT contributed equally to the manuscript writing. MT, OT, and IM coordinated the data collection and extracted the data for
Phase 2. All authors contributed to the critical interpretation of the results and approved the final version. FV is the final guarantor of the manuscript.

Acknowledgements

The authors would like to thank the CHWs in seven countries who gave their time to answer questions about their perceptions of supervision. Phase I of this research was made possible thanks to the support of Irish Aid and the people of Ireland through World Vision Ireland’s Access to Infant and Maternal (AIM-Health) programme as well as the Department for International Development and the United Kingdom through their Programmes Partnership Agreement. A special thanks to Allieu Bangura and Joseph Musa of World Vision Sierra Leone for their assistance with organising the collection of the data.

Phase 2 of this research was nested within REACHOUT, a multi-country research consortium aiming to maximize the equity, effectiveness and efficiency of close-to-community services in rural areas and urban slums. Robinson Karuga, Kingsley Chikaphupha, Daniel Gemechu, Ralicia Limato, Irin Akhter and Sozinho Ndima from the country REACHOUT teams organised the collection, cleaning and analysis of country data. The REACHOUT Consortium is funded by the European Union FP7 grant (number 306090). This document reflects only the authors’ views and the European Union is not liable for any use that may be made of the information contained therein.
References


29. Flora DB, Curran PJ: An empirical evaluation of alternative methods of
estimation for confirmatory factor analysis with ordinal data. Psychol Methods

30. Hu L, Bentler PM: Cutoff criteria for fit indices in covariance structure analysis:
Conventional criteria versus new alternatives. Structural Equation Modeling 1999,
6:1-55.

with small degrees of freedom. Sociological Methods & Research 2015, 44:486-
507.

32. Lehmann U, Sanders D: Community Health Workers: What do we know about
them? The state of the evidence on programmes, activities, costs and impact on

33. Strachan DL, Kallander K, Ten Asbroek AH, Kirkwood B, Meek SR, Benton L,
Conteh L, Tibenderana J, Hill Z: Interventions to Improve Motivation and
Retention of Community Health Workers Delivering Integrated Community
Case Management (iCCM): Stakeholder Perceptions and Priorities. Am J Trop

34. Borghi J, Lohmann J, Dale E, Meheus F, Goudge J, Oboirien K, Kuwawenaruwa A:
How to do (or not to do)... Measuring health worker motivation in surveys in

35. Wolf EJ, Harrington KM, Clark SL, Miller MW: Sample Size Requirements for
Structural Equation Models: An Evaluation of Power, Bias, and Solution