

Title: How to do (or not to do)... Measuring health worker motivation in surveys in Low and Middle Income Countries

Running Title: How to measure health worker motivation

Key words: Motivation, health worker, measurement scale, analysis

Word count: 4817

Key Messages

-A clear conceptualisation of motivation is required prior to measurement

-When measuring motivation in a new context, formative research and pre-testing is recommended to identify relevant dimensions and formulate items in local language.

-Validation of motivation measures through factor analysis is important. Where motivation dimensions are well known there is potential for greater use of confirmatory factor analysis

Abstract

A health system's ability to deliver quality health care depends on the availability of motivated health workers, which are insufficient in many low income settings. Increasing policy and researcher attention is directed towards understanding what drives health worker motivation and how different policy interventions affect motivation, as motivation is key to performance and quality of care outcomes. As a result, there is increasing interest among researchers in measuring motivation within health worker surveys. However, there is currently limited guidance on how to conceptualise and approach measurement, and how to validate or analyse motivation data collected from health worker surveys, resulting in inconsistent and sometimes poor quality measures. This paper begins by discussing how motivation can be conceptualized, then sets out the steps in developing questions to measure motivation within health worker surveys and in ensuring data quality through validity and reliability tests. The paper also discusses analysis of the resulting motivation measure/s. This paper aims to promote high quality research that will generate policy relevant and useful evidence.

Introduction

A health system's ability to deliver quality health care depends on the availability of motivated health workers, which are lacking in many low income settings (Global Health Workforce Alliance 2014). Motivation has been defined as the level of effort and desire to perform well and an important determinant of quality of care (World Health Organization. 2006). Motivation has been associated with lower levels of staff turnover (Bonenberger et al. 2014), higher retention, less job burnout and increased performance (Deci et al. 2017), including higher quality of care (Alhassan et al. 2013).

An increasing number of interventions (Willis-Shattuck et al. 2008) (Chopra et al. 2008) are designed to improve the retention of health workers and promote better service delivery by enhancing their motivation (Alhassan et al. 2103). Such interventions include financial incentives, which can be tied to performance targets (P4P) (Engineer et al. 2016), or non-financial incentives such as career development opportunities and training (Agyepong et al. 2004), upgrading facility infrastructure, resource availability (Manongi et al. 2006), strategies to improve clinical governance through supportive supervision (Bailey et al. 2016), audit or quality management processes (Bakker et al. 2011). Programme evaluators may want to assess the effect of national reforms or local programmes on health worker motivation (Huillery & Seban 2014) (Källander et al. 2015) (Chin-Quee et al. 2016). Health worker motivation studies may also be of interest in their own right to shed light on what drives motivation (Bhatnagar et al. 2017) (Mbindyo et al. 2009a), and help identify which strategies would be most effective in increasing motivation.

The desire to quantify changes in motivation and to understand motivation drivers in part explain the substantial growth in published research reporting results from health worker surveys that measure motivation (for example, (Bonenberger et al. 2014; Hotchkiss et al. 2015; Mbindyo et al. 2009b; Weldegebriel et al. 2016)). However, the measurement and analysis of motivation is not straight forward as motivation is not directly observable (Denhardt et al. 2008; Pinder 2008). A vast body of empirical literature has examined work motivation and the factors driving motivation (Pinder 2008), which have been shown to have predictive value in relation to determining health worker effort and performance (Bandura 1982). For public health researchers with no specialist background in psychology or behavioural economics, however, this literature can be daunting. The lack of guidance on the conceptualisation and measurement of motivation in health workers in particular, has resulted in inconsistent and sometimes poor quality measures within the empirical literature. Our paper aims to serve as an entry point and step-by-step guide for public health researchers new to the field and seeking to measure motivation with measurement scales within surveys. This guide can equally be applied to the measurement of related constructs (e.g. satisfaction, attitudes, perceptions) and with populations other than health workers.

This paper begins by discussing how motivation can be conceptualised, then sets out the steps in developing questions to measure motivation within health worker surveys and in ensuring data quality. The paper also discusses analysis of the resulting motivation measure/s. This paper aims to promote high quality and policy relevant research evidence.

Step 1: Conceptualising motivation

Motivation is a complex construct as indicated in this definition: *“Work motivation is a set of energetic forces that originate both within as well as beyond an individual's being, to initiate work-related behaviour, and to determine its form, direction, intensity, and duration.”* (Pinder 2008)

A list of the most prominent motivation theories is provided in Box 1. Motivation is usually either conceptualized as a unidimensional construct, where the focus is on the overall quantity of motivation available to drive behaviour (Gow et al. 2013) (Hagopian et al. 2009); or it is conceptualized as a multidimensional construct, with an additional focus for example on the composition of qualitatively different types of motivation such as intrinsic and extrinsic motivation (Lohmann et al. 2016). For definitions of key terms such as 'construct' please refer to Box 2. In some cases researchers may wish to capture multiple conceptualisations. The choice of approach depends on the research question and, in case of programme evaluation, one's theory about how a given programme will affect motivation.

The measurement of motivation is more difficult. A key question is whether to measure motivation itself (a 'direct' measure) by, for example, seeing how a programme affects intrinsic motivation; or to instead measure the things that affect or are affected by motivation ('proxies', 'indirect' measures). Direct measures of motivation are typically derived through measurement scales within a survey or through qualitative methods (Inceoglu et al. 2012) (Deci & Ryan 1985). For example, JL and ED examined whether financial incentives crowd out intrinsic motivation using measurement scales grounded in Self-Determination Theory in health worker surveys in Afghanistan and Burkina Faso (Lohmann et al. 2017) (Dale 2014; Deci & Ryan 1985). (see Figure 1). Indirect measures can be equally derived through surveys or qualitative methods or through experimental games or observations of behaviour. For instance, the Franco framework, which has been widely used in health worker motivation studies in low and middle income countries (LMICs), measures determinants of motivation with a series of psychometric scales in a health worker survey, examining the individual (e.g. self-efficacy, desire for achievement), organizational (e.g. management support, financial rewards), and external level determinants (relations with the community/patients) (see (Franco et al. 2002) (Mbindyo et al. 2009b) (Mutale et al. 2013) (Morrison et al. 2015) (Chandler et al. 2009)). JB and AK used this approach to examine the effects of primary care reforms on motivation composition and levels in Tanzania. Also in Tanzania, Leonard and Masatu (2010) made use of the Hawthorne effect (i.e. performance impact of being observed) to investigate health workers' intrinsic motivation. In choosing a motivation measure, it is important to consider whether and how a programme is likely to affect motivation and how this would affect worker performance. This paper outlines the steps in measuring health worker motivation with direct motivation measures, specifically Likert-type psychometric scales, as part of surveys, and in analysing such data, using examples from our respective research and the wider literature.

Step 2: Developing and pre-testing a tool

Having selected a conceptualisation of motivation, the first step in developing a survey tool is to identify a set of questions to measure motivation, referred to as a measurement scale (DeVellis 2012; Fowler 2009). If the aim is to understand the composition of motivation, then it is helpful to anticipate potential motivation dimensions with reference to theory and the intervention in question (Prytherch et al. 2012). Focus group discussions or in-depth interviews with health workers can also help identify dimensions and ensure appropriate communication of these concepts in the local language (Box 3) (e.g. (Agyepong et al. 2004) (Sacks et al. 2015)). Once relevant dimensions have been identified, these need to be formulated as questionnaire items (i.e. statements or questions). A good first step is to review existing scales (see (Mbaruku et al. 2014) (Hotchkiss et al. 2015) (Bonemberger et al. 2014) (Inceoglu et al. 2012) and Annex 1), and to decide on positive or negative wording, response scales and the number of response options (Box 4). It is recommended to include a minimum of three items per dimension (Little et al. 1999) (Guilford 1952), although with a new scale, 4-5 items are recommended as some items may not perform well. To enable

subsequent validity checks (Step 5), it is also important to collect data on variables that are expected to be related to motivation or motivation dimensions, such as motivational outcomes, e.g. intention to quit or organizational commitment (Hagopian et al. 2009) (Bonenberger et al. 2014), or measures of the knowledge practice gap (Leonard & Masatu 2010) or other performance measures. It is also important to collect data on variables which might influence provider responses to items (health worker or facility level characteristics) (Chandler et al. 2009) (Mbaruku et al. 2014) (Hotchkiss et al. 2015) (Franco et al. 2004). Researchers also need to decide on the mode of survey administration (see Box 5). As with all surveys, it is recommended to pre-test the motivation measurement tool with a small sample of health workers (see (Prytherch et al. 2012)), and to proceed through steps 4-5 below.

Step 3: Sample size considerations and sampling

Sample size is a further consideration prior to survey administration. The techniques used to assess the validity of the motivation measure (step 5) require certain minimum sample sizes, dependent on the number of dimensions, items and other factors (Kline 2010). Commonly used rules of thumb for factor analytical techniques are “no less than 100 observations” (Gorsuch 1983) (Kline 1979; Kline 2010) (MacCallum et al. 1999), with 50 observations often considered the absolute minimum (de Winter et al. 2009) for exploratory factor analysis; and 200 observations the minimum for confirmatory factor analysis (DeCoster 1998). If sub-group analysis is planned (for example comparing motivation between different cadres of health worker), these sample sizes should be achieved for each sub group. As in any other study, sample size requirements also depend on the planned substantive analyses (Step 7) (Borghi et al. 2009). The sample size requirements to analyse motivation determinants depends on the type of model used, with a standard linear regression model having lower sample size requirements than structural equation models (Rodriguez Pose et al. 2014). When considering the impact of a programme on motivation, power considerations are also important, but estimations of effect size tend to be difficult given that they are highly dependent on the motivation measure itself. Often, motivation surveys are administered to health workers who are present on the day of the facility visit. These health workers are likely more motivated than their counterparts who are not present at facilities, and it would be important, where possible to also make provisions to interview health workers who are absent from facilities on the day of the visit.

Step 4: Exploratory data analysis

Once the data have been collected, it is important to start with an exploration of the data, estimating mean and median scores and distributions for each item, and checking for missing data. The empirical literature on health worker motivation has tended to analyse Likert responses as continuous variable, given that the underlying motivation construct is assumed to be continuous. However, there has been some debate as to whether this approach is appropriate given the ordered categorical nature of Likert scales, though there is evidence it may make little difference in practice – see (Carifio & Perla 2008) (Carifio & Perla 2007), for more discussion of this point.

A high level of missing values may indicate that an item was not well understood by respondents (Raykov & Marcoulides 2011) (Little 1992). Where missing values exceed 10% researchers should weigh the option of dropping the item against maintaining measurement consistency across respondents. It is generally recommended to consider dropping items where more than 80% of respondents provide the same answer to a question as such items have little discriminatory value (Streiner et al. 2008). Direction of response for each item, particularly for those that were negatively worded, should also be checked for plausibility.

Step 5: Assessing validity of motivation measures

Before using motivation measurements in core analyses, researchers should ensure the measures are valid or that they measured what was intended (DeVellis 2012) (Fowler 2009).

If motivation is considered to be multidimensional, the first step in validating the measure is to determine the composition of motivation, or the underlying dimensions, to confirm or modify initial hypotheses. This is typically done with factor analytical techniques, either exploratory or confirmatory (see Box 6). Before doing so, it is important to check the factorability (or reducibility) of the data (e.g. inspect item correlation matrix; Bartlett test of sphericity test; Kaiser-Meyer-Olkin test (Yong & Pearce 2013)).

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis (CFA) is used where the researcher has strong assumptions regarding the dimensionality of the scale from prior qualitative research, theory, or prior use of the scale (Brown 2006). Researchers must specify the number of dimensions (or 'factors') and which items measure which dimension or factor. For example, the researcher may have pre-identified three motivation dimensions: 'work environment, salary, and conscientiousness'; and clearly assigned to each a number of items (for example, the item 'availability of drugs' may be associated with the dimension 'work environment').

CFA results indicate the extent to which the pre-specified dimensions are reflected in the data. Good model fit confirms that the dimensions are relevant and can be readily interpreted. Several statistical approaches can be used to confirm whether the dimensions are relevant using CFA, with structural equation modelling being the most common (Kline 2010). If health workers were sampled from facilities, it is important to account for the clustered nature of the data and the analyses described subsequently (Annex 2). In the absence of good model fit, modifications may be made by, for example, removing or reassigning items, or modifying the choice of dimensions. Careful consideration of the implications of eventual modifications for the underlying conceptualisation of motivation is recommended.

In many studies of health workers in LMICs there has been limited if any study of motivation meaning that it is unclear what the underlying dimensions or factors might be. For this reason exploratory factor analysis (EFA) has been most widely used in these settings.

Exploratory Factor Analysis (EFA)

When constructing new scales and/or applying them to novel contexts, researchers are often not entirely sure how many and which motivation dimensions the scale items measure. Unlike CFA, EFA does not impose any *a priori* assumptions on the number of motivation factors, and the assignment of items to factors. Rather, EFA is used to identify meaningful dimensions of motivation, and to determine which items measure which dimension, on the basis of respondents' answer patterns to the scale items. EFA is sometimes used to generate a theory about the relevant dimensions of motivation that are then used in a CFA. With sufficient sample size, EFA can be performed on one part of the data, and the generalizability of the extracted factors can be determined using CFA on the other (Raykov & Marcoulides 2011).

i. Factor extraction

A variety of statistical approaches can be used to extract factors using EFA. Principal component analysis (PCA), and principal axis factoring (PAF) are the most common (Williams et al. 2012) (DeVellis 2012). Rotation is used to simplify and clarify the results of EFA facilitating the

identification of factors. There are two main types of rotation: orthogonal and oblique, with the main difference being that the latter allows for some correlation between factors whereas the former does not. The former has been widely used because it is believed to be simpler (e.g. (Chandler et al. 2009)). However, as motivation dimensions are unlikely to be unrelated (for example, there will be some association between different constructs, such as drug availability and supervision or management involvement in facilities), the latter approach is preferable.

ii. Deciding how many factors and which items to retain

The full list of factors resulting from an unrestricted EFA will correspond to the number of items included. The researcher must decide how many to retain. This decision will be based in part on theoretical considerations: *how many dimensions is it reasonable to expect?*; and whether the resulting factors can be readily named and described. The following can also help determine the number of factors: a common rule of thumb is to retain factors that have eigenvalues over 1 (the Kaiser criterion) (Hayton et al. 2004) (Kaiser 1960) ; visually examine eigenvalue plots for the natural bend or break point in the data where the curve flattens out (Figure 2) (Cattell 1966) (Chandler et al. 2009); examine the total variance explained (aim to explain 50-75% with the least number of factors). In practice if using the 'factor' command in Stata, there are different cut off values for factor retention that are built into the software depending on the method of factor extraction selected¹. We encourage researchers to think critically about how many factors make sense in their context rather than to blindly accept these arbitrary cut-offs.

It is also important to examine the factor loadings for each item. In EFA, all items will load on all factors to some degree. The aim is to determine which items are most indicative of which factors, based on the degree of factor loading, with 0.3 (Tabachnick & Fidell 2007) and 0.4 being commonly used (Chandler et al. 2009) as cut-off values for 'substantive loadings'. Higher thresholds are recommended for small sample sizes. The ideal scenario is that each item has a substantive loading on only one factor and is conceptually close to the other items with substantive loadings on that factor. However, this is often not the case, and researchers will have to decide whether for instance to define a different number of factors or to eliminate items with low factor loadings. EFA is invariably an iterative process, as results change with the number of factors retained and items included.

iii. Interpreting and naming factors

When interpreting and naming factors, it is important to refer back to the exact wording of the scale items and the aspects of motivation they were designed to measure. Often, the interpretation of a factor is relatively straightforward from the items loading on it. For example, in Tanzania, the following three items: availability of drugs, supplies and equipment at the facility, had substantive loadings on the same factor. All three clearly pertained to the 'work environment'. It is possible that some items may not fit semantically with the factor they load on. For example, in the same Tanzanian study, 5 items loaded substantively on another factor. Four of the items were related to 'management and supervision at the facility', but one item did not appear to fit with that definition: 'relationship with local leaders in the community'. One explanation for such cases is a divergence between respondent and researcher interpretation of an item. In this case, respondents may have considered community leaders together with managers given their joint involvement in facility governing committees. Interviews/focus groups can be used to shed light on respondent

¹ For instance, for principal axis factor, it's all factors with eigenvalues greater zero. For PCA, it's all factors with eigenvalues greater than 1 (the Kaiser criterion).

understanding and if the item is found to be related to the dimension it can be retained. Another reason for 'lone items' is that they are related to a sub-dimension of motivation that did not emerge as a separate multi-item factor simply because the scale contained only one item pertaining to it. In such cases, researchers must decide whether to keep the item as a (psychometrically suboptimal) single-item measure, or whether to drop it. In some cases, clusters of items that do not fit well together may be a statistical artefact: EFA groups items based on response patterns without considering how these items relate to each other semantically. The idea is that people respond similarly to items of similar content, because these items tap into the same construct. However, this may not be the case. For instance, a person might feel equally motivated by intrinsic and extrinsic factors, and thus assign similar numeric values to related items. In an EFA, we might then end up with a one-factor solution combining extrinsic and intrinsic motivation items. However, this does not mean that extrinsic and intrinsic motivation are the same. Therefore, it is important to interpret factor analysis results together with theory and knowledge of the context.

Step 6: Measurement Reliability

Reliability refers to the extent to which the measurement scale produces similar results under similar conditions (DeVellis 2012). Internal consistency based on Cronbach's alpha coefficient, the average correlation between items, is the most widely used statistic to assess measurement reliability. In recent years, however, psychometricians have cautioned against the use of alpha, for conceptual reasons (Yang & Green 2011) and due to its vulnerability to outliers, non-normal data, small number of items, and low variability in total scores (Greer et al. 2006) (Cortina 1993) (Sijtsma 2009) (Cronbach & Shavelson 2004). Factor-analysis-based estimation of reliability is now preferred to alpha (Yang & Green 2011) (Raykov & Marcoulides 2011). When estimating alpha, it is recommended to use the polychoric correlation matrix instead of a Pearson correlation matrix (Gadermann et al. 2012) (Dale 2014). For multidimensional measures of motivation, alpha should be estimated for each dimension (Cortina 1993). A typical recommended cut-off level for alpha has been 0.70, however, as this parameter depends on the number of items among other things, this value should be treated cautiously.

Test-retest measures the degree to which health workers would provide the same responses to items in a repeat survey. In public health studies where the scale development is not the central focus, test-retest validation studies are unfortunately often not feasible for practical reasons. If a retest is possible, it is important to choose the time delay between test and retest in a way that the underlying construct measured with the scale can be assumed to have remained stable.

When motivation is to be compared across different subgroups (e.g. women vs men, doctors vs other cadres, different language groups, across countries), the scale should be tested for equal measurement properties across subgroups. If measurement invariance is not established, there is a risk that subgroup differences are not due to differences in motivation, but to differences in the performance of the measure in the subgroups (Vandenberg & Lance 2000). Measurement invariance testing is usually done in a CFA framework (see Annex 3).

Step 7: Core Analysis

Once validity and reliability are established, the motivation measure can be used within analysis, depending on the objective of the study. If the objective is to describe motivation levels, item responses can be combined into a composite score, typically calculated as the arithmetic mean of health worker responses. Means can be calculated either as unweighted means, i.e. all items have the same weight, or one can give more weight to some items than to others, which may be preferable

if the EFA and/or CFA results show substantially different factor loadings between items. In such cases, the loadings can serve as weights. If the motivation measure was found to be multidimensional at the EFA/CFA stage, scores are calculated separately for each dimension. Researchers have sometimes also estimated an overall motivation score by combining item scores across dimensions e.g. (Gow et al. 2013) (Bhatnagar & George 2016; Hagopian et al. 2009) (Mbindyo et al. 2009b). Where dimensions are deemed conceptually distinct, such practice makes limited sense and risks evening out important differences across dimensions. If researchers wish to capture overall motivation, a related item can be included in the measurement scale (e.g. "Overall, how motivated do you feel?").

If the objective is to understand determinants or consequences of motivation, or change in motivation over time, there are two main analytical options: using composite scores ('manifest variables'), or using a latent variable approach where the relationship between motivation and other variables of interest are inferred directly from the scale items, without the estimation of composite scores.

Composite scores can be used as predictor or outcome variables in a regression model. However, much of the variance contained in the individual items is 'averaged out' by the calculation of a mean composite score (Borsboom 2006) (Skrondal & Laake 2001).

With a latent variable approach in SEM, associations between motivation and other variables of interest are directly estimated from the items via the latent variable/s. This approach provides more accurate estimates of the relationship between motivation and other variables as all information contained in the dataset is preserved. However, large structural equation models are complex and difficult to handle, and have large sample size requirements.

Step 8: Presenting findings

When reporting findings, it is important to be transparent as to the steps taken to generate results and decisions made during this process. It is standard practice to present all items used to measure motivation along with their mean scores and standard deviations. Results for EFA and/or CFA should be reported, including factor loadings for each item and model fit. If composite scores are calculated, mean scores and standard deviations should be reported. Spider diagrams or other graphs can be helpful to visualise composite scores and variance across dimensions and changes over time (Figure 3). If SEM is used, a visualisation of the model including parameter estimates can be informative in addition to model fit information (Figure 4).

Discussion

We have highlighted the steps involved in measuring and analysing health worker motivation survey data and the importance of having a clear conceptualisation of motivation as a single or multi-dimensional construct, prior to undertaking measurement.

We have described the use of exploratory or confirmatory factor analysis to identify or confirm motivation dimensions. Most of the existing health worker motivation literature in LMICs uses EFA (Alhassan et al. 2013) (Mbindyo et al. 2009b) (Bonenberger et al. 2014) (Chandler et al. 2009). There is potential for greater use of CFA, especially in studies that have clearly articulated dimensions of motivation, based on theory or prior formative research (e.g. (Weldegebriel et al. 2016) (Agyepong et al. 2004) (Ojaka et al. 2014) (Franco et al. 2004) (Hotchkiss et al. 2015)). Some studies had pre-defined motivation dimensions and presented a descriptive assessment of item scores and means across dimensions without employing factor analysis to validate these results (e.g. (Ojaka et al. 2014) (Ssengooba et al. 2007) (Dieleman et al. 2006) (Lephoko et al. 2006)). While descriptive analysis is an important first step in any motivation study, it is difficult to definitively assess how well

items measured each dimension, and whether assumptions about composition were accurate, without doing factor analysis. However, as tools become more widely used and validated in different contexts and languages, and our knowledge of motivation dimensions in these contexts grows, factor analysis may not always be required.

Much of the empirical research has been aimed at identifying the composition of motivation and factors driving motivation, looking at how these vary between groups and over time in response to policy change. In such cases, the focus is on the relative differences/changes over time/between groups rather than absolute levels. We have shown how composite scores can be calculated if the interest is in absolute motivation levels at a certain point in time. However, researchers should be cautious in the interpretation of these scores. Responses to questions about motivation may be affected by social desirability bias. For example, respondents may provide high scores on intrinsic motivators, items relating to commitment, punctuality, or attitude to work, regardless of how they really feel. This may not mean they really are highly intrinsically motivated, but might be a result of them 'anchoring' their responses differently for different dimensions and items. Careful design of the scale can shed light on such issues and inform interpretation. Most published studies have used qualitative methods to inform the design of the scale (Sacks et al. 2015) and/or as part of the research study (Chandler et al. 2009) to maximise content validity and facilitate an accurate interpretation of findings. Checking that associations between motivation measures and motivational outcomes and/or health worker characteristics conform to expectations is also important. A number of studies have examined and reported determinants of motivation to assess construct validity (e.g (Hotchkiss et al. 2015) (Franco et al. 2004)), however, this is not done systematically. Some studies have also examined the relationship between motivation and turnover intentions and performance outcomes in health workers in low and middle income settings (Bonenberger et al. 2014) (Alhassan et al. 2103). More extensive empirical research has examined this question in relation to other types of workers in high income settings (Deci et al. 2017). More research of this kind is needed in LMIC settings, in order to assess the validity of the motivational measure and also to understand the extent to which motivation acts as a mediator of better performance in different contexts and in response to different interventions.

As increasing efforts are made to improve the performance of health workers to provide more effective care in LMICS, researcher and policy interest in measuring and understanding motivation in surveys is likely to remain high. We hope this paper provides a useful introduction for those wanting to gain a better understanding of the methodology and the process of designing surveys to measure motivation in LMICs and the methods used to analyse and interpret their findings.

References

- Agyepong IA, Anafi P, Asiamah E, et al. 2004. Health worker (internal customer) satisfaction and motivation in the public sector in Ghana. *Int J Health Plann Manage*, **19**: 319-36.
- Alhassan RK, Spieker N, van Ostenberg P, et al. 2013. Association between health worker motivation and healthcare quality efforts in Ghana. *Hum Resour Health*. , **11**: 37.
- Alhassan RK, Spieker N, van Ostenberg P, et al. 2103. Association between health worker motivation and healthcare quality efforts in Ghana. *Hum Resour Health*. , **11**: 37.
- Bailey C, Blake C, Schriver M, et al. 2016. A systematic review of supportive supervision as a strategy to improve primary healthcare services in Sub-Saharan Africa. *Int J Gynaecol Obstet*, **132**: 117-25.
- Bakker W, van den Akker T, Mwagomba B, et al. 2011. Health workers' perceptions of obstetric critical incident audit in Thyolo District, Malawi. *Trop Med Int Health.*, **16**: 1243-50.
- Bandura A. 1982. Self efficacy mechanism in human agency. *American Psychologist*, **37**: 122-147.
- Bhatnagar A, George AS. 2016. Motivating health workers up to a limit: partial effects of performance-based financing on working environments in Nigeria. *Health Policy Plan*.
- Bhatnagar A, Gupta S, Alonge O, George AS. 2017. Primary health care workers' views of motivating factors at individual, community and organizational levels: a qualitative study from Nasarawa and Ondo states, Nigeria. *Int J Health Plann Manage*, **32**: 217-233.
- Bonenberger M, Aikins M, Akweongo P, Wyss K. 2014. The effects of health worker motivation and job satisfaction on turnover intention in Ghana: a cross-sectional study. *Hum Resour Health*, **12**: 43.
- Borghji J, Ataguba J, Mtei G, Akazili J, Meheus F, Rehnberg C, Di M. 2009. Methodological challenges in evaluating health care financing equity in data-poor contexts: lessons from Ghana, South Africa and Tanzania. *Adv Health Econ Health Serv Res.*, **21**: 133-56.
- Borsboom D. 2006. The attack of the psychometricians. . *Psychometrika*. , **71**: 425-40.
- Brown TA. 2006. *Confirmatory factor analysis for applied research*. Guilford Press, New York.
- Carifio J, Perla R. 2008. Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, **42**: 1150-1152.
- Carifio J, Perla RJ. 2007. Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences* **3**: 106-116.
- Cattell RB. 1966. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, **1**: 245-276.
- Chandler CI, Chonya S, Mtei F, Reyburn H, Whitty CJ. 2009. Motivation, money and respect: a mixed-method study of Tanzanian non-physician clinicians. *Soc Sci Med*, **68**: 2078-88.
- Chin-Quee D, Mugeni C, Nkunda D, et al. 2016. Balancing workload, motivation and job satisfaction in Rwanda: assessing the effect of adding family planning service provision to community health worker duties. *Reproductive Health*, **13**: 2.
- Chopra M, Munro S, Lavis JN, Vist G, Bennett S. 2008. Effects of policy options for human resources for health: an analysis of systematic reviews. *The Lancet*, **371**: 668-674.
- Cortina JM. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, **78**: 98-104.
- Cronbach LL, Shavelson RJ. 2004. My Current Thoughts on Coefficient Alpha and Successor Procedures *Educational and Psychological Measurement*, **64**: 391-418.
- Dale E. 2014. Performance based payments, provider motivation and quality of care in Afghanistan. John Hopkins University.
- de Winter JCF, Dodou D, Wieringa PA. 2009. Exploratory Factor Analysis With Small Sample Sizes. *Multivariate Behavioral Research*, **44**: 147-181.
- Deci EL, Olafsen AH, Ryan RM. 2017. Self-Determination Theory in Work Organizations: The State of a Science. *Annu. Rev. Organ. Psychol. Organ. Behav.* , **4**: 19-43.

- Deci EL, Ryan RM. 1985. *Intrinsic motivation and self-determination in human behavior*. Plenum, New York.
- DeCoster J. 1998. Overview of Factor Analysis. <http://www.stat-help.com/factor.pdf>.
- Denhardt RB, Denhardt JV, Aristigueta MP. 2008. *Managing human behaviour in public and nonprofit organisations*. . Thousand Oaks, CA.,.
- DeVellis RF. 2012. *Scale Development: Theory and Applications*. Sage.
- Dieleman M, Toonen J, Touré H, Martineau T. 2006. The match between motivation and performance management of health sector workers in Mali. *Hum Resour Health*, **4**.
- Engineer CY, Dale E, Agarwal A, et al. 2016. Effectiveness of a pay-for-performance intervention to improve maternal and child health services in Afghanistan: a cluster-randomized trial. *Int J Epidemiol*, **45**: 451-9.
- Fowler FJ. 2009. *Survey Research Methods*. 4th Edition edn. Sage.
- Franco LM, Bennett S, Kanfer R. 2002. Health sector reform and public sector health worker motivation: a conceptual framework. *Soc Sci Med*, **54**: 1255-66.
- Franco LM, Bennett S, Kanfer R, Stubblebine P. 2004. Determinants and consequences of health worker motivation in hospitals in Jordan and Georgia. *Soc Sci Med*, **58**: 343-55.
- Gadermann AM, Guhn M, Zumbo BD. 2012. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research and Evaluation*, **17**: 1-13.
- Global Health Workforce Alliance. 2014. A universal truth: no health without a workforce. Report of Third Global Forum on Human Resources for Health, Recife, Brazil. . World Health Organization.
- Gorsuch RL. 1983. *Factor analysis* Lawrence Erlbaum Associates. , Hillside, NJ.,.
- Gow J, George G, Mwamba S, Ingombe L, Mutinta G. 2013. An evaluation of the effectiveness of the Zambian Health Worker Retention Scheme (ZHWRs) for rural areas. *Afr Health Sci*, **13**: 800-7.
- Greer T, Dunlap WP, Hunter ST, Berman ME. 2006. Skew and internal consistency. *J Appl Psychol*. , **91**: 1351-8.
- Guilford JP. 1952. When not to factor analyze. *Psychological Bulletin*, **49**: 26-37.
- Hagopian A, Zuyderduin A, Kyobutungi N, Yumkella F. 2009. Job satisfaction and morale in the Ugandan health workforce. *Health Aff (Millwood)*, **28**: w863-75.
- Hayton JC, Allen DG, Scarpello V. 2004. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organisational Research Methods*, **7**: 191-205.
- Hotchkiss DR, Banteyerga H, Tharaney M. 2015. Job satisfaction and motivation among public sector health workers: evidence from Ethiopia. *Hum Resour Health*, **13**: 83.
- Huillery E, Seban J. 2014. Pay for performance, motivation and final output. Experimental evidence from the Democratic Republic of Congo.
- Inceoglu I, Segers J, Bartram D. 2012. Age-related differences in work motivation. *Journal of Occupational and Organizational Psychology*: 300-329.
- Kaiser HF. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**: 141-151.
- Källander K, Strachan D, Soremekun S, et al. 2015. Evaluating the effect of innovative motivation and supervision approaches on community health worker performance and retention in Uganda and Mozambique: study protocol for a randomised controlled trial. *Trials*, **16**: 157.
- Kline P. 1979. *Psychometrics and psychology*. . Acaderric Press, London.
- Kline RB. 2010. *Principles and practice of structural equation modeling*. . 3rd revised edition. edn. Taylor & Francis, New York.
- Leonard KL, Masatu MC. 2010. Professionalism and the know-do gap: exploring intrinsic motivation among health workers in Tanzania. *Health Econ*, **19**: 1461-77.

- Lephoko CS, Bezuidenhout MC, Roos JH. 2006. Organisational climate as a cause of job dissatisfaction among nursing staff in selected hospitals within the Mpumalanga Province. *Curationis*, **29**: 28-36.
- Little RJA. 1992. Regression With Missing X's: A Review. *Journal of the American Statistical Association*, **87**: 1227-1237.
- Little TD, Lindenberger U, Nesselroade J. 1999. On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, **4**: 192-211
- Lohmann J, Houfort N, De Allegri M. 2016. Crowding out or no crowding out? A Self-Determination Theory approach to health worker motivation in performance-based financing. *Social Science & Medicine*, **169**: 1-8.
- Lohmann J, Souares A, Tiendrebéogo J, et al. 2017. Measuring health workers' motivation composition: validation of a scale based on Self-Determination Theory in Burkina Faso. *Hum Resour Health*, **15**: 33.
- MacCallum RC, Widaman KF, Zhang S, S. H. 1999. Sample size in factor analysis. *Psychological Methods*, **4**: 84-99.
- Manongi RN, Marchant TC, Bygbjerg IC. 2006. Improving motivation among primary health care workers in Tanzania: a health worker perspective. *Hum Resour Health*, **4**: 6.
- Mbaruku GM, Larson E, Kimweri A, Kruk ME. 2014. What elements of the work environment are most responsible for health worker dissatisfaction in rural primary care clinics in Tanzania? *Hum Resour Health*, **12**: 38.
- Mbindyo P, Gilson L, Blaauw D, English M. 2009a. Contextual influences on health worker motivation in district hospitals in Kenya. *Implement Sci*, **4**: 43.
- Mbindyo PM, Blaauw D, Gilson L, English M. 2009b. Developing a tool to measure health worker motivation in district hospitals in Kenya. *Hum Resour Health*, **7**: 40.
- Morrison J, Batura N, Thapa R, Basnyat R, Skordis-Worrall J. 2015. Validating a tool to measure auxiliary nurse midwife and nurse motivation in rural Nepal. *Human Resources for Health* **13**.
- Mutale W, Ayles H, Bond V, Mwanamwenge MT, Balabanova D. 2013. Measuring health workers' motivation in rural health facilities: baseline results from three study districts in Zambia. *Hum Resour Health*, **11**: 8.
- Ojaka D, Olango S, Jarvis J. 2014. Factors affecting motivation and retention of primary health care workers in three disparate regions in Kenya. *Hum Resour Health*, **12**: 33.
- Pinder CC. 2008. *Work Motivation in Organizational Behavior*. Psychology Press.
- Prytherch H, Leshabari MT, Wiskow C, et al. 2012. The challenges of developing an instrument to assess health provider motivation at primary care level in rural Burkina Faso, Ghana and Tanzania. *Glob Health Action*, **5**: 1-18.
- Raykov T, Marcoulides GA. 2011. *Introduction to Psychometric Theory* Routledge, East Sussex.
- Rodriguez Pose R, Engel J, Poncin A, Manuel S. 2014. Against the odds: Mozambique's gains in primary health care. Overseas Development Institute.
- Sacks E, Alva S, Magalona S, Vesel L. 2015. Examining domains of community health nurse satisfaction and motivation: results from a mixed-methods baseline evaluation in rural Ghana. *Hum Resour Health*, **13**: 81.
- Sijtsma K. 2009. On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, **74**: 107-120.
- Skrondal A, Laake P. 2001. Regression among factor scores. *Psychometrika*, **66**: 563-75.
- Ssengooba F, Rahman SA, Hongoro C, et al. 2007. Health sector reforms and human resources for health in Uganda and Bangladesh: mechanisms of effect. *Human Resources for Health*, **5**.
- Streiner DL, Norman GR, Cairney J. 2008. *Health Measurement Scales: A practical guide to their development and use*. Oxford University Press, Oxford.
- Tabachnick BG, Fidell LS. 2007. *Using multivariate statistics in medicine*. Allyn & Bacon, Boston, MA.

- Vandenberg RJ, Lance CE. 2000. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organisational Research Methods*, **3**: 4-70.
- Weldegebriel Z, Ejigu Y, Weldegebreal F, Woldie M. 2016. Motivation of health workers and associated factors in public hospitals of West Amhara, Northwest Ethiopia. *Patient Preference Adherence*, **10**: 159-69.
- Williams B, Brown T, Onsmann A. 2012. Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, **8**: 1-13.
- Willis-Shattuck M, Bidwell P, Thomas S, et al. 2008. Motivation and retention of health workers in developing countries: a systematic review. *BMC Health Services Research*, **8**: 247.
- World Health Organization. 2006. Working Together for Health. The World Health Report 2006. World Health Organization.
- Yang Y, Green SB. 2011. Coefficient Alpha: A Reliability Coefficient for the 21st Century? *Journal of Psychoeducational Assessment*, **29**.
- Yong AG, Pearce S. 2013. A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, **9**: 79-94.