TITLE: Measuring hope amongst Tanzanian women who participate in microfinance: An evaluation of the Snyder hope scale

RUNNING TITLE: An evaluation of the Snyder hope scale

JOURNAL: Global Public Health

AUTHORS:

Christian Holm Hansen[1 2], Shelley Lees[3], Saidi Kapiga[4 5], Janet Seeley[1 3], Tony Barnett[3 6]

AFFILIATIONS:

1 MRC/UVRI and LSHTM Uganda Research Unit, Entebbe, Uganda

2 MRC Tropical Epidemiology Group, London School of Hygiene & Tropical Medicine, UK

3 Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, UK

4 Mwanza Intervention Trials Unit, National Institute for Medical Research, Tanzania.

5 Department of Infectious Diseases Epidemiology, London School of Hygiene & Tropical Medicine, UK

6 Humanitarian and Conflict Response Institute, University of Manchester, UK

CORRESPONDING AUTHOR: Dr Christian Holm Hansen, MRC Tropical Epidemiology Group, London School of Hygiene & Tropical Medicine, Keppel Street, WC1E 7HT, UK, christian.hansen@lshtm.ac.uk, tel. +44(0)207 958 8255.

WORD COUNT: circa 4,600

TABLES / FIGURES: 2 tables and 3 figures

TRIAL REGISTRATION NUMBER: ClinicalTrials.gov– ID NCT02592252

ABSTRACT

Measuring hope reliably and accurately remains an important research objective, not least in less prosperous settings where 'holding on to hope' may be critically important in the struggle against adverse life conditions. The State Hope Scale devised by Snyder (1996) was designed for use in the US. Despite reported application in diverse cultures and using translations the scale has not been extensively validated outside US populations. This study contributes to a larger project exploring the measurement of hope and provides a critique of Snyder's scale as used in a Tanzanian female population of 1,021 urban microfinance participants. We evaluate the scale's validity through assessment of the empirical distribution of scores, item response profiles, internal consistency and discriminatory ability.

Participants mostly scored very high and many reached very near the maximum attainable score. Hardly any endorsed the negative half of the response scale. Several problems are discussed including poor discrimination and strong evidence of acquiescence response bias. We also found little association of the scale scores with hypothesised correlates of hope. Future improvements on the measurement of hope are recommended, especially in studies outside the narrow Western context in which the scale was devised.

KEY WORDS: Snyder state hope scale, acquiescence response bias, scale validity, ceiling effect.

INTRODUCTION

This paper contributes to a larger project exploring the measurement of 'hope' as a way to distil what are described as the `structural factors' in the transmission of HIV infection (Barnett, Seeley, Levin, & Katongole, 2015) building on previous theoretical work (Barnett, Fournié, Gupta, & Seeley, 2015; Barnett and Weston, 2008; Bernays, Rhodes, & Barnett, 2007). Barnett, Seeley, et al. (2015) describe the application of a hope scale (the State Hope Scale, devised by Snyder (1996), hereafter referred to as the Snyder scale) which had been developed and validated in the United States of America, in Kuwait (Abdel-Khalek and Snyder, 2007), and in two different rural settings in Uganda. They conclude that in Uganda the Snyder scale, which had been translated into the local language, Luganda, was understood by study participants and that differences between two populations living in different settings could be detected in the scores obtained from the scale. However, Barnett and colleagues conclude that while differences were detected, which may reflect the perceptions and actions of people living in two different structural settings (an agriculture-based community and a fishing community), more work is required with larger sample sizes to test further the utility of the scale.

The Snyder scale consists of two axes: Agency (the process of thinking of one's goals along with the motivation to move towards them); and Pathway (ways to achieve these goals). This approach to measuring hope defines the concept `as a state where goals are formulated and realistic pathways are identified through which those goals may be pursued' (Barnett, Seeley, et al., 2015, p. 423). As such the scale is designed to capture both the person's aspirations and their knowledge of how those aspirations might be achieved.

Understanding and measuring hope reliably and accurately remains an important research objective, not least in less prosperous settings where 'holding on to hope' may be critically important in the struggle against adverse life conditions. The Snyder scale was designed for use in the US, and with only two exceptions it was validated for use in US populations as an aid to understanding and assisting people's career choices (Snyder, et al., 1996). The origin of this approach is in the US tradition of 'positive psychology' – helping people to be motivated to achieve their goals. It remains unclear how well this Western derived concept of hope and resultant Snyder scale fits outside the context in which they were developed.

There have been reported tests of the scale in a number of language or culture groups including French (Dubé, Lapierre, Bouffard, & Alain, 2007), Dutch (Carfio and Rhodes, 2002), Slovak (Halama, 2001), Chinese (Ho, 2003), Spanish (McDermott et al., 1997) and Korean (Yu, 2003). The reported findings of these tests are generally partial and, in some cases, unpublished or not peer reviewed and developed variously for use with adolescents, undergraduates and, in one case, retired people.

In this study we examine data collected using the Snyder hope scale from 1,021 Tanzanian women who participated in microfinance loan group schemes. We provide a critique of the scale as used in this population and evaluate its validity through assessment of the empirical distribution of scores, item response profiles, internal consistency and the scale's discriminatory ability.

STUDY PARTICIPANTS

The data presented in this article were collected between September 2014 and June 2015 as part of the MAISHA trial baseline survey. Set in Mwanza city (northwest Tanzania) MAISHA (Component A) was a 2-arm randomised trial of a participatory group intervention to improve relationship dynamics and reduce intimate partner violence. Participants were all-female microfinance loan group members with BRAC (one of the leading microfinance providers in Tanzania). A total of 66 loan groups from Mwanza city were included in the study and within each group the study team invited women to participate who reported ever having been in an intimate relationship with a man. Across all loan groups 1,021 women agreed to participate and completed the baseline survey questionnaire. The baseline survey was conducted prior to randomisation and any intervention activities and was administered by a team of research assistants using handheld tablets programmed in ODK for electronic data capture.

Alongside the quantitative data, qualitative data were collected from a subset of women. Immediately following the baseline survey, in-depth interviews were completed with 18 women, and 12 focus groups were conducted with women in their microfinance groups to explore their everyday lives, experiences with intimate partner relations and violence. In order to understand hope conceptually the items on the Snyder scale were developed into in-depth questions at the end of the interview or group discussion. Further details of the background, aims and methods of the trial can be found in the published protocol (Harvey et al., 2018).

The median age of women in the study was 39 years (interquartile range 33-46 years). Most women were either married or living with a male partner (73%) and only a minority had

completed education beyond primary school (22%). Half of all participants reported living in a household with six or more people. 27% reported that they had been subjected to either physical or sexual violence in the last 12 months by their male partners. The prevalence of violence and abuse experiences, and associated factors, among the study participants have been reported elsewhere along with further results from the baseline survey (Kapiga et al., 2017).

THE SNYDER HOPE SCALE AND HOW IT IS SCORED

The Snyder hope scale that was administered to the study participants is shown in table 1. The instrument consists of four pathway subscale items (1, 3, 4 and 5) and four agency subscale items (2, 6, 7 and 8). The remaining four are so-called 'filler' items and do not contribute to the total or subscale scores. Each of the 12 statements was read out to the study participants who used a six-point scale to indicate the extent to which they agreed with the statement. (For this study the scale was translated into Swahili and back-translated into English. The English version varies slightly in wording from the original hope scale. A Likert scale of 6 points was used, rather than the 8 points in the original scale as Swahili does not allow for such subtle differences as between 'slightly' and 'somewhat' true or false. The modified response scale was first used in Uganda where is was adapted from the original 8-point version and then tested in two distinct populations (Barnett, Seeley, et al., 2015).) The six answer options were, with the scores attributed to each answer shown in parentheses, definitely true (6), mostly true (5), slightly true (4), slightly false (3), mostly false (2), and definitely false (1). For each participant a hope total score was derived as the sum of their eight item scores (ignoring the four 'filler' items).

Since item scores range from 1 to 6, the hope total score ranges from a possible minimum of 8 to a maximum of 48 which indicates the highest possible level of hope. The pathway and agency subscale scores are similarly calculated as the sum of the relevant four item scores, and both range from a possible minimum of 4 to a maximum of 24.

Before examining the data, it is useful to reflect briefly on the limitations of the algorithm that generates a respondent's hope score. Common to many psychometric multi-item scales the method for deriving the overall hope scale score relies on strong assumptions. The validity of the total score, derived simply as the sum of the item scores, rests on the assumption that each item of the questionnaire is measuring the same theoretical construct, and to the same extent; or at least that the constituent parts of the hope construct, as measured by the eight items, are equally important and should be given equal quantitative weight in the summation. The other key assumption is the validity of the regular spacing of integer answer values which implies, for instance, that the difference between considering a statement 'mostly false' and 'slightly true' is exactly equivalent to the difference between considering the (same or another) statement 'slightly true' and 'definitely true'. There are examples of multi-item scales that employ more involved, well-validated algorithms for deriving a single quantitative measure to summarise a respondent's level of the underlying construct. Such examples include the EQ-5D (measuring health-related quality of life) and the SF-36 mental and physical component summaries (measuring general health status). [5,6,7]

THE DISTRIBUTION OF HOPE?

In total 1,021 participants completed the hope scale. Histograms showing the hope total scores, and the pathway and agency subscale scores, are shown in figure 1. The average hope total score was 39.2; however, between the scale anchor points of 8 (least hopeful) and 48 (most hopeful), scores are not easily interpreted. A score of 39.2 could be interpreted as being (39.2-8)/(48-8) = 82% hopeful, or, if divided by eight thereby mapping the total score back onto the six-point answer scale, the interpretation of the overall score might be that it is (39.2/8 ≈ 5) 'mostly true', overall, that the participant has hope.

[Table 1 and figure 1 around here]

Aside from the average score, it is useful to consider the spread of individual participant scores across the entire scale. Although the possible range of scores extends from 8 to 48 the lower part of the scale is hardly utilised at all in the sample. In fact, 96% of scores are concentrated on the range of the scale that is above 28, the midway point between least and most hopeful. Perhaps the MAISHA trial participants, despite adversity in life, are actually full of hope. After all, the respondents have self-selected to participate in a trial of an intervention aimed at improving their situation (see, for reference, the literature on the participation in microfinance programmes such as Halder and Mosley, 2004; Steinert et al., 2018); it seems likely that the least hopeful of women would see little point and hence avoid participating in such a trial. However, for a scale to have good discriminatory ability it needs to accommodate the full spectrum of experiences in the population, whether at the extreme lower or upper end. While there are discernible differences in the levels of hope among those who have little of it, at the

other end of the spectrum, it is not possible to differentiate easily between those with high hope. In other words, it would seem there is a frequently reached natural upper limit to how much hope one can harbour whereas no such limit exists as regards hopelessness (or at least it is not a limit that is reached). If this is congruent with our understanding of hope, the Snyder scale may well be the perfect instrument for measuring it. More likely, the scale is inadequate in describing hope in this population; it is unable to differentiate among respondents with higher hope and there is a clear ceiling effect whereby many respondents attain the maximum (or near-maximum) score.

Some effects of this are poor reliability and poor sensitivity to change, e.g. when trying to correlate the measure with associated factors, possible predictors or interventions received. Stochastic error can be considerable in self-reported outcomes. When a tool fails to discriminate well between participants with differing levels of the underlying construct (due to observations being mostly squeezed up against the maximum attainable score) even limited random variation in a respondent's answers can affect a large shift in the nonparametric ranking of their score and consequently the percentile that the respondent falls within. Those scoring in the 90th percentile, for example, may differ qualitatively very little from those in the 80th or 70th percentiles.

A close look at the item scores

Having considered the skewed distribution of the Snyder total scores, we now turn our attention to the individual item scores from which the totals were derived. The responses given by study participants to each of the eight statements are shown in figure 2. Overwhelmingly

participants reported that they agreed with the statements and were most likely to respond 'definitely true'. Two of the statements measuring agency, '*You've been pretty successful in life'* and '*You meet the goals that you set for yourself*' received relatively more mixed responses but were still very much regarded as true. Overall, the negative half of the item-response scale was hardly endorsed at all.

'Definite agreement' was the most common response by study participants which might be a sign that the six-point response scale was not adequately explained to or embraced by respondents. It is possible that many considered their answer a binary choice between categorical agreement or disagreement and that the nuances of the response scale did not translate well for cultural or language reasons.

[Figure 2 and 3 around here]

THE 'YES' BIAS

It is well-established in the context of surveys that respondents are more likely to agree than to disagree with a statement (Messick and Jackson, 1961). To mitigate this response bias, known as acquiescence, many scales are designed with half the items worded such that agreement contributes negatively to the outcome in question while the other half are worded such that agreement contributes positively. All eight items of the Snyder hope scale are worded such that agreement with the statements indicate more hope on the part of the respondent. It seems possible, then, that the scale might be artificially exaggerating or overestimating what it is trying to measure.

An opportunity arose to investigate the possibility of acquiescence bias when, a couple of years after the baseline survey was completed, the participants were asked to complete the hope scale again as part of the MAISHA trial follow-up questionnaire. The hope scale that was presented to participants on this occasion was modified slightly in that half the statements were now negatively phrased. Thus in the modified version items a, f, i and l were changed respectively to '*You cannot think of many ways to get out of a difficult situation when life evolves unexpectedly*', '*You cannot think of many ways to get the things in life that are important to you*', '*Your past experiences have not prepared you well for your future*' and '*You do not meet the goals that you set for yourself*'. All other items were left unchanged from the baseline survey. In total 919 of the participants completed this repeat administration of the questionnaire.

The pattern of responses to the four *un*changed questionnaire items at follow-up was almost identical to that observed at baseline with about 90% of responses expressing agreement with the statements presented, i.e. answers of either 'slightly true', 'mostly true' or 'definitely true'. However, for the four statements that had been negatively worded on the second occasion there was a marked shift in the response pattern. Whereas 90% were in agreement at baseline with the positively worded statements, once reworded, only 73% expressed *dis*agreement with the negatively worded statements. This represents a 2- to 3-fold increase in expressions of poor hope. The inconsistency is illustrated in figure 3 which shows that the two sets of statements (a, f, i, l versus b, d, h, j) drew comparable responses at baseline, and that responses to the statements that remained unchanged (b, d, h, j) were comparable at baseline and follow-up. Hence, there is little suggestion of confounding by time or question content. When estimated in

a linear model adjusted for timepoint and question set (as well as randomised treatment allocation), the effect of re-wording half the statements, as compared to leaving them unchanged, on the total score is a reduction of 2.1 units (95% CI: 1.82 to 2.35; $p<0.001$) providing strong evidence that the 'yes' bias has a real effect on this measure.

ASSOCIATIONS WITH HYPOTHESISED CORRELATES

Another important step in assessing the validity of the scale is to evaluate the extent to which the Snyder scale scores vary according to measured factors that are theoretically associated with hope. If the hope construct is captured well by the scale, we would expect the scale scores to be strongly associated with variables thought to be determinants or correlates of hope such as education, wealth, mental health status and relationship violence. Table 2 shows the mean Snyder scale scores, and accompanying standard deviations, stratified by differing levels of the four variables just mentioned. Also shown are the results of regression analysis with these variables.

The results reveal very little systematic variation in the scores irrespective of participants' level of education, wealth, exposure to relationship violence or mental health score. The mean Snyder hope scores in the sub-groups defined by levels of the first three variables diverge no more than 1.5 units in either direction away from the overall average of 39.2, which is negligible in magnitude relative to a 41-point scale. The strongest predictor of hope was the mental health score as measured on the SRQ-20, a screening tool for common mental disorders developed by the WHO for use in low- and middle-income countries (World Health Organisation, 1994). When modelled as a continuous variable in a linear regression, this

variable reduced the hope score by 0.25 for every unit increase on the 20-point SRQ scale. Despite the small p-value for the association between the mental health score and hope, the correlation coefficient (r = 0.21) was very low, and the SRQ-20 variable explained just R^2 = 4.5% of the total variation in the data.

The standard deviations around the stratum-specific means shown in table 2 are high and not dissimilar from the overall standard deviation of 5.5 units. This suggests poor discrimination according to the four selected attributes and is further evidence of little systematic variation and poor ability to predict a person's hope score from underlying person characteristics thought to be associated with hope. The very low R^2 statistics also reflect that hardly any of the variation in the Snyder hope scores is explained by associations with theoretical correlates.

[Table 2 around here]

INTERNAL CONSISTENCY

Cronbach's alpha coefficient is a measure of internal consistency (Cronbach, 1951). The alpha coefficient derived from these data was 0.73 and measures consistency across the scale's eight items, i.e. the extent to which the component items all measure the same thing. A score of α=0 indicates no relationship between items, and α=1 indicates perfect correlation between item scores. The alpha coefficient from these data can be interpreted as 73% of the total variation in the Snyder hope scores being attributable to differences in respondents' actual levels of hope. This is possibly high enough for a research tool with which to compare groups but is much too low for reliable prediction of individual respondents' level of hope and is consistent with the lack of strong associations with hypothesised correlates in the data (Bland and Altman, 1997).

BEYOND THE NUMBERS

The questions on the Snyder scale were expanded as in-depth questions to explore the themes that constructed the hope scale. The results revealed that all the women were aware of structural and interpersonal barriers that made their lives difficult but reported being optimistic about the future and their ability to improve their lives and those of their children. When asked 'do you think that there is hope (*matumaini*) for a better future?', most of the women said 'yes'. When asked what they were hopeful about, participants focused their discussions on the success of their (often petty) businesses and the future education and achievements of their children. Importantly, many women talked of emancipation from men's power over their financial matters, their businesses, their property, and their children. Their hopes for this were, however, contingent on their microfinance loans which enabled them to expand their businesses (albeit in a limited way) and pay for their children's school fees. This hope was also articulated as goals (*malengo*), which included access to repeated loans, acquiring assets (cars, property, sewing machines), educating their children, and growing businesses. Their ultimate goal, in achieving these, was to increase their standing in the community. Women's hope for their better futures was thus largely limited to increasing income and gaining some power within the household but did not encompass aspirations to change their lives through political and structural change. These qualitative data suggest that, drawing on Giddens' theory of structuration (Giddens, 1984), these women have high hope but focus on aspirations that are achievable, exhibiting *pragmatic agency* (Giddens, 1984; Kockelman, 2007; Ortner, 2006). Their hope does not encapsulate expectations of great wealth but pragmatic marginal improvement in their material circumstances. This is not reflected by the Snyder scale scores, perhaps

reflecting the pragmatic modesty of these women's aspirations in the face of structural constraints to their agency.

DISCUSSION

In this study we examined data obtained with the Snyder hope scale from Tanzanian urban women who were members of microfinance loan groups and had agreed to participate in a trial of an intervention to address challenging relationship dynamics and intimate partner violence. Despite adversity in life among the respondents, the hope scores were heavily skewed towards the upper range of the scale and many reached the maximum attainable score or very near to it. Only a very small minority scored below 28, the halfway point between the maximum and minimum score., and the distribution indicated a strong ceiling effect with the instrument being unable to measure variety and discriminate among respondents with high hope. This renders the scale inadequate in the same way that a temperature scale that is capped at 20 degrees Celsius would be inadequate for measuring annual day-time temperature variations in England. A useful measurement scale requires a range of potentially attainable scores that extends beyond the range of actually observed scores.

The individual item responses showed that definite agreement was the most common response given, and for many of the statements, practically none of the participants expressed any disagreement at all. Designing instrument items that yield more varied and nuanced responses is necessary for ensuring a less skewed total score and an overall measurement scale with improved discriminatory ability.

There was evidence that the high levels of agreement with statements were at least partially the result of acquiescence response bias which may have accounted for increases in the hope total score comparable in magnitude to effects associated with the hypothesised correlates presented in table 2. The phrasing of items should be balanced such that half express less of the underlying trait being measured and the remaining half express more of the trait.

There was very little evidence of systematic variation in participants' scores according to factors that are in theory related to hope. Less than 6% of the variation was explained by the combined effects of education, wealth, mental health and relationship violence. With so little of the total hope score explained by presumed correlates it is unclear how useful the scale is as a research tool in this population and arguable what it measures.

The Snyder scale was devised in a US population and has been used successfully in the US and apparently with undergraduates in Arabic translation in Kuwait, data from our study shows that the scale does not translate well in settings in Africa (Uganda and Tanzania) with people without any considerable formal education who speak related languages from the Bantu language group. (The study by Boyce in South Africa where some people speak languages from the Bantu group used an English version of the scale (Boyce, 2013).)

Thinking of alternative instruments for measuring hope in this population, it is interesting that most variation was observed among responses to the statements 'You've been pretty successful in life' and 'You meet the goals that you set for yourself', both on the agency subscale and the only items requiring the respondent to assess accomplishment. In a recent mixed qualitative and quantitative methods study from South Africa, Abler et al. (2017) devise

an alternative to the Snyder scale for measuring hope among adolescent women in a low resource, high prevalence HIV setting. The scale is carefully developed with items based on formative research aimed at understanding young South African women's concept of hope through in-depth interviews and focus group discussions and is shown to have high reliability and internal consistency after being tested on 2,533 participants. Despite the formative research suggesting a hope construct consisting of three distinct domains, the quantitative factor analysis did not support this, resulting in a final 12-item uni-dimensional scale with the authors proposing that the scale could possibly be shortened further. Also of note, the study only demonstrates weak correlations of the new hope scale with hypothesised correlates and, like the Snyder scale, all the items are worded such that agreement with the statements indicate higher levels of hope. The authors conclude that a limitation of the study is the little variation that was observed in the hope scores due to most of the individual statements being endorsed by 90% of the respondents which could have resulted in a ceiling effect.

The concept of hope is abstract: a subjective feeling which arguably manifests itself differently in diverse cultures and settings around the world with external structure imposing differently on individual agency and pathways. In our qualitative interviews, we asked women a simple question, 'whether they were hopeful about the future', allowing them to define – independently of the researchers' interpretation of the construct – what hope is for them. In recent work by C. Desmond and colleagues, Centre for Rural Health, University of KwaZulu-Natal, South Africa, (personal communication), the idea of a single-item instrument to measure hope is developed further.

The main strengths of our study include the large number of participants who completed the scale and the standardised mode of electronic data capture executed by a total of 12 carefully trained research assistants. Further, the availability of follow-up data from a repeat administration of the Snyder scale, but with half the items rephrased negatively, provided a unique opportunity to assess acquiescence bias.

Due to the nature of the MAISHA trial, the participants were female-only, and exclusively members of microfinance loan groups. While it is possible that microfinance might have positively influenced participants' hope, this effect is unlikely to have been so strong and generalised (and in such stark contrast to the negligible influences observed from other hypothesised correlates of hope) to have caused the heavy right skew and ceiling effect observed in our study. Although the participants may not constitute a true general population sample, the characterisation of the Snyder scale, with its limitations that have been presented in this article, is unlikely to be a mere artefact of the somewhat select trial population which in any case constitutes a subset of the general population. The Snyder scale was administered towards the end of a much longer questionnaire and it is possible that some respondents experienced a degree of questionnaire fatigue towards the end which could have resulted in poor engagement with the instrument.

Conclusion

In this study we have shown that the Snyder scale, as applied to a population of Tanzanian women who participate in microfinance loan groups, has several undesirable properties including strong skewness and ceiling effect and resultant poor discrimination, reliability and

sensitivity to change possibly explaining the lack of strong associations with hypothesised correlates. There was also strong evidence that the instrument invites significant acquiescence bias.

Future studies that aim to measure hope through multi-item scales should ensure that endorsement of individual questions or statements contributes positively to the outcome for exactly half of the items and negatively for the remaining half. To avoid a strong ceiling effect measurement scales should be designed to ensure that only a very small minority of respondents answer in the extreme, e.g. by asking about levels of agreement or disagreement with statements such as 'I feel confident that from now on things are only going to get better' or 'I have never been more hopeful for the future than I am now'. Finally, for settings where hope might manifest itself differently due to a relative lack of prosperity and individual freedom to affect life changes, consideration should be given to the inclusion of a single-item instrument asking respondents to rate their levels of hope for the future on a Likert or continuous scale.

FUNDING

MAISHA is supported by the STRIVE Research Programme Consortium funded by UK Aid from the Department for International Development (DFID) and another donor, who wishes to remain anonymous. The views expressed in this paper do not necessarily reflect the Department's official policies. The funding bodies and sponsor have had no role in the design of the study or in writing this manuscript, and will not have any role in its conduct, analyses and interpretation of data, or decisions to disseminate the results. CH is supported by the UK Medical Research Council (MRC) and the UK Department for International Development (DFID)

under the MRC/DFID Concordat agreement, which is also part of the EDCTP2 programme supported by the European Union (Grant Ref: MR/R010161/1)

ACKNOWLEDGEMENTS

We wish to thank all study participants for their time and commitment to the study. We are also grateful to the MAISHA study team for their contribution and tireless dedication to implementing the study in Tanzania and to the administration teams at MITU and LSHTM for their support.

REFERENCES

Abdel-Khalek, A., & Snyder, C. R. (2007). Correlates and predictors of an Arabic translation of the Snyder Hope Scale. *Journal of Positive Psychology, 2*(4), pp. 228-235.

Abler, L., Hill, L., Maman, S., DeVellis, R., Twine, R., Kahn, K., . . . Pettifor, A. (2017). Hope Matters: developing and validating a measure of future expectations among young women in a high HIV prevalence setting in rural South Africa (HPTN 068). *AIDS and Behavior, 21*(7), pp. 2156-2166.

Barnett, T., Fournié, G., Gupta, S., & Seeley, J. (2015). Some considerations concerning the challenge of incorporating social variables into epidemiological models of infectious disease transmission. *Global public health, 10*(4), pp. 438-448.

Barnett, T., Seeley, J., Levin, J., & Katongole, J. (2015). Hope: A new approach to understanding structural factors in HIV acquisition. *Global public health, 10*(4), pp. 417-437.

Barnett, T., & Weston, M. (2008). Wealth, health, HIV and the economics of hope. *AIDS, 22*(Suppl 2), p S27.

Bernays, S., Rhodes, T., & Barnett, T. (2007). Hope: a new way to look at the HIV epidemic. *AIDS, 21*, pp. S5-11.

Bland, J., & Altman, D. (1997). Cronbach's alpha. *British Medical Journal, 314*, p 572.

Boyce, G. (2013). *The relationship between perceptions of racial identity, hope and educational aspirations amongst young people in South Africa* (University of KwaZulu-Natal, Durban.

Carfio, J., & Rhodes, L. (2002). Construct validities and the empirical relationships between optimism, hope, self-efficacy and locus of control. *Journal of Prevention, Assessment and Rehabilitation, 19*, pp. 125-136.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, pp. 297-333.

Dubé, M., Lapierre, S., Bouffard, L., & Alain, M. (2007). Impact of a personal goals management program on the subjective well-being of young retirees. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 57*(3), pp. 183-192.

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration* Oxford: Polity Press.

Halama, P. (2001). The Slovak version of Snyder's Hope Scale. *Ceskoslovensak-Psychologie, 45*, pp. 132-142.

Halder, S. R., & Mosley, P. (2004). Working with the ultra-poor: learning from BRAC experiences. *Journal of International Development, 16*(3), pp. 387-406.

Harvey, S., Lees, S., Mshana, G., Pilger, D., Holm Hansen, C., Kapiga, S., & Watts, C. (2018). A cluster randomised controlled trial to assess the impact on intimate partner violence of a 10-session participatory gender training programme delivered to women taking part in a group-based microfinance loan scheme in Tanzania (MAISHA CRT01): study protocol. *BMC Women's Health, 18*(1), p 55.

Ho, S. M. Y. (2003). *Hope in Hong Kong*. University of Hong Kong. Hong Kong, China.

Kapiga, S., Harvey, S., Muhammad, A., Stoeckl, H., Mshana, G., Hashim, R., . . . Watts, C. (2017). Prevalence of intimate partner violence and abuse and associated factors among women enrolled into a cluster randomised controlled trial in northwestern Tanzania. *BMC Public Health, 17*, p 190.

Kockelman, P. (2007). Agency: The Relation between Meaning, Power, and Knowledge *Current Anthropology, 48*(3), pp. 375-401.

McDermott, D., Hastings, S., Gariglietti, K., Gingerich, K., Callahan, B., & Diamond, K. (1997) *A cross-cultural investigation of hope in children and adolescents*. Paper presented at the Annual Meeting of the American Psychological Association August 15-19, Chicago, Illinois, USA.

Messick, S., & Jackson, D. (1961). Acquiescence and the factorial interpretation of the MMPI. *Psychological Bulletin, 58*, pp. 299-304.

Ortner, S. B. (2006). *Anthropology and Social Theory: Culture, Power, and the Acting Subject* Durham: Duke University Press.

Snyder, C. R., Sympson, S. C., Ybasco, F. C., Borders, T. F., Babyak, M. A., & Higgins, R. L. (1996). Development and Validation of the State Hope Scale. *Journal of Personality and Social Psychology, 2*, pp. 321-335.

Steinert, J. I., Zenker, J., Filipiak, U., Movsisyan, A., Cluver, L. D., & Shenderovich, Y. (2018). Do saving promotion interventions increase household savings, consumption, and investments in Sub-Saharan Africa? A systematic review and meta-analysis. *World Development, 104*, pp. 238-256.

World Health Organisation. (1994). *A user's guide to the self-reporting questionnaire (SRQ)*. Geneva

Yu, N. M. (2003). *Translating the Hope Scale into Korean*. Yonsei University. Soeul, Korea.

Table 1 The Snyder hope scale consists of four pathway and four agency items in addition to four filler items that do not contribute to the hope total or subscale scores. Respondents are asked to indicate how much they agree with each statement.

Item #	Subscale	Statement
1	Pathway	You can think of many ways to get out of a difficult situation when life evolves unexpectedly
2	Agency	You energetically pursue your goals
Filler	-	You feel tired most of the time
3	Pathway	There are lots of ways around any problem
Filler	-	You are easily downed in an argument
4	Pathway	You can think of many ways to get the things in life that are important to you
Filler	-	You worry about your health
5	Pathway	Even when others get discouraged, you know you can find a way to solve the problem
6	Agency	Your past experiences have prepared you well for your future
7	Agency	You've been pretty successful in life
Filler	-	You usually find yourself worrying about something
8	Agency	You meet the goals that you set for yourself

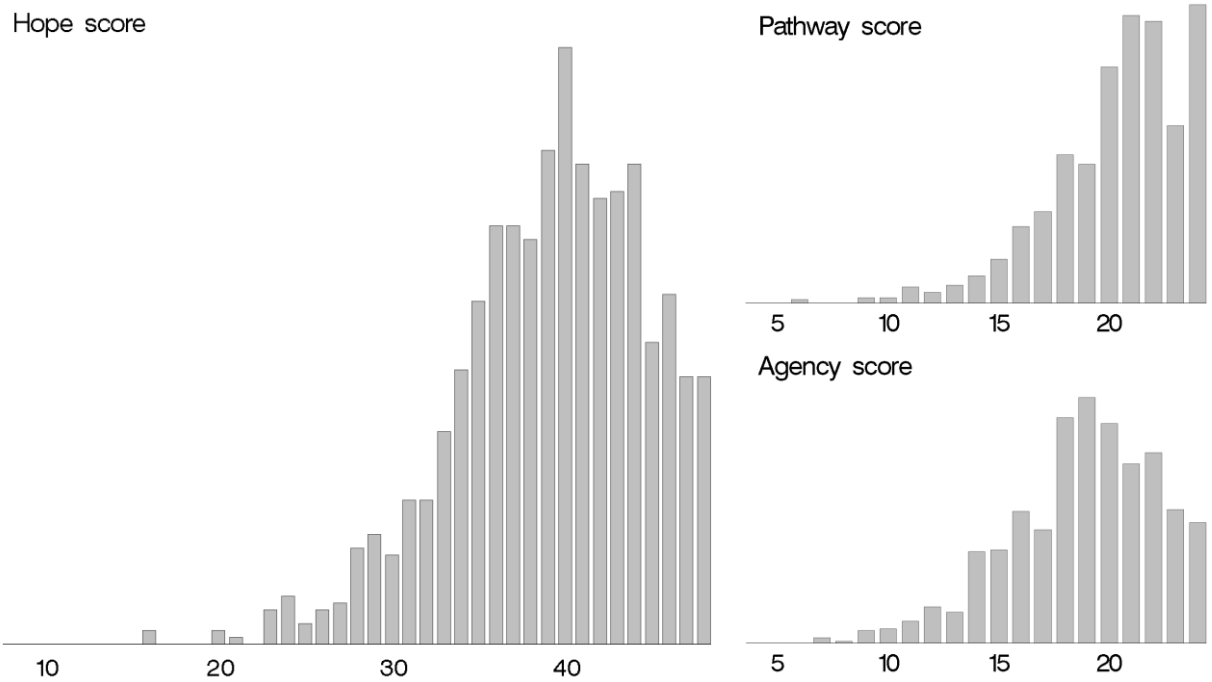
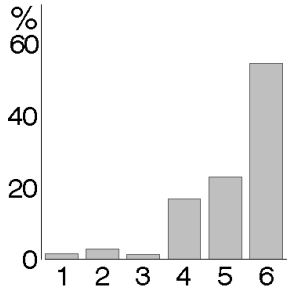
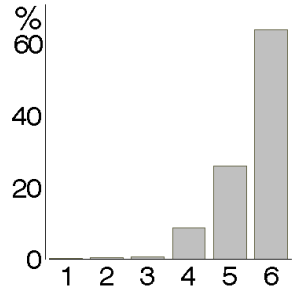


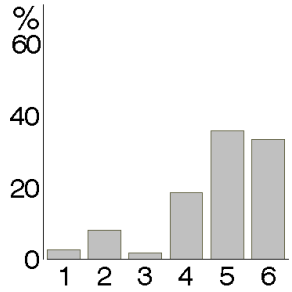
Figure 1 Distribution of hope total scores (left), and of the pathway (top right) and agency (bottom right) sub-scale scores



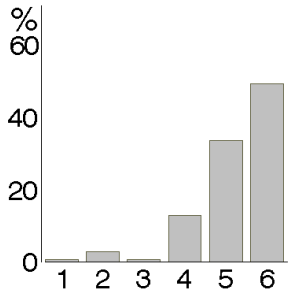
Item #1 You can think of many ways to get out of a difficult situation when life evolves unexpectedly



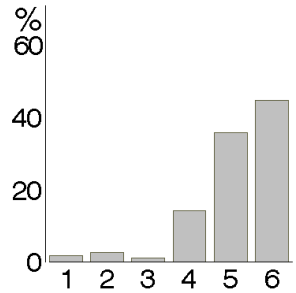
Item #2 You energetically pursue your goals



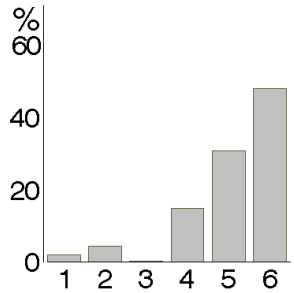
Item #3 There are lots of ways around any problem



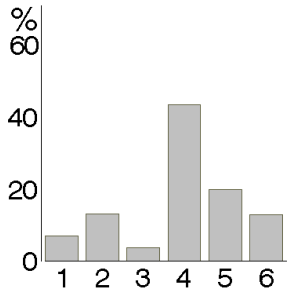
Item #4 You can think of many ways to get the things in life that are important to you



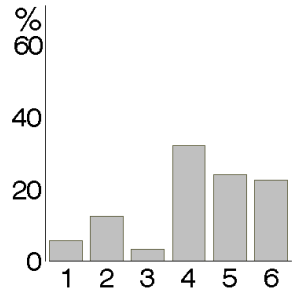
Item #5 Even when others get discouraged, you know you can find a way to solve the problem



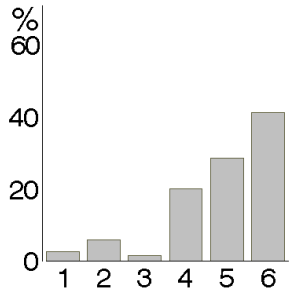
Item #6 Your past experiences have prepared you well for your future



Item #7 You've been pretty successful in life



Item #8 You meet the goals that you set for yourself



Overall n= 1021 x 8

Figure 2 Distribution of item responses shown for each of the eight scale items and overall. Participants indicated on a six-point scale how much they agreed with each statement. Answer options were (1) definitely false, (2) mostly false, (3) slightly false, (4) slightly true, (5) mostly true, (6) definitely true.

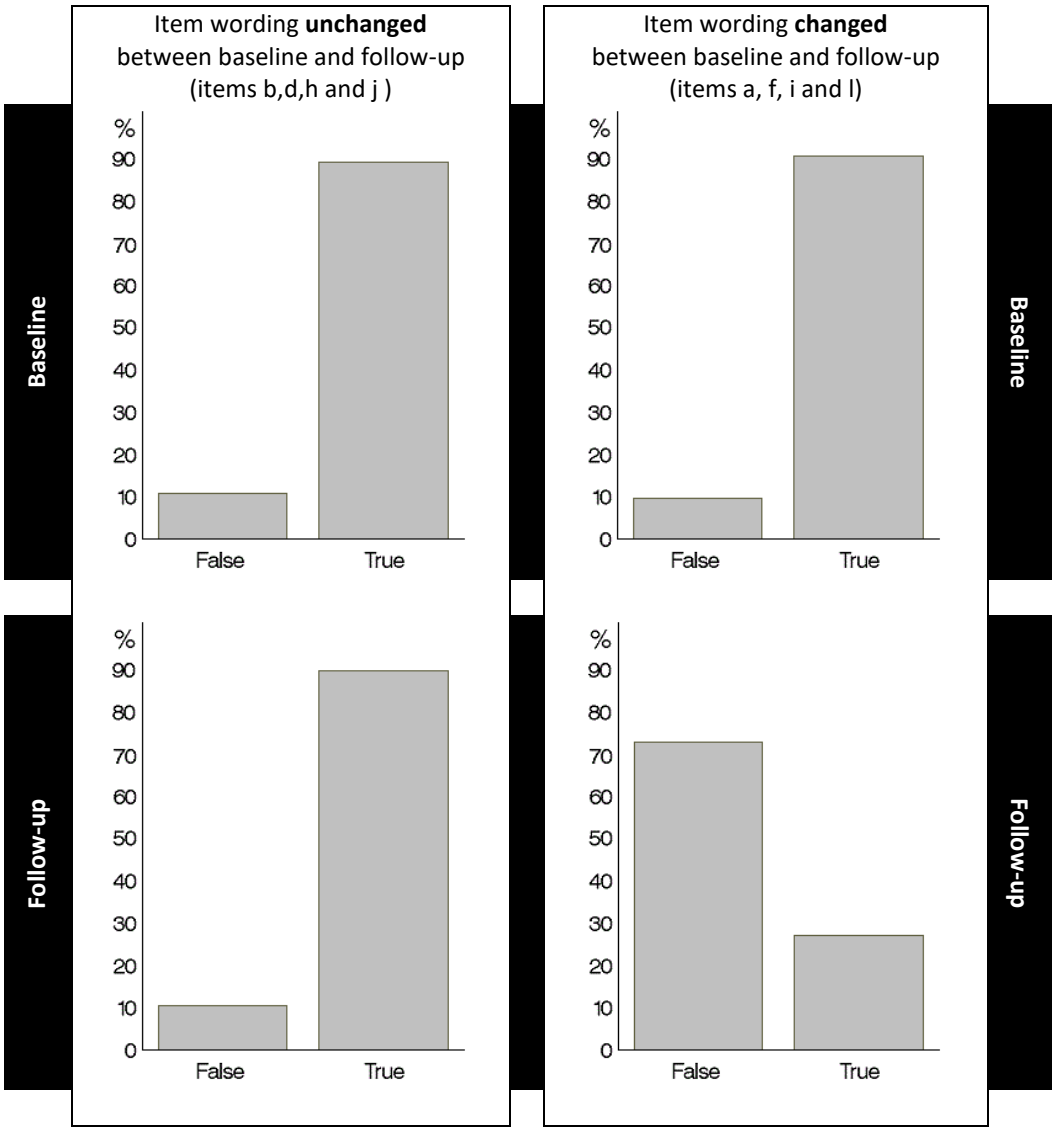


Figure 3 Overall proportion of responses expressing agreement ('slightly', 'mostly' or 'definitely true') with the statements presented at baseline (top) and follow-up (bottom) shown separately for the group of four statements (items b, d, h and j) that remained unchanged (left) and the four statements (items a, f, i, and l) that were negatively worded at follow-up (right).

Table 2 Snyder hope scores regressed on attributes that are in theory related to hope.

	n	mean (SD)	average (crude) effect on hope (95% CI)	p	R^2	effects from fully adjusted model[$] (95% CI)	p
Total	1021	39.2 (5.5)	-				
Education				**0.039**	**0.4%**		**0.097**
Secondary or higher	224	39.8 (4.8)	-			-	
None or primary only	797	39.0 (5.7)	-0.87 (-0.05; -1.69)			-0.68 (-1.49; 0.12)	
Wealth				**0.001**	**1.5%**		**0.008**
Owns a car	60	40.7 (4.0)	-			-	
Owns a TV (no car)	386	39.6 (5.2)	-1.14 (-2.64; 0.35)			-1.02 (-2.49; 0.45)	
Owns mobile phone only	429	39.0 (5.6)	-1.71 (-3.20; -0.22)			-1.36 (-2.82; 0.10)	
Owns none of the above	146	37.8 (6.4)	-2.87 (-4.52; -1.21)			-2.50 (-4.12;-0.88)	
Relationship violence[¥]				**0.110**	**0.3%**		**0.892**
No	750	39.3 (5.6)	-			-	
Yes	271	38.7 (5.2)	-0.63 (-1.40 to 0.14)			0.05 (-0.72; 0.83)	
Mental health				**<0.001**	**4.5%**		**<0.001**
Unit increment in poor mental health score[*]	1021	40.7 (5.4)[#]	-0.25 (-0.33; -0.18)			-0.24 (-0.32;-0.17)	

SD=standard deviation. CI=confidence interval. [$]Adjusted for all four variables shown in the table; R^2 (full model) = 5.9%. [¥]Reported suffering physical and/or sexual violence from an intimate partner in the past 12 months. [*]Poor mental health as measured on the SRQ-20 which ranges from 0 (perfect mental health) to 20 (worst possible mental health). [#]Modelled mean and SD shown for unexposed, i.e. for SRQ=0 in a simple linear regression of hope on SRQ.