

Assessment of Implementation Outcome Measures for Positive Deviance Approach as a New Strategy to Promote Exclusive Breastfeeding: A Psychometric Follow-Up Study

Yibeltal Siraneh ¹, Mirkuzie Woldie ^{1,2}, Zewdie Birhanu³

¹Department of Health Policy and Management, Faculty of Public Health, Jimma University, Jimma, Ethiopia; ²Fenot Project, Department of Population and Public Health, University of British Columbia, Addis Ababa, Ethiopia; ³Department of Health, Behavior and Society, Faculty of Public Health, Jimma University, Jimma, Ethiopia

Correspondence: Yibeltal Siraneh, Email yibeltal_siraneh@yahoo.com

Background: Implementation outcomes are the effects of deliberate actions while implementing new interventions which explain the process and are preconditions to achieve the desired outcomes. Positive Deviance Approach (PDA) was implemented as a new strategy to mitigate the gaps of exclusive breastfeeding (EBF) practice in Jimma, Ethiopia. However, there was no evidence to what extent intervention participants' perceive/rate these outcomes. In addition, psychometric outcome measures lack conceptual clarity, have unknown reliability and validity to the context. Therefore, we aimed to assess implementation outcomes, determine valid and reliable scales and analyze their correlation and variation explained.

Methods: A post-implementation follow-up study was conducted between August 28, 2020 and September 30, 2020 among the intervention participants of an earlier trial. A sample of 384 participants was invited to respond to an interviewer (or self)-administered structured questionnaire. The developed tool was refined through experts' comment, pre-test, and using Exploratory Factor Analysis (EFA) to determine the measurement scales, and Total Variability Explained (TVE). The reliability test was done. The mean scores were determined which indicates the extent of each scale's rate. Pearson's correlation, independent sample *t*-test/One-way ANOVA were used as needed. The explained variation (interdependency) of scales was checked using coefficient of determination (r^2). Level of significance was declared at 95% CI and *p*-value of <0.05.

Results: This study identified nine implementation outcomes of PDA as a new strategy to promote EBF with a TVE 72.1%. 52 valid and reliable items were developed to measure these outcomes. The maximum mean score was for "appropriateness scale" - 27.81 (6.5SD), while the minimum was for "implementation cost scale" - 11.37 (5.2SD). The overall mean score was 164.18 (26.8SD). The emerged scales explained 72.1% of the total variability in implementation outcomes. Majority (6) of outcomes of PDA were highly rated. Implementation fidelity, penetration, organizational readiness, and sustainability scores were positively and significantly correlated with acceptability of the approach.

Conclusion: This study identified nine valid, reliable, well explained and correlated implementation outcomes which help to understand the success of PDA as a new strategy to promote EBF. Majority of outcome measures were highly rated by intervention participants, which suggests it is a promising approach in promoting EBF in urban community by PDs.

Keywords: implementation outcomes, measurement scales, construct validity, reliability, correlation, variation explained, follow-up study, Jimma, Ethiopia

Background

Implementation outcomes are the effects of deliberate and purposive actions while implementing new interventions. These outcome measures are essential for evaluating whether the intervention is successful^{1,2} and they are the central meaning of what implementation science is.^{1,3} It helps to identify the health system challenges and solutions, and is

particularly useful in developing countries where the reachability of interventions is under-reported.⁴ These outcomes are crucial in implementation science and are indicators of the implementation processes and preconditions to achieve the desired outcomes.³ Hence, reliable and valid measures of these outcomes are important to understand the mechanisms of success or failure.

However, currently available implementation outcome measures are not clear, not context-based, and less qualified to our interest.⁵ This raises questions in building common knowledge in guiding implementation efforts. The available measures were not tested for validity, reliability and their correlation/variability. Due to this, they are not clearly distinguishable including measuring items that seem similar to examine different constructs.^{3,6} So what these generic measures are assessing is not exactly clear. In fact, some measures of implementation outcomes have been assessed for scale validity and reliability even if not generalizable for specific approach, context, or behavior.^{5,7} These measures are somehow conceptually distinguishable though it has been difficult to differentiate empirically, and to what extent intervention participants rate/perceive them is unpredictable.^{3,5}

In our study, we want to be clear with outcome measures of Positive Deviance Approach (PDA) intervention used to promote EBF practice in the community as stated in the following paragraphs. EBF is an effective intervention to the health of infants, and a maternal practice used to feed breast milk only, except supplements/medicines in the first 6 months of age.^{8,9} The available measures are like “empty cup” such as acceptability, appropriateness, feasibility, fidelity, penetration, sustainability, organizational readiness, adaptability and implementation cost. The reason why we selected those measures are a) they seem conceptually distinct but they are complex empirically,^{3,7} b) they are mostly used in formative studies as the main indicators of implementation success/failure but not studied,^{3,10} c) evidence shows those outcomes are highly correlated and proxies to each other but we fail to justify contextually, d) our preliminary intervention needs to address definitional and measurement ambiguities and we need to understand whether the means justifies the end outcomes of such intervention, e) to contribute to the development of valid, reliable and qualified outcome measures in implementation science,⁷ and f) to understand how and to what extent intervention participants (implementors and end-users) perceive them. This will also answer the questions of measurement perspectives of those outcomes as we can assess them flexibly. These measures do not have the same meaning for many researchers and rather are conceptualized and measured contextually. There is no agreed-upon type of measurement to develop for different intervention participants. Hence, developing, operationalizing, measuring, and validating the constructs are important in order to understand it.^{1,5,7}

Applying a new approach and measuring its implementation outcomes is very relevant. This is because assessing intermediate constructs is equally important as measuring effectiveness to explain the means. These outcomes are the key elements of the implementation process and they are proxy indicators of the service/client outcome, while distal indicators to the desired outcome.^{1,3,5,7} Those constructs can be measured at any implementation stage (pre, throughout, post-implementation).³ Studies are inconsistent regarding when to measure such outcomes. However, we have not yet identified empirical studies that explain which outcomes are best suited for different types of studies at different stages of implementation.^{3,7} These outcomes are themselves interrelated in dynamic and complex ways.^{11,12} Although we implemented a new strategy (PDA), we have no evidence on how to design the most acceptable and sustainable mechanisms in developing feasible and effective EBF promoting strategy.

There are key questions in evaluating implementation research such as how to conceptualize and measure success or failure of implementation processes and their impact on ultimate outcome. Implementation outcomes need to be developed, identified, and assessed, distinct from intended behavioral outcomes. The reason why it is important to differentiate these outcomes is to understand how services/approaches are successful and the reasons are equally important in understanding its effect.¹³ In implementation research, the challenge is not only applying a certain approach but also measuring its implementation outcomes at the same time. These outcomes may be differentially salient to various stakeholders depending on their perception.^{3,13} To ensure applicability of implementation outcomes across a range of settings and to maximize their external validity, all stakeholders should be represented.

This is because established evidence of a “proven” approach elsewhere does not ensure successful implementation in another context. Implementation also requires addressing a number of important contextual factors, such as provider attitudes, end user behavior, and the service system/approach. These outcomes have potential to capture the provider and users’ attitude

(acceptability) and behaviors (adoption, uptake) as well as contextual factors (system penetration, appropriateness, implementation cost).^{3,14} In this regard, we believe that successful implementation should be considered in light of many factors, including implementation outcome measures. Implementation success should be reflected by the implementation outcomes. Therefore, reliable and valid measures of implementation outcomes will enable empirical testing of the success of the efforts. In most current initiatives including our intervention, the success is assumed and evaluated at behavioral or health outcome level. However, beyond an exclusive focus on such outcomes only, understanding the process of implementation with contextual factors is paramount to have comprehensive understanding of how the approach works or not.

Theoretical Framework

Scholars defined acceptability, appropriateness, feasibility, fidelity, adaptability, organizational readiness, penetration, sustainability and implementation costs as follows: acceptability is the perception among intervention participants that a given intervention or approach is agreeable, while appropriateness is the perceived fit, relevance, or compatibility of the intervention at a given setting for the provider, or end-users; and/or its fitness to address specific problems. It is conceptually similar to acceptability and it seems overlapping.^{3,13} Feasibility is the extent to which a new intervention can be successfully implemented within a given setting,³ while implementation fidelity is to confirm whether it delivered as intended.¹⁵ Adaptation is the process of implementers or users bringing changes to the original design of an intervention. Adaptive interventions are those for which adaptation is allowed or even encouraged.¹⁶ Organizational readiness is a shared psychological state in which members feel committed to implement an intervention through receptive means.¹⁷ Penetration is the integration of a practice within a service setting and its subsystems.¹⁵ We found infrequent use of this term in the implementation research rather known in terms of institutionalization. Sustainability is the extent to which a new intervention is maintained within a setting.¹⁵ Implementation cost is the cost impact of an implementation effort that can be perceived, actual or opportunity cost.³

All the aforementioned constructs evaluate the fitness of something. Some measures are personal which means different people can view the same approach in different ways based on their needs, preferences/expectations, while some others are somehow technical or social. These constructs seem to overlap. However, they can be measured operationally to fit with the purpose. For example, a PDA could be judged both as unacceptable and inappropriate if PDs see it as contrary to the existing social norms related with the thing under promotion. In the same way, a PDA could be judged both as inappropriate, when considered as ineffective in meeting mothers' needs, and infeasible if PDs see it as difficult to implement at community setting with the existing resources, social networks and values. Therefore, perceptions of PDA could be measured with different conditions under observation. An implementation framework shows that¹⁸ acceptability is likely to vary to a greater extent than other constructs. This is because of willingness to try new things (trialability), while appropriateness is also highly variable based on the consistency with mothers' needs (relevance). Similarly, feasibility is also variable based on the cost or time (resource availability) or ease of implementation or use (complexity). From this understanding, we assumed as all those outcome measures are correlated.^{14,18}

The PDA was implemented as a new strategy to promote EBF which assumed in any community there are people whose uncommon but successful behaviors enable them to find better solutions to a problem than their peers.^{19,20} However, we were not clear about measurements and correlation of implementation outcomes of this approach. The successful implementation of the PDA depends on these outcome measures as perceived/rated by the study participants. However, there is no evidence to what extent they perceive/rate those outcomes as far as EBF is concerned. In addition, currently available measures lack conceptual clarity and have largely unknown reliability and validity to the context. Hence, assessing outcome measures helps to understand the mechanism how intervention participants perceived and engaged in promoting EBF practice at community-level using a PDA. Therefore, we assessed implementation outcomes, determined valid and reliable measurement scales and analyzed their correlation and variation explained.

Methods and Participants

Study Design and Setting

A post-implementation follow-up study was conducted in Jimma town, Ethiopia between August 28, 2020 and September 30, 2020 among the intervention participants of an earlier trial.²¹ Jimma town is one of the twelve Oromia town administrations which is found 352 Kilometers away from Addis Ababa to the southwest part of the country. It has 17 kebeles (smallest administrative units) all of which are working sites for urban Health Extension Professionals (HEPs). Total population was 192,000, and 43,955 households. There were a total of 59 HEPs assigned and working for community-based health services in the town. In the intervention arm of the trial, there were 8 HEPs working within 3 clusters/Kebeles.²²

In the earlier study, the intervention arm of the trial integrated (supported) with the implementation activities. This was done because one of the implementation study designs is a cluster Randomized Controlled Trial (RCT) at the community setting. In a community setting, RCT study has many drawbacks that could be minimized by integrating implementation research with participatory approach.^{1,2} The trial was conducted to compare the change in the primary outcome (EBF proportion) between the two arms. However, in testing the newly approached intervention, implementation study could identify how the intervention works or not in a given context in addition to the effectiveness study.^{1,2} The study area/clusters and population that was managed during the earlier cluster RCT study were clearly defined. Six clusters (Kebeles) were randomly selected. Three of the clusters were randomized as an intervention area (Mendera-Kochi, Ginjo-Gudiru and Awetu-mendera) and the remaining as control areas.²²

Description of the Approach and Implementation Activities

A PDA was followed to promote EBF practice at the community-level through home-to-home visit. Different stakeholders (as stated in the study population section) including the end-users actively participated in developing implementation activities. The positive deviance (PD) inquiry and approach were defined and contextualized following six principal steps. Why and how the integration was made, implementation steps, and process were explained and discussed with stakeholders. The home-based informational counseling and social support activities were implemented for 7 months in 2018. It was implemented in 3 clusters with six visits per mother. Positive deviants were incentivized with minimum acceptable refreshment cost permitted by ethical committee (0.5US\$/visit) which means 5US\$ when she finished 10 home visits at each round (total rounds-6). The informational counseling and social support activities were assumed to take 45 minutes per visit while those in the control arm received the usual service from urban HEPs. The main implementation activities were;

- Enrolling eligible mothers
 - This was conducted by the project team and urban HEPs
 - Facility and community based surveys were conducted
 - The main activity was listing participants from family folder (Health extension registry book) with identifiers and completing rapid community survey not to miss those non-registered.
- Capacity building training for Health Development Army (HDA) leaders and HEPs on PDA
 - This was conducted by the project team
 - It was arranged at urban HEPs working sites/Kebeles
 - All HEPs and HDA leaders received formal communication and training on how to proceed with PDA principles and the way forward for this intervention study.
- Identifying the presence of deviated behaviors (in favoring EBF practice), selecting positive deviants and training them
 - This was conducted by the project team, Women HDA (WHDA) leaders and HEPs
 - It was facility-based and arranged at urban HEPs working sites/Kebeles
 - Through qualitative methods and together with HEPs and WHDA leaders, PDs were recruited/screened from the community and qualified by training

- The training modules focused on informational support (Module one) and emotional and appraisal support (Module two).
- Provision of personalized and comprehensive informational counseling and social support
 - This was conducted by the trained PDs as per the protocol for intervention groups only
 - It was a community-based home to home visit
 - Informational counseling and social support to mothers together with their family or relevant others on EBF practice was performed. The time of visit, activities per visit, and the total number of visits were also clearly stated in the protocol and performed accordingly
 - All the added support was provided to the participants in the intervention arm. Selected and trained PDs were the agents for each visit in addition to the usual routine community-based services. The intervention was also considered the participation of their relevant others. The counseling contents of EBF were adapted from WHO guideline²³ and used to train PDs. At each of the six visits, the psychosocial support (emotional and appraisal) and verification of their understanding was the “pillar” to make sure about the plan of care, experience sharing and expectations, beliefs, and myths related to EBF. The PDs followed the five key points during the visits as appropriate;
 - Developing an empathic relationship
 - Collaborating with the family
 - Gently probing the individual’s and family’s BF beliefs, attitudes and stimulating alternative ideas
 - Practical support to the mother (skill transfer)
 - Appraisal of any difficulty faced during BF was done.

The details of the study design, implementation process, main activities, visiting time and intervention package were reported in an earlier published article.²¹

Study Population

This post-implementation study was conducted among intervention arm of the earlier trial. Among the six clusters (kebeles) which were randomly selected of the total, three of the clusters were randomized into the intervention arm. The total number of mothers in the intervention and control arm were 260 (130 each). During enrollment, the name of Kebele and villages, personal contacts/phone number, and house number of each participant were documented to be accessed again for this follow-up study. All mothers who received the intervention nominated one relevant other who supported her in practicing EBF and engaged during the counseling and social support sessions. In addition to this, All HEPs, health extension supervisors, Community Health Workers (CHWs) (those are community volunteers who are trained on basic health promotion and/or specific intervention to facilitate new interventions, mass drug administration, sanitation campaign, immunization campaign, and any other community level health related activities with close supervision from HEP and who plan and work together with HEP), WHDA leaders/volunteers and model mothers certified by HEP and all positive deviants were considered as a study population in this follow-up study. They all participated in the implementation process/activities as an implementer and/or facilitator or counselor or, end-user or combination of these. All mothers who received the intervention and their relevant others were included. The remaining study participants (HEPs, health extension program supervisors, CHWs, WHDA leaders, model mothers who were volunteers, and PDs) were included based on their active participation during the implementation process. The implementers and/or facilitators received basic training on PDA, steps of deviance inquiry, how to set selection criteria to identify deviants, selection of deviant mothers, basic skills of facilitation to promote EBF using this approach at community setting.

Sample Size and Sampling Procedure

The sample size was estimated considering all active participants/stakeholders who were involved during implementation. Those were 8 HEPs, 10 CHWs, 45 WHDA leaders, 45 model mothers, 3 health services extension program supervisors, and 13 positive deviants who participated in the inception training, review meeting and workshops and in the

subsequent implementation and intervention process, and 130 mothers who received the intervention with their one relevant other. Therefore, the total sample size was the submission of $8+10+45+45+3+13+130+130=384$ intervention participants. Hence, the final sample size was 384. All these eligible study participants were invited to participate in this study.

Study Variables and Measurements

The study variables were socio-demographic characteristics of the study participants (age, sex, educational status, ethnicity, religion, marital status, occupational status, monthly income), participation in the intervention (role in the implementation, level of participation during implementation, self-confidence to run the program, and value to the approach, perceived level of understanding of PDA and perceived context of support), and implementation outcome measures (acceptability, appropriateness, feasibility, fidelity, adaptability, organizational/structural readiness, penetration, sustainability and implementation costs). The first two parts of variables were measured using nominal and ordered responses, and some variables such as age and monthly income in continuous manner. The remaining variables (outcome measures) are generic constructs conceptualized from literature^{1-3,5,7,13,14} and identified to consider while measuring certain implementation outcomes. In this study, crude numbers of items that were assumed to measure these constructs in context were developed for validation and to identify reliable measurement scales of the intervention outcome of the PDA. We employed a deductive approach to generate those items,²⁴ whereby we used the definitions and conceptual framework described in the background section to ascertain whether items could adequately measure/capture the theoretical content/concept of the construct in relation to the PDA. In the process of item and construct validation analysis, it was determined whether items represented the intended construct more than the other constructs through removing the cross-loading items. All outcome measures were measured using a five-point Likert agreement scale. The scale ranged from 1 (strongly disagree) to 5 (strongly agree). The total number of initially developed items was 78. Responses to negatively worded statements were reverse-scored before analysis. After identifying valid constructs, the scores of each respondent were summed up and the highest score indicates higher relevance of the factor in promoting the approach with the perception of the study participants. The relative weight of constructs was identified using PCA/factor analysis. The pre-assumed constructs were operationalized and measured as follows;

Acceptability is the perception among implementation stakeholders and end users that a PDA is agreeable or appealing to them. It was assessed based on the study participants' experience with various dimensions of the intervention implemented, and users' survey which is more specific, referencing a particular set of intervention processes. This construct directly explored the study participants' perceptions that the PDA is beneficial to them and to others in terms of preferences, needs, compatibility, suitability, usefulness, practicability regarding the EBF situation. Ten¹⁰ items were developed to capture their perceptions about the approach in addressing BF problem in context. Responses were added up to form a composite score that was interpreted as a higher value of the score indicating the acceptability of the PDA and relative importance of this factor. The higher acceptability of the PDA implied that it is a promising strategy in promoting EBF practice in urban settings by positive deviants.

Appropriateness is the relevance, or compatibility of the PDA to the intervention participants, provider, or consumer; and/or perceived fit of the intervention to address EBF practice in the setting. It is perception based on respondents' view as a PDA intervention may be perceived as appropriate but they might not accept it, and vice versa. This construct is important to capture some "pushback" to implementation efforts, or if it is not consistent with PDs' skill set, role, or job expectations. Hence, a variety of stakeholders would likely have perceptions about a new intervention's appropriateness to a particular service setting, objective, providers, and clients. Seven⁷ items were developed and designed to capture their perceptions about the perceived fit of the approach in addressing BF problems. Responses were summed-up to get a composite score that was interpreted as a higher value of the score indicating the appropriateness of the PDA. High composite scores implied a higher importance of this factor. A higher score implied that it is an appropriate strategy in promoting EBF practice in urban settings by positive deviants.

Feasibility is the extent to which a PDA (new/intervention) can be successfully used or carried out within urban community setting to promote EBF that could be measured by the actual and perceived experience of implementation stakeholders. Typically, the concept of feasibility is the potential to explain the PDA intervention's success or failure, as

reflected by recruitment, retention, or participation rates. Study participants could perceive that the PDA may be appropriate for their setting — in that it is compatible with the setting's mission or service mandate, but may not be feasible due to resource or training requirements. Eight⁸ items were developed and designed to capture their perceptions about the extent to which level it was carried out using the available resources. Composite score was analyzed to interpret a higher value of the score showing appropriateness of the approach and relative importance of this factor. A higher score implied that it is a feasible strategy in promoting EBF practice by positive deviants.

Fidelity is to confirm whether implementation was delivered as per the plan and quality of implementation used to ensure internal validity. Implementation stakeholders could have a perception whether the intervention maintained its intended effects. It has a direct relation with outcome of an intervention as a mediating variable between the context and intervention effectiveness. They can have their own judgement whether the PDA intervention was implemented as it was prescribed in the original protocol by comparing the original evidence with the actually implemented activities in terms of six dimensions such as adherence to the program protocol, dose or amount of program delivered, quality of program delivery, program component differentiation, exposure to the intervention, and participant responsiveness or involvement. This perceived fidelity (individuals' perceptions about the degree to which an intervention was delivered as intended) of implementation was measured through self-report/ratings, using 10 items on a five-point Likert agreement scale, by individuals delivering the services or receiving them, and what was achieved as compared to the shared plan at the beginning. Items were designed based on the major PDA intervention components and activities. Composite score was analyzed and interpreted as higher score indicating the intervention was conducted as intended in a valid way.

Adoption/adaptability is the degree of uptake or actual use of a program at different levels. It is the intention, initial decision, or action of the end-users to try or employ the EBF practice promoted by PDs as facilitated by many other stakeholders. Adaptation is the process of implementers or users adopting the practice and the way it is promoted as designed or making it fit them without compromising its intent. Depending on the nature of the modifications, adaptation could either be potentially positive or could carry the risk of threatening the theoretical basis of the intervention, resulting in a negative effect on expected outcomes. Adaptive interventions are those for which adaptation is allowed or even encouraged as we promoted in this PDA intervention. In this study, adaptations could be deliberate or accidental and were measured in response to the 5 developed items. Respondents were expected to consider whether no modification was made, need or no need for modification, additions of new components, deletions/ignoring of an intervention component and minor or major modifications to an existing intervention component from their implementation experiences. Composite was created with higher scores indicating high degree of adoptions or uptake of the PDA interventions for promoting EBF by the PDs.

Organizational readiness is reflected in terms of features such as culture, climate, structures, policies, routines, and resources can create a receptive context for the new intervention or approach. However, receptive context does not translate directly into organizational readiness, rather the content and approach of intervention matters as of the context of implementation. Project stakeholders might work in a very receptive organizational context, but they might not value all proposed changes or new approach equally. It is a shared psychological state in which stakeholders feel committed to implement the intervention and confident in their collective abilities to do so and how much they value the effort and how favorably they appraise factors of implementation capability. Hence, this construct was measured through stakeholders' perceived readiness for the task demands, resource availability, and situational factors, readiness for change, commitment to sustain the change, and change efficacy. Seven⁷ items were designed to measure this component. Responses were composited with a high value indicating high perceptions of organizational/community readiness. When organizational readiness is high, stakeholders and the community are more likely to promote EBF using PDs, exert maximum effort, show more cooperative behavior, and feel confident that they can do it.

Penetration is defined as the integration of PDA intervention activities within a setting and its subsystems. The concept of intervention penetration to the users (the number of eligible persons who use a service, the number of providers who deliver a given service and the number of facilitators of the implementation as compared to the expected ones) was measured to assess feelings of intervention participants regarding the degree of reachability of the intervention. The project reported 130 mothers got the intervention and 13 trained PDs did what was expected from them as per the protocol. However, the intervention stakeholders could perceive different that was assessed using 7 developed items. The

created composite score was interpreted as higher value indicating highly successful/penetrating intervention was conducted using the PDA in addressing EBF practice.

Sustainability is the extent to which a newly implemented PDA intervention is integrated or institutionalized within a service setting's ongoing and stable operations to transit from temporary to permanent support, from periodic to routine practice and attaining long-term viability. Implementation stakeholders might perceive sustainability as different from penetration, in that higher penetration may contribute to long-term sustainability. This study measured the stakeholders' perception about how likely the PDA intervention in promoting EBF will be continued and functional in the community. There were a total of 8 items designed to measure this dimension. Intervention participants were asked their opinion/to rate across a range of factors that were related with the sustainability of this program. High composite scores were interpreted as a higher importance of this factor. This means; stakeholders who have higher perceptions toward the viability and ability to implement the PDA to promote EBF will result in future sustainability.

Implementation cost is the cost impact of an implementation effort that is incurred in addition to the usual baseline cost. Implementation costs vary according to interventions' complexity, way of delivery, interests, settings and overheads. The direct and indirect expense/cost of implementing an intervention was measured and analyzed in another cost-effectiveness study. However, in this study, the perception of stakeholders' about the measurable, non-measurable and opportunity cost of the approach/intervention was assessed using 8 items. The responses to the positively stated items were reverse-scored before the analysis. A higher composite score indicates that the implementation of PDA in addressing the problem of EBF practice was perceived as costly.

Tool Development and Validation

Data were collected using a structured questionnaire developed to address the socio-demographic and other intervention related variables, and implementation outcome measures. Initially the questionnaires were prepared in English. To ensure face validity, the draft questionnaire was reviewed by an expert panel including 3 MPH students, 2 Ph.D. students, and 2 academic staff with MPH. They were given 78 items developed for measuring implementation outcomes to review and comment on items' understandability, logical order, duplication of items with the same meaning, readability, relevance, simplicity, language clarity such as wording and sentence structures. Based on their comments, relevant modification was made, and 8 items were dropped. The final modified tool was translated into Amharic and Afan Oromo by language experts, and then back-translated into English by a third party who had experience to ensure meaning equivalence. Then, based on the comments from the pretest (conducted on 5% (n=19) of the total sample size), relevant modifications were made. Next, four health science graduates in the nursing and public health profession were recruited for data collection and supervision. Three data collectors and one supervisor were trained on the procedure of data collection, the content of the pre-tested tool and ethical considerations. Trained interviewers collected the data from each intervention participant through self-administered or interviewer-administered ways considering the participant's ability to read and write. The data collection was conducted at the study participant's home or office. A noise-free area for the interview was selected considering each context. The interview was estimated to take 30 to 45 minutes. Data were checked for completeness after each day of data collection by supervisors.

Using the collected data, the construct validity and items reliability test were checked for the implementation outcome measures. Construct validity was tested using Exploratory Factor Analysis (EFA) to indicate valid scale constructs with a total variability explained. Principal Component Analysis (PCA) method was used. Then reliability test (inter-item consistency) was checked using Cronbach's alpha values for each measurement scale.

Data Analysis

The data were entered into Epidata 3.1 and analyzed using the statistical packages for social sciences (SPSS) software version 21.0. Descriptive statistical measures (socio-demographic characteristics of the participants) such as frequency, mean, proportion, and standard deviation were computed and presented.

To assess the validity of the conceptual measurements, we performed EFA on the same data. Factor analysis using PCA method was conducted to reduce the data/items and to test construct validity. The purpose of doing EFA was to explore and validate constructs and respective items, reduce the data, and create composite factor scores. All the

assumptions of principal component analysis (PCA) were checked. All (70) items were fitted once to create components. The Eigenvalue was fixed as 9 (number of pre-assumed constructs). The case to variable ratio was 5.3 to 1. Bartlett's Test of Sphericity was significant at $p < 0.05$. Overall sampling adequacy was checked using Kaiser-Meyer-Olkin (KMO) which was > 0.5 . The correlation matrix showed more than 2 cells with a coefficient of 0.3 to 0.9. Varimax rotation was employed during factor extraction to minimize the cross-loading of items. Cross-loaded items (items created complex structure) on more than one component were removed. Reliability of items (inter-item consistency) was checked using the reliability coefficient (Cronbach's alpha values) > 0.70 .

The measure of correlation and variations among the implementation outcome measuring scales were analyzed. Pearson's correlation analysis was carried out to examine the relationship between the psychometrically measured implementation outcome variables as bivariate analysis. Similarly, an independent sample *t*-test was carried out to compare the mean score between the dichotomous variables (two groups). One-way ANOVA was also used to see the significant differences between more than two groups of the selected background characteristics. All assumptions for each test were checked. Multi-collinearity between covariates was assessed with individual correlation values. The explained variation (interdependency) of scales was checked using coefficient of determination (r^2). All statistical analyses were performed at the 95% confidence interval with a 5% level of significance. Accordingly, correlated variables and compared and significantly different mean scores were interpreted.

Results

Background Characteristics

The response rate was 96.9% (372/384). The mean age was 30.7 (± 6.9 SD) ranging from 18 to 60 years. The majority (66.1%) were between the age of 25–34 years. More than three-fourths (80.9%) of the study participants were females, and 83.6% were married. The majority (62.6%) could read and write through formal education. More than one third (34.9%) were merchants, and the majority (60.2%) had estimated monthly net income of 33.1US\$ or less (during data collection period, the average exchange rate was 1US\$=36.50Birr). The majority (68.3%) of study participants were end-users who were mothers and their relevant others. More than half (54.3%) reported they were highly or actively engaged during implementation. Similarly, 54.6% and 54% of them reported they were highly confident, and competent enough to promote EBF through PDA respectively (Table 1).

Description of Implementation Outcome Measurement Scales

Implementation outcome measurement scales were identified using PCA/EFA. Each emerged latent variable was named considering the common concept of items loaded to it. Fifty-two items having a communality of > 0.50 were retained on nine components/scales. These were named as acceptability, fidelity, appropriateness, feasibility, penetration, sustainability, adaptability, organizational readiness, and implementation cost with 7, 7, 7, 6, 6, 4, 5, 5 and 5 items retained respectively. Sixteen¹⁶ items were removed due to a commonality value < 0.5 . The other two items were removed due to cross-loading value of > 0.4 . An item with a highest mean score indicates that the item was highly rated in that scale. In this study, highly rated item with a mean score of 4.09 ± 1 SD was loaded to "acceptability". On the contrary, an item with the lowest mean score (2.09 ± 1.1 SD) was loaded to "feasibility" scale/factor. The Cronbach's alpha reliability coefficients for these items loaded to appropriateness, acceptability, fidelity, adaptability, penetration, implementation cost, feasibility, organizational readiness, and sustainability were 0.964, 0.906, 0.920, 0.952, 0.887, 0.923, 0.840, 0.851, and 0.911 respectively (Table 2).

Summary Measures for Implementation Outcomes

Implementation outcomes of the PDA as an intervention to promote EBF were summarized using the scale mean scores and the Total Variability Explained (TVE) by the emerged components. The higher factor score created on a factor analysis indicates a higher degree of importance of that factor with regard to implementing PDA to promote EBF. The minimum mean score belonged to implementation cost scale - 11.37 (5.2SD), while the maximum mean score was for appropriateness scale - 27.81 (6.5SD). The highest variance score (Var: 42.3) was also recorded for appropriateness scale

Table 1 Background Characteristics of the Participants, a Follow-Up Study, Jimma Town, 2020

Variables (n=372)	Category	N (%)
Age (in years)	15–24 years	45(12.1)
	25–34 years	246(66.1)
	≥35 years	81(21.8)
Sex of the respondent	Female	301(80.9)
	Male	71(19.1)
Religion	Orthodox	147(39.5)
	Muslim	136(36.6)
	Protestant	78 (21)
	Others	11(2.9)
Ethnic group	Oromo	162(43.5)
	Amhara	122(32.8)
	Dawuro	44(11.8)
	Kaffa	33(8.9)
	Other	11 (3)
Marital status	Married	311(83.6)
	Divorced	26 (7)
	Widowed	17(4.6)
	Single	18(4.8)
Educational status	Cannot read and write at all	101(27.2)
	Can read and write without formal education	38(10.2)
	Can read and write through formal education	233(62.6)
Occupational status	Housewife	76(20.4)
	Merchant	130(34.9)
	Gov't employee	87(23.4)
	Private employee*	62(16.7)
	Other	17(4.6)
Estimated monthly net income	≤1000Birr	224(60.2)
	1001–2000Birr	77(20.7)
	≥2001Birr	71(19.1)
Role in the PDA implementation	Facilitator and/or trainee	105(28.2)
	Counselors/change agent	13(3.5)
	End user	254(68.3)

(Continued)

Table 1 (Continued).

Variables (n=372)	Category	N (%)
Perceived level of engagement during PDA intervention	Highly/actively engaged	202(54.3)
	Medium	128(34.4)
	Low engagement	42(11.3)
Perceived level of confidence to promote EBF through PDA	Highly confident	203(54.6)
	Medium	58(15.6)
	Low confidence	92(24.7)
	Uncertain	19(5.1)
Perceived level of competence to implement or run the PDA	Highly competent	201(54)
	Medium	59(15.9)
	Low competence	81(21.8)
	Uncertain	31(8.3)
Importance of relevant others to promote/practice EBF	Yes, important	294(79)
	Not	48(12.9)
	I am not sure	30(8.1)

Notes: *Employed at private and non-governmental organizations.

Table 2 Descriptive Parameters (Emerged Scales) for the Measures of Implementation Outcomes of PDA as an Intervention to Promote EBF, Jimma, 2020

Measurement Scales	Items	Communalities	Loading	Mean (SD)	α
Appropriateness (Factor 1)	This PDA seems proper /applicable to promote EBF	0.952	0.971	3.97(\pm 0.9)	0.964
	The PDA with informational counseling and social support service seems right at home level	0.952	0.971	3.97(\pm 0.9)	
	The criteria used to select PDs are reasonable to work with PDA	0.952	0.971	3.97(\pm 0.9)	
	The PD approach was fit for the purpose in promoting EBF	0.952	0.971	3.97(\pm 0.9)	
	The PDA with informational counseling and social support service was useful/relevant	0.847	0.909	4.01(\pm 1.0)	
	This PDA seems well aligned/a good match with the objective	0.847	0.909	4.01(\pm 1.0)	
	PDA seems suitable/compatible with the providers and client need	0.547	0.665	3.93(\pm 1.2)	
Acceptability (Factor 2)	Promoting EBF practice using the PDA is compatible with mother's interest	0.728	0.819	3.86(\pm 1.1)	0.906
	This PDA is appealing/interesting	0.733	0.809	4.09 (\pm 1.0)	
	The PDA is useful in promoting EBF	0.648	0.783	3.79(\pm 1.3)	
	PDA is pretty good and has skillful promoters of EBF	0.701	0.767	4.06(\pm 1.1)	
	This PDA is okay to promote EBF	0.660	0.746	4.02(\pm 1.2)	
	Promoting EBF practice using the PDA is suitable/fine	0.615	0.732	3.91(\pm 1.0)	
	Promoting EBF through the PDA seems good enough (palatable)	0.545	0.679	3.75(\pm 1.2)	

(Continued)

Table 2 (Continued).

Measurement Scales	Items	Communalities	Loading	Mean (SD)	α
Fidelity (Factor 3)	There was information redundancy at different visits [®]	0.765	0.808	3.93(±1.1)	0.920
	Program participants complied to the program protocol	0.680	0.783	3.82(±1.1)	
	Dose or number of home visits for counseling was enough to initiate and sustain EBF practice	0.763	0.776	4.06(±1.1)	
	Emotional responsiveness (ability to respond empathically) during informational counseling was well considered as hallmark of this approach	0.717	0.765	3.79(±1.0)	
	Collecting signature of PD by the end user/mother was well performed technique to follow performance	0.735	0.751	3.88(±1.0)	
	The PD approach would maintain its intended effects	0.643	0.724	3.80(±1.1)	
	The intervention was implemented as it was prescribed in the original protocol	0.598	0.706	3.95(±0.9)	
Adaptability (Factor 4)	Mother involved in the process can be a counselor in the future	0.920	0.889	2.35(±1.1)	0.952
	Intention to practice EBF would be a fertile ground to inherit the PDA	0.920	0.889	2.35(±1.1)	
	The PD approach needs additions of new components/ services [®]	0.914	0.881	2.32(±1.1)	
	The PD approach needs major modification to promote EBF [®]	0.783	0.797	2.49(±1.2)	
	The PDA would improve the uptake of EBF information	0.709	0.741	2.62(±1.2)	
Penetration (Factor 5)	The approach may not be successful through home visit [®]	0.728	0.797	2.84(±1.2)	0.887
	The approach may not fit with the existing community structure [®]	0.701	0.777	2.85(±1.3)	
	Full engagement of participants is done to ensure reachability of the intervention	0.686	0.769	2.75(±1.3)	
	The spillover effect will increase the intervention's coverage in the community	0.639	0.729	2.75(±1.3)	
	Every eligible mother could be reached with such approach in urban setting	0.598	0.719	3.05(±1.2)	
	The approach is well integrated with the existing practice of urban HEP.	0.598	0.686	3.01(±1.3)	
Implementation Cost (Factor 6)	Implementing this approach is less costly in terms of cash	0.892	0.867	2.12(±1.2)	0.923
	The cost (time and/or cash) that is spent is more than the benefit received from this intervention [®]	0.892	0.867	2.12(±1.2)	
	The cost of implementing this approach is less as compared to the social/health value to be promoted	0.820	0.841	2.26(±1.2)	
	Implementing this approach is less costly in terms of time	0.685	0.731	2.37(±1.1)	
	Refreshment cost/incentive is enough to refresh during intervention activities	0.641	0.713	2.49(±1.2)	

(Continued)

Table 2 (Continued).

Measurement Scales	Items	Communalities	Loading	Mean (SD)	α
Feasibility (Factor 7)	This PDA seems workable/implementable	0.646	0.738	2.09(±1.1)	0.840
	This PDA seems realistic/doable	0.612	0.728	2.20(±1.2)	
	The PDA seems practical at urban setting/Jimma Town	0.603	0.724	2.16(±1.2)	
	Number of visits is adequate but needs to re-adjust its date [®]	0.581	0.713	2.34(±1.1)	
	This PDA seems easy to promote EBF	0.548	0.670	2.47(±1.2)	
	The PDA seems viable/possible to promote EBF	0.511	0.593	2.47(±1.2)	
Organizational Readiness (Factor 8)	Intervention participants (we) feel committed to implement PDA and confident in collective abilities	0.748	0.808	2.92(±1.3)	0.851
	We as program implementers/participants value the social support given to mothers	0.745	0.805	2.85(±1.3)	
	Resource availability matters as part of readiness for this new approach	0.618	0.679	2.95(±1.3)	
	There is cooperation among implementers in operating the new approach and readiness for change	0.547	0.677	3.35(±1.2)	
	There is supportive community structure to use PDA	0.584	0.651	2.78(±1.2)	
Sustainability (Factor 9)	PDA is institutionalized within a community structure	0.777	0.776	2.97(±1.3)	0.911
	PDA is maintained within a community	0.759	0.776	3.06(±1.3)	
	There is full community/ stakeholders support enabling the program	0.746	0.761	3.13(±1.3)	
	Volunteerism is a central aspect of this approach that would support maintenance of the intervention	0.749	0.760	3.18(±1.3)	

Note: [®] denotes the response for that item was reversed before analysis.

Abbreviations: SD, standard deviation; α , Reliability coefficient.

which indicated that the response variance (change) in this scale was high. The overall mean score of the nine outcome measures was 164.18(26.8SD). The study participants highly rated appropriateness scale which was followed by acceptability scale. On the contrary, the low rated scale was implementation cost. These highest and lowest mean scores imply the intervention participants hold a perception that the PDA to promote EBF is more likely appropriate, while costly to implement. The nine components/scales explained 72.1% of the total variability in psychometrically measured implementation outcomes of PDA. This is greater than the recommended cut-off point - 60%. This means the implementation outcomes of the approach were well explained or measured by the nine emerged scales. Appropriateness, acceptability, fidelity, penetration, organizational readiness and sustainability scales were rated above the expected potential mean score, while the mean score for feasibility, adaptability, and implementation cost scales were rated lower. This means, the PDA was perceived as appropriate, acceptable, fidable, penetrable, organizationally/structurally ready, and sustainable to promote EBF, while not feasible, adaptable, and costly (Table 3).

Inter-Scale Correlation (Relationship Between Outcome Measures)

Multiple correlation tests were done between the measurement scales of implementation outcomes of PDA as an intervention to promote EBF practice. The findings from Pearson's correlation coefficients (r) showed that the majority of the psychometric measurement scales were significantly correlated with each other either positively or

Table 3 Summary Measures for Implementation Outcomes of PDA as an Intervention Used to Promote EBF, Jimma Town, 2020

S. No	Measurement Scale (N=372)	Mean	Std. Deviation	Variance	Total Variance Explained (TVE)
1.	Appropriateness	27.81	6.502	42.279	72.1%
2.	Acceptability	27.47	6.278	39.414	
3.	Fidelity	27.23	6.011	36.132	
4.	Penetration	17.26	5.910	34.924	
5.	Organizational readiness	14.85	4.907	24.074	
6.	Feasibility	13.73	5.248	27.536	
7.	Sustainability	12.33	4.403	38.824	
8.	Adaptability/adoption	12.14	5.132	26.338	
9.	Implementation cost	11.37	5.155	26.574	
Composite score (minimum score: 66 and maximum score:260)		164.18	26.817	719.171	

negatively with different levels of strength. Implementation fidelity ($r=0.600$, $p<0.01$), penetration ($r=0.132$, $p<0.05$), organizational/community readiness ($r=0.195$, $p<0.01$), and sustainability ($r=0.265$, $p<0.01$) scores were positively and significantly correlated with acceptability of the approach. Of these, acceptability score was strongly correlated with the implementation fidelity score. Except for appropriateness and implementation cost; all the remaining scales (feasibility, penetration, sustainability, adaptability, and organizational readiness) were significantly and positively correlated (weak to moderate) with the fidelity score. However, the perceived appropriateness score for PDA was only negatively correlated with two scales: feasibility score ($r=-.104$, $p<0.05$) and sustainability score ($r=-.161$, $p<0.01$). Except for acceptability score, all the other scales' scores were correlated (weak to moderate) with the feasibility score. Except for appropriateness score, all other constructs were positively correlated (weak to moderate) with the intervention penetration score. Sustainability score was weakly negatively correlated with appropriateness score, while moderately positively correlated with all the remaining scales' scores. The highest and lowest correlation with the sustainability score of the approach was that of perceived penetration ($r=0.500$, $p<0.01$) and appropriateness ($r=-.161$, $p<0.01$) score respectively. Adaptability score of the approach was positively correlated with the other six scales, of which it was strongly correlated with implementation cost ($r=0.517$, $p<0.01$), while weakly correlated with implementation fidelity ($r=0.127$, $p<0.05$). Organizational readiness score was positively correlated ($P<0.01$) with seven scales, except with appropriateness score. Among the five scales significantly correlated with implementation cost, the lowest positive and significant correlation was observed with sustainability scores ($r=0.314$, $p<0.01$) (Table 4).

Mean Difference by Background Characteristics, and Explained Variation

One-way ANOVA (or independent samples *t*-test) showed that except the measurement scales reported at Table 5, all the remaining scales' scores were not significantly different by any of the background characteristics reported in Table 1. In this table, only the significant mean differences against background characteristics were reported. Since there was no significant composite mean difference between any of these grouping variables, the respective mean values and F/T-statistics were not reported. The test showed that appropriateness mean score was significantly ($F=27.18$ and $M=30.49$; $p=0.000$) different between male and female participants. The mean score for implementation scale was significantly different between those who cannot read and write and can read and write with/without formal

Table 4 Descriptive Statistics and Pearson's Correlation (Relationship) Coefficients Between Measurement Scales Used for Implementation Outcomes of PDA, Jimma Town, 2020

Variables/Scales		1	2	3	4	5	6	7	8	9
1.	Acceptability	1								
2.	Fidelity	0.600**	1							
3.	Appropriateness	-0.022	-0.031	1						
4.	Feasibility	0.099	0.110*	-0.104*	1					
5.	Penetration	0.132*	0.225**	-0.101	0.355**	1				
6.	Sustainability	0.265**	0.347**	-0.161**	0.372**	0.500**	1			
7.	Adaptability	0.050	0.127*	-0.039	0.448**	0.381**	0.335**	1		
8.	Organizational readiness	0.195**	0.245**	-0.091	0.454**	0.326**	0.402**	0.447**	1	
9.	Implementation cost	0.049	0.096	-0.006	0.396**	0.427**	0.314**	0.517**	0.332**	1
Number of items		7	7	7	6	6	4	5	5	5
Range (in score)		28	28	28	24	24	16	20	20	20
Number of respondents		372	372	372	372	372	372	372	372	372

Note: **Correlation is significant at the 0.01 level (2-tailed) and *correlation is significant at the 0.05 level (2-tailed).

Table 5 The Scale Mean Difference by Background Characteristics of the Study Participants, a Follow-Up Study, Jimma Town, 2020 (Significant Differences Only Reported)

Scale with Grouping Variable	M(SD) with 95% CI	F/T-Statistics(df)	p-value
Appropriateness by sex			
Female (n=301)	27.18(6.9); 26.39–27.96	-3.941(df=370)	0.000
Male (n=71)	30.49(2.9); 29.81–31.17		
Implementation cost by educational status			
Cannot read and write at all (n=101)	11.17(5.1); 10.16–12.18	4.247 (df=369)	0.033
Can read and write without formal education (n=38)	13.669(5.4); 11.87–15.45		
Can read and write without formal education (n=38)	13.669(5.4); 11.87–15.45	4.247 (df=369)	0.013
Can read and write with formal education (n=233)	11.099(5.1); 10.43–11.74		
Acceptability by their role played during implementation			
Facilitator (n=105)	26.10(6.1); 24.92–27.29	3.787(df=369)	0.021
End-user (n=254)	28.07(6.3); 27.28–28.85		
Fidelity by their role played during implementation			
Facilitator (n=105)	25.69(5.7); 24.59–26.78	5.002(df=369)	0.007
End-user (n=254)	27.80(6.1); 27.04–28.56		
Appropriateness by their role played during implementation			
Counselors/change agent (n=13)	33.54(1.9); 32.37–34.71	5.580 (df=369)	0.003
End-user (n=254)	27.46(7.6); 26.51–28.40		

(Continued)

Table 5 (Continued).

Scale with Grouping Variable	M(SD) with 95% CI	F/T-Statistics(df)	p-value
Feasibility score by their role played during implementation			
Facilitator (n=105)	15.01(5.5); 13.94–16.08	5.021(df=369)	0.006
End-user (n=254)	13.15(5.1); 12.52–13.77		
Organizational readiness by their role played during implementation			
Facilitator (n=105)	16.27(4.4); 15.42–17.11	6.668(df=369)	0.002
End-user (n=254)	14.35(4.9); 13.72–14.98		
Implementation cost by perceived level of engagement			
Highly/actively engaged (n=202)	10.99(5.2); 10.27–11.71	3.890(df=369)	0.017
Low engagement (n=42)	13.40(4.8); 11.92–14.89		
Appropriateness by perceived level of confidence to promote EBF using PDA			
Highly confident (n=203)	28.31(6.5); 27.41–29.20	9.616(df=368)	0.000
Uncertain (n=19)	21.11(8.9); 16.82–25.39		
Medium (n=58)	29.55(3.4); 28.66–30.44	9.616(df=368)	0.000
Uncertain (n=19)	21.11(8.9); 16.82–25.39		
Low confidence (n=92)	27.00(6.7); 25.62–28.38	9.616 (df=368)	0.001
Uncertain (n=19)	21.11(8.9); 16.82–25.39		

education. The mean scores for perceived acceptability, fidelity, feasibility and organizational readiness scales were significantly different between facilitators and end-users as a role played during implementation. However, the mean score for appropriateness scale was significantly different between counselors (PDs=change agent) and end-users (Table 5).

In addition to identifying the mean differences, the explained variation (interdependency) among those moderately to strongly correlated scales was checked using coefficient of determination (r^2). Accordingly, less than half of the variation (36%) in the acceptability scale score was explained to be due to the implementation fidelity. Among all the explained variations between the measurement scales, the maximum variation explained was observed between acceptability of the approach and implementation fidelity. It is worth explaining, quarter of the variation (25%) in

Table 6 Explained Variation (Interdependency) of Scales Among Those Moderately to Strongly Correlated Once, a Follow-Up Study, Jimma Town, 2020

S. No	Correlated Measurement Scales	r^{**}	r^2	%
1.	Acceptability*fidelity	0.600	0.360	36.0%
2.	Fidelity* sustainability	0.347	0.120	12.0%
3.	Feasibility* penetration	0.355	0.126	12.6%
4.	Feasibility* sustainability	0.372	0.138	13.8%
5.	Feasibility* adaptability	0.448	0.201	20.1%
6.	Feasibility* organizational readiness	0.454	0.206	20.6%

(Continued)

Table 6 (Continued).

S. No	Correlated Measurement Scales	r**	r ²	%
7.	Feasibility* implementation cost	0.396	0.157	15.7%
8.	Penetration* sustainability	0.500	0.250	25.0%
9.	Penetration* adaptability	0.448	0.201	20.1%
10.	Penetration* organizational readiness	0.326	0.106	10.6%
11.	Penetration* implementation cost	0.427	0.182	18.2%
12.	Sustainability* adaptability	0.335	0.112	11.2%
13.	Sustainability* organizational readiness	0.402	0.162	16.2%
14.	Sustainability* implementation cost	0.314	0.099	9.9%
15.	Adaptability* organizational readiness	0.447	0.200	20.0%
16.	Adaptability* implementation cost	0.517	0.267	26.7%
17.	Organizational readiness * implementation cost	0.332	0.110	11.0%

Note: **Correlation is significant at the 0.01 level (2-tailed), while *denotes “verses” or “with”.

the penetrability of the PDA intervention was due to the sustainability score as perceived by the study participants. As we can see in the table, the minimum coefficient of determination/variation explained (r^2) was observed between sustainability and implementation cost scales, in which 9.9% of the variation in sustainability score was due to the perceived implementation cost of the approach. The minimum variation explained implies that the intervention participants were not worried much about the implementation cost while rating the sustainability of the approach (Table 6).

Discussion

This study assessed and examined implementation outcomes of PDA as a strategy to promote EBF practice. As part of this assessment, valid and reliable measurement scales were developed. The nine newly developed scales (with a total of 52 reliable items) were acceptability, fidelity, appropriateness, feasibility, penetration, sustainability, adaptability, organizational readiness, and implementation cost. During the validation process, the number of items was significantly reduced, which will make it easy for future research use. For those valid and reliable items, the mean score was determined to understand the degree of importance as rated by intervention participants. An item with the highest mean score indicates that the item was highly rated in that scale. In this study, highly rated items were loaded to acceptability scale/factor, while an item with the lowest mean score was loaded to feasibility scale. Besides the variation in score (degree of rate), all valid items were internally consistent (reliable) to measure implementation outcomes of PDA to promote EBF practice. Scholars agreed that the development of such tested measurement scales would advance implementation science, especially to conduct a newly approached intervention.^{14,17,21} This means, contextually developed items reflect and truly measure the conceptual content of the nine implementation outcomes.

Implementation outcomes of PDA were summarized using the scale mean scores and the TVE by the emerged components. The higher factor score created on a factor analysis indicates a higher degree of importance of that factor with regard to implementing PDA to promote EBF. The maximum mean value was determined for appropriateness scale and the highest variance was belongs to it. The respective mean values indicated that the study participants rated appropriateness and acceptability scale highly, while they rated implementation cost scale low. These highest and lowest mean scores imply the intervention participants hold a perception that the PDA to promote EBF is more likely

appropriate and acceptable, while costly to implement. In sum, the nine contextually developed scales explained 72.1% of the total variability in psychometrically measured implementation outcomes of PDA. This means these outcomes were well measured by these scales. The finding showed that appropriateness, acceptability, fidelity, penetration, organizational readiness and sustainability scales were rated considerably above the mean score, while feasibility, adaptability, and implementation cost scales were rated below. This means, the PDA was perceived as appropriate, acceptable, fidable (implementable as planned with quality), penetrable through organizationally/structurally ready community setting, and sustainable to promote EBF, while it was not feasible, adaptable, and costly. This negative perception might be due to the fact that much of the intervention activities were conducted through volunteerism and minimal incentives, besides the fact that the deviance inquiry process is time-consuming and needs much effort to be adapted to the context.^{21,25} Many of the highly rated scales were more theoretically relevant and can be judged from mere understanding of PDA from the inception training, whereas all those scales rated low might be due to the perception developed through practical experience. However, the low rating of the three scales and high rating of the sustainability of the new approach need further study since empirically “less feasible”, “difficult to adapt” and “costly” approach could not be sustainable.^{3,5}

Studies reported that these implementation outcomes are best represented from an empirical perspective to have differentiable constructs though they are highly complex.^{3,5,14} Hence, this study contributes to the science by developing valid and reliable measures of implementation outcomes that would fit the context, clarify the concept, promote common terms as discussed by another study,²⁶ and concerns about the state of measurement that can explain the process of the PDA implementation.^{6,27,28} Although such outcome measures are central to understanding the extent to which implementation is successful, valid and reliable measures are lacking and mostly not documented well.⁵ This study fills that gap by developing valid and reliable measures of implementation outcomes that are important to replicate in a range of implementation studies that will use a PDA in a pilot or effectiveness studies. As far as conducting an implementation study to disseminate for further implementation,²⁹ integrating effectiveness with implementation study is very important. As an effectiveness study needs ultimate outcome measure to understand its effect, the implementation outcome aspect is equally relevant and needs to be measured to understand whether it is acceptable, appropriate, sustainable, feasible and so on.^{30,31} These outcomes are relevant to assess intervention participants’ perceptions of such public health interventions as well as assessing perceptions of implementation strategies, which are assumed to be new or complex.^{3,32,33} Assessing these outcomes early in the research process may ensure that interventions and implementation strategies are optimized and fit with end-users’ preferences. However, we assessed such outcomes during post-implementation period just to learn from the process, and to contribute to future implementation science advancement, rather our intervention/trial was over in the study area. Most importantly, this study will benefit the health system if it wants to adapt the PDA or to sustain it as per our scale-up plan, in the study area, to promote EBF.

The majority of the psychometric measurement scales were significantly correlated with each other either positively or negatively with different levels of strength. Implementation fidelity, penetration, organizational/community readiness and sustainability were positively and significantly correlated with acceptability of the approach. This means these outcome measures are interdependent upon each other and increased positively in order to influence the practice of EBF through PDA intervention. This also implies that the PDA intervention was implemented as per the protocol with expected quality (fidelity), penetrated through the existing organizational and community structure in an acceptable way to the participants so that it would be sustainable. The strongest³⁴ relationship observed between the acceptability and implementation fidelity supports this concept. Once a certain thing is acceptable, it can be implemented as designed, or the reverse can happen.^{3,14} The other relationship shows a feasible, penetrable, sustainable, adaptable approach can be integrated within an organizationally ready setting if it was implemented as expected (fidelity). However, we are not clear why the perceived appropriateness and implementation cost were not correlated with fidelity, which disproves the pre-existing theoretical relation.^{3,13,28} In addition, why the perceived appropriateness had reverse relationship (when appropriateness score increased, the other two scales’ scores decreased) with feasibility and sustainability of the approach, needs further study. Indeed, this tells us the appropriate approach may not be feasible and sustainable, which also confirms the pre-existing conceptual difference.^{3,13} The other finding also conforms to the generic understanding that acceptable intervention may not be feasible, and its appropriateness will not be a guarantee of its penetration. However, to have a sustainable approach/intervention, all the remaining six implementation outcomes

could be rated positively regardless of its appropriateness, which could be perceived reversely. Practically this study revealed that sustainability is highly related with penetration otherwise its continuity will be threatened.

Adaptability of the approach was positively correlated with many of the other scales, of which it was strongly correlated with implementation cost, while weakly correlated with implementation fidelity. This is in agreement with the classical conceptual framework of fidelity which does not address the issue of how to adapt an intervention while still maintaining its effectiveness. In theory, the implications of adaptations on fidelity may be different. When adding something new, fidelity can be easily maintained. When a component is suppressed or modified, fidelity is threatened. In general, adaptability may or may not be affected by fidelity depending on the modifications made regardless of the challenge to measure those outcomes.^{14,17} In addition to this, perceived appropriateness and implementation cost may not be necessary causes to have receptive organization, rather the remaining measures could matter. Organizational readiness is likely to be highest when organizational members not only want to implement an organizational change but also feel confident, empowered and accept the approach so that they can do so.^{3,14}

Our study considered all the intervention participants to measure the implementation outcomes of PDA in promoting EBF, which is the first of its kind in integrating implementation and trial study. In this regard, its appropriateness was differently perceived by male and female participants, which males rated highly. This may be due to the stronger feeling of males wanting to help lactating mothers through other experienced deviants. This suggests some clue about the importance of engaging males while intervening to promote EBF. Also, implementation cost was differently perceived by those who cannot read and write as compared to their counterparts. This scale was rated low by those who cannot read and write which is difficult to fully explain in this study. Roughly, it might be due to the fact that learning and understanding concepts and complex ideas increase with increasing educational status. This implies that considering educational status while designing and implementing such community-based counseling and social support intervention is paramount. The role participants had during implementation affected their perception, so they viewed its acceptability, fidelity, feasibility, appropriateness and organizational readiness, differently from facilitators, end-users and counselors/PDs. Acceptability and fidelity were highly rated by the end-users as compared to the facilitators, while feasibility and organizational readiness were highly rated by the facilitators. But appropriateness was highly rated by the counselors/PDs. This generally implies the role they have, the activity they conducted and the actual and perceived level of engagement can affect participants' understanding to rate given implementation outcomes differently. This also provides information regarding using PDs as a social actors and credible sources of EBF information and skill sharing resources in the urban community. Surprisingly, the implementation cost was rated highly by those who perceived as they had low level of engagement during implementation. Self-report and information biases were clearly reflected here which may mislead the future action. Otherwise, successful community-based programs are often characterized by increased acceptability, appropriateness and fidelity of the intervention that in turn ensure its sustainability.^{35,36} The social support and experience sharing nature of learning how to EBF might have contributed to the increased fidelity,³⁷ in addition to the shared plan at the beginning with the stakeholders at inception training. Our finding suggests the need to consider individuals' background while dealing with such a new approach to minimize a perception difference so as to boost effort together for success and sustainability issues.

The observed higher and positive perceptions toward the implementation outcomes reported in our study may be due to many reasons such as multi-stakeholders' engagement from inception to implementation until evaluation stages, influences of the positive perception to the acceptability and appropriateness of the PDA among the key stakeholders, and the improved community/parental acceptance of the informational and social support given by the PDs. Such findings imply that there is a possibility to intensify and replicate PDA as a new strategy to promote EBF and learn from different realities at urban setting. However, perception of the implementation by participants toward feasibility, implementation cost and adaptability need great attention in order for it to be successful during implementation, and to achieve the desired outcome.

In addition to considering background characteristics, the explained variation (interdependency) among those correlated scales is important. Accordingly, less than half of the variation in the acceptability was due to the implementation fidelity which is the maximum variation explained. The next highest explained variation was seen between adaptability of the approach and its perceived implementation cost. This means one-fourth of the variation observed in

adaptability was due to the perceived implementation cost. It is worth explaining, quarter of the variation in the penetrability was due to the perception of raters about sustainability. Minimal variation observed in sustainability was due to the perceived implementation cost of the approach. This implies that the intervention participants were not worried much about the implementation cost while rating the sustainability of the approach highly. Implementation costs may vary according to variations in their complexity, type of implementation strategy used, settings of varying complexity and overheads, and the overall costs of delivery will vary by setting.

Evidence shows that sustainability of a certain intervention approach is affected by the characteristics of the implementing individuals (coordinators, supervisors, facilitators, change agents, and end-users) in addition to the organizational and community support.^{15,38,39} It was reported that the quality of staff training and level of engagement could empower the stakeholders for future performance.⁴⁰ This is also supported by previous studies which showed that programs with higher levels of fidelity and adaptability are likely to be sustainable.^{41,42} However, sustainability can be affected in two ways such as ignoring the supporting parameters at designing and implementation stage⁴³ and/or a short-lived trial study which has no plan for scale-up. This calls for proactive thinking and careful design of the program components such as training, partnerships, stakeholders' engagement, and role identification in order to enhance the continuation of the program in an acceptable way. Another study revealed that the adoption of community-based interventions depends on the extent of community and stakeholders' involvement and acceptance, which in turn influences sustainability.³⁷ The influence of adaptability on sustainability might be due to the degree of the intervention's "fitness" with the local context, needs, and health program. This means an intervention that is well-integrated with the existing health system would be institutionally maintained.^{38,44} With a sense of mutual benefit, if intervention participants have clear duties and responsibilities, the implementation outcomes will be highly rated. High rates for such outcomes of PDA depend on the capabilities of the raters to implement the approach, the resources and material support, and community involvement and readiness to build sense of ownership.⁴⁵ Otherwise, a service giver and receiver type of approach does not work with PDA.

Strengths/Implications, and Limitations

We intended to assess and develop specific scales related with implementation outcomes of PDA as an intervention to promote EBF. Our aim is supported by other scholars⁴⁶ who argue that measures need to be pragmatic if they are to be useful outside the context of research. Among the pragmatic features of these measures, psychometric strength, actionability, conceptual and content clarity and relevance to stakeholders were considered in our study. It was developed with experts' and stakeholders' input and therefore might reflect what stakeholders view as important. We worked with experts to comment on these measures, with the goal of developing new scales that can be used to assess the pragmatic nature of the outcomes with psychometric properties that address/fit our stakeholders.⁷ We recommend that researchers who are working on a similar approach and settings apply these pragmatic implementation outcome measures as they are useful in informing decision-making, compatible with the settings, easy to use, and acceptable,⁴⁷ and they are especially relevant for resource-limited settings.^{4,48} Among the strengths, we developed brief outcome measures that were started with developing as much as high number of items per constructs, tested through psychometric measurement ways, and developed each item as specific as possible to the approach, behavior under promotion, and context. For instance, acceptability items are specific to the purpose. The context- or treatment-specific nature of the tool will increase the chances of its use broadly in implementation research and practice. It is known that in implementation science, the majority of current measures are developed for the purpose of a generic study (usually with minimal conceptual clarity and psychometric testing) which is difficult to use again. However, as a limitation, we cannot conclusively declare whether the developed measures can address a study conducted in different contexts for different behaviors, as it is. Specifying a wide range of the PDA principles, assumptions, and activities would ultimately enhance ownership and develop a positive perception about implementation outcomes. The measurements used in our study to examine these outcomes of PDA could provide methodological solutions to the recently growing implementation sciences.⁴⁹ Thus, this result contributes to fill gaps in evidence on the application of the PDA from the design to sustainability, as perceived by the participants in resource-limited settings. The lessons learned from the perspectives of the key stakeholders would help to advance effective methods to use, to engage, empower and retain participants to promote EBF. As a strength, our study

conceptualized and measured implementation outcomes which act as intermediate indicators between the intervention and its service outcome in short way, while between the intervention and desired effects in long way. This is useful to understand the relationship between the intervention and desired effects.³

However, self-report, recall, social desirability and information bias could be the possible limitations. To minimize reporting bias, items were prepared to verify each other (negatively stated items were included to help catch attention). Future research would benefit from further analyzing the discriminant and predictive validity using Structural Equation Modeling (SEM) between latent constructs at different level of structure. Replication is also recommended. Sensitivity to change and specificity of those measures/scales can be tested using advanced regression model. In the process, discriminate and structural validity with test-retest reliability and prediction modeling could be done. This would give an opportunity to see the cause-effect relationship beyond mere correlation. Causation always implies correlation but correlation does not necessarily imply causation.³⁴ Testing the measures with people who have different background characteristics and roles (during implementation), as we did in this study, would benefit generalizability. Further study can be done to examine the rate of decline or improvement in implementation outcomes over time to predict EBF practice while using PDA. We encourage researchers to replicate this methodology and suggest further refinements for the same approach and behavior in different contexts.

Conclusion

This study identified nine valid implementation outcome measures (acceptability, fidelity, appropriateness, feasibility, penetration, sustainability, adaptability, organizational readiness, and implementation cost) of PDA as a new strategy to promote EBF, which are important indicators used to explain the implementation process. The process revealed 52 items which are valid and reliable measures of these outcomes. The correlation and variation explained in our context support the theoretical/conceptual relationship. Except for perceived feasibility, adaptability and implementation cost, the remaining six outcomes of PDA were highly rated by intervention participants. These findings suggested that the higher rate for an outcome measure implies it is a promising approach in promoting EBF in urban community by PDs. Therefore, the informational counseling and social support intervention using PDA should consider such intermediate outcomes to be successful in improving EBF practice. Moreover, addressing the perception of intervention participants about feasibility, adaptability and cost issues needs great attention throughout the implementation period to minimize the implementation challenges. These specific psychometric measurement scales can be used in a formative evaluation potential to understand the process) of pragmatic properties by adapting to the context.

Abbreviations

BF, Breast Feeding; CFA, Confirmatory Factor Analysis; CHWs, Community Health Workers; CI, Confidence Interval; cRCT, Cluster Randomized Controlled trial; EBF, Exclusive Breast Feeding; EFA, Exploratory Factor Analysis; HDA, Health Development Army; HEPs, Health Extension Professionals; IRB, Institutional Review board; KM, Kilo meters; MPH, Masters of Public Health; PCA, Principal Component Analysis; PD, Positive Deviance; PDA, Positive Deviance Approach; PDI, Positive Deviance Inquiry; PDs, Positive Deviants; RCT, Randomized Controlled trial; SD, Standard Deviation; SEM, Structural Equation Modeling; TVE, Total Variability Explained; UHEPs, Urban Health Extension Professionals; US\$, United States Dollar; WDA, Women Development Army; WHDA, Women Health Development Army.

Data Sharing Statement

The authors confirm that the data supporting the findings of this study are available within the article. For some methodological details with regard to the trial, a previously published article was cited in this paper.²¹

Ethical Approval and Consent to Participate

The letter of ethical approval was secured from the Institutional Review Board (IRB), Jimma University, Institute of Health with Ref. No: IHRPGD/728/2020 and dated 27/08/2020, to collect data for this post-implementation follow-up study. A letter of support was obtained from Jimma town health office to respective Kebeles/clusters. Informed consent

was taken from each participant. They were informed about their full right to participate or refuse participation in this follow-up study. Confidentiality was kept using codes. The contact details taken during enrollment and implementation period were used to reach out to them for this follow-up study. This study was conducted in accordance with the Declaration of Helsinki.

Acknowledgments

We acknowledge Jimma University (JU) for the financial support provided to conduct this follow-up study. In addition, thanks to International Institute for Primary Healthcare-Ethiopia (IIFPHC-E) and John Hopkins University-USA for the financial and technical support provided for the previous trial, without which this follow-up study could not have been done. We also thank Jimma Town Health Office, study participants, and data collectors.

Disclosure

The authors have disclosed that they do not have any potential conflicts of interest for this work, financial or otherwise.

References

- Leykum LK, Pugh JA, Lanham HJ, Harmon J, McDaniel RR. Implementation research design: integrating participatory action research into randomized controlled trials. *Implement Sci.* 2009;4(1):1–8. doi:10.1186/1748-5908-4-69
- Allen KD, Bierma-Zeinstra SMA, Foster NE, Golightly YM, Hawker G. OARSI clinical trials recommendations: design and conduct of implementation trials of interventions for osteoarthritis. *Osteoarthritis Cartilage.* 2015;23(5):826–838. doi:10.1016/j.joca.2015.02.772
- Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health.* 2011;65:76.
- World Health Organization. Implementation research [Internet]. WHO; 2014. Available from: <https://tdr.who.int/home/our-work/strengthening-research-capacity/implementation-research-training-materials>. Accessed November 5, 2022.
- Lewis CC, Fischer S, Weiner BJ, Stanick C, Kim M, Martinez RG. Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. *Implement Sci.* 2015;10(1):1–17. doi:10.1186/s13012-015-0342-x
- Martinez RG, Lewis CC, Weiner BJ. Instrumentation issues in implementation science. *Implement Sci.* 2014;9(1):1–9. doi:10.1186/s13012-014-0118-8
- Lewis CC, Weiner BJ, Stanick C, Fischer SM. Advancing implementation science through measure development and evaluation: a study protocol. *Implement Sci.* 2015;10(1):1–10. doi:10.1186/s13012-015-0287-0
- World Health Organization. Exclusive breastfeeding for optimal growth, development and health of infants; 2019: 85.
- World Health Organization. Guidelines for essential newborn care encompass. 2017: 1255.
- Bowen DJ, Kreuter M, Spring B, et al. NIH Public Access: how to design feasibility study. *Am J Prev Med.* 2010;36(5):452–457. doi:10.1016/j.amepre.2009.02.002
- Rogers E. Diffusion of Innovations; 1995: 260.
- Hovmand PS, Gillespie DF. Implementation outcome study. *J Behav Health Serv Res.* 2010;37(79):30–45.
- Proctor E. Implementation outcomes. 2011 National Child Welfare Evaluation Summit Washington DC; 2011: 5–15.
- Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci.* 2017;12(1):1. doi:10.1186/s13012-017-0635-3
- Rabin BA, Brownson RC, Haire-joshu D, Kreuter MW, Weaver NL. A glossary for dissemination and implementation research in health. *J Public Health Manag Pract.* 2008;14(2):117–123. doi:10.1097/01.PHH.0000311888.06252.bb
- Pérez D, Stuyft Van Der P, Zabala C, Castro M, Lefèvre P. A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. *Implement Sci.* 2016;11(91):1–11. doi:10.1186/s13012-016-0457-8
- Weiner BJ. A theory of organizational readiness for change. *Implement Sci.* 2009;4(67):1–9. doi:10.1186/1748-5908-4-67
- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009;4(1):1–15. doi:10.1186/1748-5908-4-50
- Zeitlin M, Vynckt S. The use of nutritional “positive deviants” to identify approaches for modification of dietary practices. *Am J Public Health.* 1976;66(1):38–42.
- Bolles K, Speraw C, Berggren G. LJTF (hearth). community-based nutrition activities informed by the positive deviance approach in Leogane, Haiti: a programmatic description. *Food Nutr Bull.* 2002;23(4):11–17. doi:10.1177/15648265020234S103
- Siraneh Y, Woldie M, Birhanu Z. Effectiveness of positive deviance approach to promote exclusive breastfeeding practice: a cluster randomized controlled trial. *Risk Manag Healthc Policy.* 2021;14:3483–3503. doi:10.2147/RMHP.S324762
- Jimma Town health office. Jimma Town health office annual report of 2017; personal communication; 2017.
- World Health Organization. Breastfeeding counselling a training course participants’ manual part one Sessions 1–9; 2009: 1–9.
- Hinkin TR. A brief tutorial on the development of measures for use in survey questionnaires. *Organ Res Methods.* 1998;1(1):104–121. doi:10.1177/109442819800100106
- Srivastava A, Gwande K, Bhattacharya S, Singh VK. Impact of the positive deviance approach on breastfeeding practices among Tribal pregnant women: a before – after intervention study. *Chrimed J Health Res.* 2019;6(2):222–228.
- McKibbin KA, Lokker C, Wilczynski NL, et al. A cross-sectional study of the number and frequency of terms used to refer to knowledge translation in a body of health literature in 2006: a Tower of Babel? *Implement Sci.* 2010;5(1):1–11. doi:10.1186/1748-5908-5-16
- Eton & Lepore. NIH public access. *J Am Acad Child Adolesc Psychiatry.* 2008;23(1):1–7.

28. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. *Implement Sci.* 2013;8(1):1–11. doi:10.1186/1748-5908-8-139
29. Brownson RC, Jacobs JA, Tabak RG, Hoehner CM, Stamatakis KA. Designing for dissemination among public health researchers: findings from a national survey in the United States. *Am J Public Health.* 2013;103(9):1693–1699. doi:10.2105/AJPH.2012.301165
30. Curran GM, Bauer M, Mittman B, Pyne JM. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care.* 2012;50(3):217–226. doi:10.1097/MLR.0b013e3182408812
31. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. *BMC Psychol.* 2015;3(1):1–12. doi:10.1186/s40359-015-0089-9
32. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci.* 2015;10(1):1–14. doi:10.1186/s13012-015-0209-1
33. Powell BJ, Proctor EK, Glisson CA, et al. A mixed methods multiple case study of implementation as usual in children’s social service organizations: study protocol. *Implement Sci.* 2013;8(1):1–12. doi:10.1186/1748-5908-8-92
34. Hemphill JF. Interpreting the magnitudes of correlation coefficients. *Am Psychol.* 2003;58(1):78–79. doi:10.1037/0003-066X.58.1.78
35. Palinkas LA, Chou CP, Spear SE, Mendon SJ, Villamar J, Brown CH. Measurement of sustainment of prevention programs and initiatives: the sustainment measurement system scale. *Implement Sci.* 2020;15(1):1–15. doi:10.1186/s13012-020-01030-x
36. Moore G, Audrey S, Barker M, et al. Process evaluation in complex public health intervention studies: the need for guidance. *J Epidemiol Community Health.* 2014;68(2):101–102. doi:10.1136/jech-2013-202869
37. Ceptureanu SI, Ceptureanu EG, Luchian CE, Luchian I. Community based programs sustainability. a multidimensional analysis of sustainability factors. *Sustainability.* 2018;10(3):1–15. doi:10.3390/su10030870
38. Walugembe DR, Sibbald S, Le Ber MJ, Kothari A. Sustainability of public health interventions: where are the gaps? *Health Res Policy Syst.* 2019;17(1):1–7.
39. Savaya R, Elsworth G, Rogers P. projected Sustainability of Innovative Social Programs. *Eval Res.* 2009;33:189–205. doi:10.1177/0193841X08322860
40. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008;41(3–4):327–350. doi:10.1007/s10464-008-9165-0
41. Barrera M, Berkel C, Castro FG. Directions for the advancement of culturally adapted preventive interventions: local adaptations, engagement, and sustainability Manuel. *Prev Sci.* 2020;18(6):640–648.
42. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implement Sci.* 2013;8(1):1–11. doi:10.1186/1748-5908-8-117
43. Pluye P, Potvin L, Denis JL. Making public health programs last: conceptualizing sustainability. *Eval Program Plann.* 2004;27(2):121–133. doi:10.1016/j.evalprogplan.2004.01.001
44. Olumide AO, McGuire C, Calhoun L, Speizer I, Babawarun T, Ojengbode O. Factors promoting sustainability of NURHI programme activities in Ilorin and Kaduna, Nigeria: findings from a qualitative study among health facility staff. *BMJ Open.* 2020;10(10):1–11. doi:10.1136/bmjopen-2019-034482
45. Iwelunmor J, Blackstone S, Veira D, et al. Toward the sustainability of health interventions implemented in sub-Saharan Africa: a systematic review and conceptual framework. *Implement Sci.* 2016;11(1). doi:10.1186/s13012-016-0392-8
46. Glasgow RE. Pragmatic measures: what they are and why we need them. *Am J Prev Med.* 2013;45(2):237–243. doi:10.1016/j.amepre.2013.03.010
47. Powell BJ, Weiner BJ, Stanick CF, Halko H, Dorsey CLC. Stakeholders’ perceptions of criteria for pragmatic measurement in implementation: a concept mapping approach (oral presentation). In: *9th Annual Conference on the Science of Dissemination & Implementation.* Washington, D C: Academy Health and the National Institutes of Health; 2016.
48. Beidas RS, Stewart RE, Walsh L, et al. Free, brief, and validated: standardized instruments for low-resource mental health settings. *Cogn Behav Pract.* 2015;22(1):5–19. doi:10.1016/j.cbpra.2014.02.002
49. Domitrovich CE, Bradshaw CP, Poduska JM, et al. Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework Celene. *Adv Sch Ment Health Promot.* 2008;1(3):6–28. doi:10.1080/1754730X.2008.9715730