



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

From Internal Capability to External Outcome: A Systematic Review of Personality Development Programmes and Women's Empowerment

Swati Saini

¹Research Scholar, Department of Education, Sunrise University

Dr. Hari Chand

²Associate Professor, Department of Education, Sunrise University

Abstract

Women receive personality development programmes, life skills education and soft skills training in large numbers, and behind the delivery lies a premise about where the obstacle to their advancement sits: inside them, in confidence, communication and assertiveness. What the accumulated evidence establishes about that premise is what this review asks. A protocol drove the work, which is reported to PRISMA 2020. Alongside eight electronic databases, repositories, hand-searching and citation searching were used. Two reviewers screened independently; design-matched instruments appraised what was included, and the synthesis was thematic. Searching turned up 2,010 records; 1,483 reached screening, 196 were read in full, and the synthesis took in 42 studies carried across 47 publications. A gradient runs through the findings. Psychological outcomes came from 41 of the 42 studies at a mean standardised effect of +0.63, social outcomes from 29 at +0.46, educational from 16 at +0.31, civic from 14 at +0.28, economic from 17 at +0.20. Five analytical themes were constructed. Internal change is the most consistent of the effects and the best evidenced. Whether it transfers to external outcomes is conditional and not automatic, with transfer showing up mainly where programmes ran longer and sat within a continuing structure. Leadership self-perception outruns leadership attainment. Household and institutional response gatekeeps the transfer. Durability, meanwhile, goes essentially unmeasured: 32 of the 42 studies took no measurement once the programme had closed.

Keywords: personality development; women's empowerment; life skills education; self-efficacy; systematic review; thematic synthesis

1. Introduction

Today women enter education and the labour market in numbers two generations ago would have found hard to imagine. Leadership, enterprise and the rooms where decisions are taken still hold few of them. A gap separates qualification from position, and what produces it is contested.

Dominance, institutionally, belongs to one answer. Inside the woman is where it puts the binding constraint, in her confidence, her communication, her willingness to put herself forward. From which the remedy follows: a structured programme that develops those



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

capacities. On that premise, colleges, employers, non-governmental organisations and government schemes run personality development courses, life skills education, soft skills training and women's leadership programmes, reaching very large numbers of participants at substantial public and private cost.

Systematic answers to whether the premise holds have not come from the field, partly for structural reasons. Evaluations are not scarce. They are dispersed, though, over psychology, education, management, development studies, social work and gender studies, each of which has its own name for the same class of intervention. A similar conclusion came out of Nasheeda et al.'s (2019) review of life skills education for adolescents: psychosocial effects reported commonly, methodological quality weak. What nobody has done is collect the female-specific evidence across those vocabularies, appraise it by one standard, and ask what it establishes taken together.

The review in question is reported here. Three questions drive it. Under a personality development programme, what changes, and how dependably? Of the dimensions of empowerment, which respond, and in what sequence? And on what conditions, if any, does internal change turn into external outcome? Answering that last one is the paper's principal contribution, and the answer sits less comfortably than the programme literature assumes.

2. Conceptual Background

The analysis is framed by two bodies of theory, developed for the most part in isolation from each other. One of them explains why a woman's reading of her own capability should shift under a structured programme. Efficacy information, Bandura (1977) proposed, reaches a person through four channels: enactive mastery in practice, vicarious experience from watching others, verbal persuasion, and reinterpreted physiological arousal. More or less by construction a developmental programme opens all four, which makes a measurable shift in efficacy beliefs the result one would expect rather than one to be surprised by.

What empowerment is concerns the second body of theory. Resources, agency and achievements were the three parts Kabeer (1999) decomposed it into, with the insistence that they relate to one another without being interchangeable: act on agency and you have neither supplied the resources nor delivered the achievements. The same cut appears in different language in Sen's (1999) distinction between a capability and an achieved functioning. Acquiring the capacity to act is one thing; meeting the occasion to exercise it is another, and a woman may never meet it.

Beside both runs a third, critical line of thought. On Batliwala's (2007) account, the collective and political content for which the word empowerment was coined has steadily leaked out of it, until what is left can be installed in an individual on commission. Where interventions do succeed, Cornwall (2016) finds after asking what actually works in women's empowerment programming, relationships and structures have been altered and not dispositions on their own. The same case is made for women's leadership programmes by Ely et al. (2011), who put the



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

operative constraint in organisational practice and in second-generation gender bias, not in the woman attending.

Read together, the two traditions produce an expectation that can be tested: the internal term should move reliably under programmes of this kind, the external terms weakly, and whatever does move the external terms should be something more than the programme. Whether the evidence behaves that way is the question the review takes up.

3. Method

3.1 Design and protocol

Systematic review with thematic synthesis is the design adopted. A protocol went in for approval and was approved before any searching, and reporting follows the PRISMA 2020 statement (Page et al., 2021). Meta-analysis got consideration and was turned down. At one end of this field's outcome instruments sit validated psychological scales, at the other unstandardised questionnaires the programme providers wrote themselves, and anything pooled across that span would be numerically precise with nothing stable underneath it.

3.2 Search strategy

This topic is spread over psychology, education, management and development studies, and no one database takes in the whole of it, so eight went into the search in combination, backed by four supplementary methods (Booth et al., 2016). The terms sat in four concept blocks, joined with OR inside each block and with AND between them. Pilot searching over three rounds, against ten studies known beforehand to be relevant, refined the strategy. Table 1 gives the sources.

Table 1. Information sources searched

Source type	Sources	Rationale
Electronic databases	Scopus; Web of Science; PsycINFO; ERIC; JSTOR; EBSCO Business Source; Shodhganga; Google Scholar (first 200 results)	Multidisciplinary coverage; Shodhganga for Indian doctoral theses held by no international index
Repositories	UN Women; UNESCO; UNDP; World Bank; NITI Aayog; NSDC	Evaluation reports of large programmes not published in journals
Citation searching	Backward and forward searching on all studies passing full-text screening	Recovers work missed by database indexing or vocabulary



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

Source type	Sources	Rationale
Hand-searching	Most recent five volumes of six core journals	Compensates for indexing lag

Note. Searches covered publications from 2005 to 2026 in English, with Hindi-language theses included where full text was accessible.

3.3 Eligibility criteria

Written into the protocol, the criteria were applied as written. To qualify, a study had to evaluate a structured programme whose content, duration and delivery mode were stated, whose declared aim covered communication, confidence, self-awareness, interpersonal capability, emotional regulation, decision-making or leadership behaviour, whose participants were women or adolescent girls aged 15 and above, and which reported at least one psychological, social, economic, educational or leadership outcome. Mixed-sex samples qualified only where women's results stood separately. Table 2 summarises the criteria.

Table 2. Inclusion and exclusion criteria

Dimension	Included	Excluded
Participants	Women and adolescent girls aged 15 and above; mixed-sex samples with results disaggregated by sex	Mixed samples without disaggregation
Intervention	Structured programmes with stated content, duration and mode, however labelled	Incidental exposure; purely technical or vocational training with no developmental component
Outcome	At least one psychological, social, economic, educational or leadership outcome	Satisfaction, attendance or completion alone
Design	Experimental, quasi-experimental, survey, mixed-methods, qualitative and formal evaluations	Conceptual essays without primary data
Quality	Summary appraisal score of 0.55 or above	Studies below the threshold

3.4 Study selection and appraisal

Screening of every de-duplicated record was independent and duplicated across two reviewers, title and abstract first, full text second, with the criteria piloted beforehand on a shared sample of fifty records. Agreement was put in numbers with Cohen's kappa and read against the benchmarks of Landis and Koch (1977). The value at title and abstract came to 0.78,



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

substantial; at full text to 0.84, almost perfect. Each sat above the 0.70 threshold the protocol had fixed, so no re-specification was needed before screening continued. Thirteen records went to discussion and four to a third reviewer.

The fourteen-item checklist of Kmet et al. (2004) appraised quantitative and mixed studies, its ten-item companion the qualitative ones, and mixed-methods designs were cross-checked against the Mixed Methods Appraisal Tool (Hong et al., 2018). Each score is a proportion of the maximum applicable. Inclusion required 0.55, which kept out studies of limited quality while retaining adequate ones and identifying them as such.

3.5 Extraction and synthesis

A structured extraction form, piloted on five studies, captured the data. Synthesis ran through the three stages Thomas and Harden (2008) describe: findings coded line by line, then descriptive themes assembled without stepping past what the evidence says, then analytical themes built to answer the review questions. Studies supplying the statistics for it have their effect sizes given as Hedges' *g*, read against the conventional thresholds of 0.20, 0.50 and 0.80 for small, moderate and large effects (Cohen, 1988); the rest are reported as direction of change. Confidence in every analytical theme was rated on an approach adapted from GRADE-CERQual (Lewin et al., 2018), which weighs methodological limitations, coherence, adequacy of data and relevance.

4. Results

4.1 Study selection

Searching the databases returned 1,847 records, with a further 163 from the supplementary sources, for 2,010 in total. Removing duplicates took out 527 and left 1,483 unique records to screen. Title and abstract excluded 1,268 of those. Full texts were sought for 215; 19 proved unobtainable and are reported as unretrieved rather than excluded. That left 196 assessed in full, of which 141 were excluded with reasons. Below the quality threshold fell a further 13, leaving 42 studies reported in 47 publications. The flow appears in Figure 1.

Figure 1. Flow of studies through the review

Stage	Records at this stage	Removed or excluded
Identification	Databases 1,847 Other sources 163 Total 2,010	—
	↓	
	Unique records after de-duplication 1,483	Duplicates removed 527
	↓	
Screening	Screened on title and abstract 1,483	Excluded 1,268



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

Stage	Records at this stage	Removed or excluded
	↓	
	Full texts sought 215	Not retrieved 19
	↓	
Eligibility	Full texts assessed 196	Excluded with reasons 141
	↓	
	Quality-appraised 55	Below 0.55 threshold 13
	↓	
Included	Studies included 42 Reports 47	—

Note. Flow reported in accordance with the PRISMA 2020 statement (Page et al., 2021).

Mixed-sample studies that never break their results out by sex supplied 34 of the 141 exclusions at full text, more than any other ground. Treat that as a finding about the field, not as housekeeping. Developmental programmes delivered to mixed groups are evaluated in quantity and the results published pooled, and a question about what happens to the women in them cannot be answered from any of it.

4.2 Characteristics of the included studies

Between them the 42 studies cover 15 countries and the years 2006 to 2025, with 17 from India and 21 from South Asia as a whole. Another 7 came from sub-Saharan Africa and 5 from high-income Western settings. Students and city sites prevail, a fact about where these programmes are easiest to run and evaluate rather than about where they are most wanted. Quasi-experimental and pre-experimental designs take up most of the set, which holds just 2 randomised trials. Contact hours stretch from 6 to 120, a wide span for something the literature treats as a single class of intervention, and the chief source of the heterogeneity that ruled meta-analysis out.

Across the 42 studies there are 41 distinct outcome instruments. Much of the quantitative measurement rests on two well-validated scales, which nearly always arrive with reliability evidence attached. Facing them are author-constructed empowerment indices and confidence questionnaires, drawn on almost as often and validated almost never. Which constructs fall on which side of that line is what makes it matter: the outcomes measured best are the narrowest.



Table 3. Outcome instruments used in three or more studies

Instrument	Construct	Studies	Reliability reported
Generalised Self-Efficacy Scale	Self-efficacy	9	8 of 9
Rosenberg Self-Esteem Scale	Self-esteem	8	7 of 8
Author-constructed empowerment index	Composite empowerment	7	2 of 7
Author-constructed confidence questionnaire	Self-confidence	6	1 of 6
Psychological Capital Questionnaire	Psychological capital	4	4 of 4
Emotional intelligence inventory (various)	Emotional intelligence	4	3 of 4
Semi-structured interview schedule	Qualitative outcomes	9	Not applicable
Employment or enrolment record	Objective attainment	5	Not applicable

Note. Counts exceed 42 because most studies used more than one instrument.

4.3 Methodological quality

Across the included set, summary appraisal scores average 0.68. The good or strong bands at 0.75 and above were reached by eleven studies; the other 31 are adequate. Weaknesses sit behind that average which matter more than the mean does. Objectives, designs and participants are described well. Randomisation and blinding are handled poorly, partly for structural reasons, a training programme not being something one can blind. Control of confounding is the item where performance is genuinely weak and the weakness is not structural: where women enrol voluntarily, whatever disposition brought them to enrol stands as a live alternative explanation for any change that follows, and few studies confront it.

More than any other feature, one thing constrains this evidence base. Measurement simply stopped when the programme did in 32 of the 42 studies. Place that next to how commonly the trainer administered the self-report, and much of this literature turns out to consist of what participants said about themselves, to whoever had just trained them, on the day the training ended.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

4.4 Findings by dimension of empowerment

Down the dimensions the reported effects decline in order, and that ordering is the review's central quantitative result. Table 4 sets it out.

Table 4. Reported findings by dimension of empowerment

Dimension	Studies reporting	Positive	No change	Mean g	Character of the finding
Psychological	41 of 42	39	2	+0.63	Consistent across designs, settings and providers
Social	29 of 42	24	3	+0.46	Moderate; leans on qualitative accounts
Educational	16 of 42	11	5	+0.31	Concentrated on participation, not attainment
Civic	14 of 42	10	4	+0.28	Too few studies for a firm statement
Economic	17 of 42	8	7	+0.20	Small; conditional on embedding

Note. Two studies reported a negative social effect and two a negative economic effect; these are included in the row totals. Effects expressed as Hedges' g.

This literature concentrates on psychological empowerment, and that is where its evidence is strongest. Self-efficacy, self-esteem, assertiveness or a near relative of these shows gains in 39 of the 42 studies. Narrow the set to the strongest studies, or to those with independent assessors, or to those on validated instruments, and the finding holds in every case, with the magnitude dropping each time. Measurement is overwhelmingly by self-report, which makes the mean of +0.63 an upper bound rather than an estimate.

Economic empowerment occupies the far end. Under half the included studies measured an economic outcome at all, and of those that did, seven of 17 found no detectable change. The gap between how strongly these programmes are claimed to work and how strongly the evidence supports the claim is at its widest here. That is not evidence of economic failure. It is closer to evidence that nobody has asked the question properly, an economic outcome being measurable only if participants are followed long enough for one to happen.

Two things thin out together as the eye moves down Table 4. Evidence grows scarcer as effects grow smaller, so some part of the gradient is a record of how seldom the outer dimensions were



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

measured rather than of how little occurred in them. Separating the two fully is beyond the available studies, and no conclusion below assumes otherwise.

4.5 Comparative analysis

Five comparisons went into the protocol before anyone looked at the data. Duration gave the clearest pattern; timing of measurement gave the most consequential one. Table 5 reports both.

Table 5. Mean effect by programme duration and by timing of measurement

Comparison	Studies	Psychological	Social	Economic	Reporting economic outcome
15 contact hours or fewer	14	+0.53	+0.33	+0.03	5 of 14
16 to 30 hours	12	+0.65	+0.52	+0.36	6 of 12
31 to 60 hours	10	+0.81	+0.44	+0.18	4 of 10
More than 60 hours	6	+0.49	+0.67	+0.19	2 of 6
Immediate post-test only	32	+0.68	+0.51	+0.24	13 of 32
With follow-up measurement	10	+0.47	+0.31	+0.08	4 of 10

Note. Duration bands and measurement timing were specified in the protocol before analysis.

Effects reported from longer programmes are larger, and the association grows stronger the further a dimension sits from the individual. Among the shortest programmes an economic effect is almost never reported. Two readings fit the pattern and nothing in the evidence decides between them. Under the first, dose matters: change follows contact, and the distant dimensions want more contact than the near ones. Under the second, selection does: the organisations able to run long programmes tend also to be better resourced, better connected to employers, and likelier to provide the follow-on support that this review identifies below as the condition for economic outcomes.

Follow-up studies report smaller effects right across the dimensions. The obvious reading is attenuation, and the qualitative accounts fit it, participants describing a spell of confidence after the programme that held where circumstances supported it and faded where they did not. That reading cannot stand on the comparison by itself, however. Only 10 of the 42 studies measure twice, and the ten are unrepresentative: better resourced, better designed, likelier to have brought in an independent evaluator, and better design yields smaller effects for reasons that have nothing to do with how much time has passed. Stated honestly, the position is this. Smaller



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

effects under follow-up are consistent with attenuation; attenuation is plausible; and nothing in this evidence base tells attenuation apart from design quality.

4.6 Analytical themes

Out of the descriptive coding came five analytical themes. Table 6 states each one, how many studies contributed to it, and what confidence it was given.

Table 6. Analytical themes and confidence in each

No.	Theme	Studies	Confidence	Principal bound
T1	Internal change is the most consistent and best-evidenced effect	39	High	Magnitude uncertain; self-report, often trainer-administered
T2	Transfer to external outcomes is conditional, not automatic	26	Moderate	Conditions inferred from study characteristics, not measured
T3	Leadership self-perception outruns leadership attainment	14	Moderate	Small evidence base; short follow-up
T4	Household and institutional response operates as a gatekeeper	11	Low to moderate	Qualitative only; never quantified
T5	Durability of effect is largely unmeasured	42	High	A finding about absence, not about effect

Note. Confidence assessed using an approach adapted from GRADE-CERQual (Lewin et al., 2018).

A conditional model rather than a linear one follows from the second and fourth themes read together, and Figure 2 sets it out. External outcomes need internal change, which the programmes produce reliably, and needing it is not the same as having enough of it. Conversion depends on conditions mostly outside the programme's control, and where those are missing the sequence stops there, leaving the participant with a capability and no occasion to exercise it.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

Figure 2. Conditional transfer from internal change to external outcome

Programme inputs	→	Internal change	→	Enacted change	→	External outcome
Content, duration, delivery, facilitator		Self-efficacy, self-esteem, assertiveness		Speaking, negotiating, seeking, applying		Employment, enterprise, role, attainment
↑ Measurement in this literature concentrates on these two stages ↑						
Conditioning factors (Theme 4): family and spousal response; institutional openness; local labour market; childcare and mobility; continuity of support after the programme						
Where the conditioning factors are absent the sequence terminates at internal or enacted change: the capability is expanded but the functioning is not achieved						

Note. Model constructed by the authors from Themes T2 and T4.

Nothing in the coding framework anticipated the fourth theme; it surfaced inductively while data were being extracted. Almost every participant who describes change that lasted also describes someone who made room for it: a husband or father willing to accept the new behaviour, a mother-in-law who took the childcare, a teacher who called on her in class, a supervisor who handed her a task she would not have volunteered for. Among those describing change that faded the picture reverses, or there was simply never an occasion to use what had been learned. Between the two sets of accounts, programme content hardly differs. What differs is what was waiting when the participant got home.

5. Discussion

5.1 Principal findings

Firm ground exists in this review at exactly one point, and qualification covers nearly all the rest. Under personality development programmes, how women read their own capability changes; it changes reliably; and it does so across every design, setting and provider category represented here. That is what theory predicted (Bandura, 1977) and what the evidence delivers, and no part of the review stands on firmer footing.

Programme documentation often takes a further step that does not follow. Internal change founds empowerment, the reasoning runs, so showing internal change shows empowerment. Three things point the other way at once: the gradient in Table 4, the way economic and civic outcomes cluster around the longer and embedded programmes, and the gatekeeper accounts in the qualitative studies. Transfer is conditional. Evaluate a programme on internal measures alone and no size of gain will let it demonstrate what is claimed for it.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

5.2 Relation to prior work

A distinction drawn in the conceptual literature for thirty years, with little systematic evidence ever attached to it, is matched closely by these findings, and the pattern is anticipated exactly by Kabeer's (1999) separation of resources, agency and achievements, and the gap this literature demonstrates at scale is named by Sen's (1999) distinction between capability and achieved functioning. So read, the review does not contradict the theoretical argument; it fills it with empirical content.

The critical strand gains support too. Empowerment has been emptied of its structural content, Batliwala (2007) charges, and in evidence that charge looks like the present finding, where an intervention reaching only the participant leaves everything around her as it was. From the other direction, Cornwall's (2016) conclusion that what works alters relationships and structures rather than dispositions restates the gatekeeper theme. On leadership, Ely et al. (2011) hold that programmes site the deficit in the woman while the constraint sits in organisational practice, and Theme T3 is precisely that in empirical form: many more women who believe they could lead, not many more women leading.

Brody et al. (2015) offer the most instructive comparison, their review running in the opposite direction to this one. Programmes that start from economic participation and sometimes yield psychological change are what they examine; programmes that start from psychological development and seldom yield economic change are what this review examines. Placed together, the pair suggest the more reliable path may run from economic participation to psychological empowerment. That inference belongs to neither review alone and is put forward as a hypothesis. The experimental record is not uncomfortable with it: measurable economic and social effects followed the multifaceted programmes Bandiera et al. (2020) evaluated, which paired vocational and life skills training with sustained delivery, while the training-only components reviewed by Buvinić and Furst-Nichols (2016) largely did not.

5.3 Implications for practice and policy

Below are four implications. They are put as implications and not as instructions, the evidence behind them being of uneven strength.

- Where duration is settled, it belongs with the design choices and not with the budget line: more contact hours bring larger effects, and the climb is steepest on the dimensions that sit furthest from the individual.
- Once the sessions finish, something should still be there for the participant to attach to, whether a placement service, a credit group, an employer relationship or a community organisation. It is around arrangements of that sort that the studies reporting external outcomes gather.
- Household and institutional response belongs inside the intervention rather than outside it as context. The theme supporting this is the least well evidenced of the five, so piloting and evaluation should come before wholesale adoption.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

- What is measured should set the calibration of the claim. A programme that measures confidence should report a change in confidence, not a change in empowerment.

For funders, nothing else on offer matches the value of specifying the evaluation instead of merely requiring one. Two conditions, written in for a single funding cycle, would change what this literature holds: an outcome measured no sooner than six months after the programme closes, and an outcome that is objectively verifiable rather than self-reported. Against the cost of running the programme the addition is small, and the present review shows what has come of leaving it out.

At low cost, three changes on the research side would lift the quality of this evidence base. The first is to use validated instruments wherever they exist, since author-constructed confidence questionnaires yield systematically larger effects and offer little that the free, widely translated alternatives do not, and a fifth of this literature already uses those. The second is to hand post-test administration to someone who did not deliver the training, since a measure taken by the trainer is compromised beyond what any later analysis can repair. The third is to instrument the conditioning variables of Theme T4 next to the psychological scales; household decision-making and spousal support are already measured in the demography literature, and the adaptation needed is slight.

6. Conclusion

There is a narrow sense in which these programmes work, it is a real one, and it comes to this: what women believe about their own capability shifts, the shift arrives reliably, and participants set store by it and put it in concrete terms. Silent at a family meeting before, a woman who now speaks there has crossed a line that means something to her, and scepticism about self-report scales, however well founded methodologically, leaves that fact standing.

An account has grown up around these programmes which the evidence does not bear out. Confidence is missing, the programme supplies it, empowerment follows. Read that way, a woman still standing where she started took in too little of what she was taught. The evidence gathered here points in another direction. She may have taken it in entirely, and then gone back to arrangements with no use for it. Siting the deficit decides who carries the burden, and for two decades the care in this literature has gone to the participant and almost none of it to what surrounds her.

What follows is a modest and specific agenda. Take a second measurement six months after the programme ends. Measure the household and institutional conditions as well as the psychological ones. Report an outcome that someone else can verify. None of it waits on a methodological advance, and until it is done the field settles neither that these programmes empower women nor that they fail to.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

References

1. Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., & Sulaiman, M. (2020). Women's empowerment in action: Evidence from a randomized control trial in Africa. *American Economic Journal: Applied Economics*, 12(1), 210–259.
2. Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215.
3. Batliwala, S. (2007). Taking the power out of empowerment: An experiential account. *Development in Practice*, 17(4–5), 557–565.
4. Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). Sage.
5. Brody, C., De Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., & Dworkin, S. L. (2015). Economic self-help group programs for improving women's empowerment: A systematic review. *Campbell Systematic Reviews*, 11(1), 1–182.
6. Buvinić, M., & Furst-Nichols, R. (2016). Promoting women's economic empowerment: What works? *The World Bank Research Observer*, 31(1), 59–101.
7. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
8. Cornwall, A. (2016). Women's empowerment: What works? *Journal of International Development*, 28(3), 342–359.
9. Ely, R. J., Ibarra, H., & Kolb, D. M. (2011). Taking gender into account: Theory and design for women's leadership development programs. *Academy of Management Learning & Education*, 10(3), 474–493.
10. Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291.
11. Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Development and Change*, 30(3), 435–464.
12. Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields* (HTA Initiative No. 13). Alberta Heritage Foundation for Medical Research.
13. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
14. Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., Bohren, M. A., Tunçalp, Ö., Colvin, C. J., Garside, R., Carlsen, B., Langlois, E. V., & Noyes, J. (2018). Applying GRADE-CERQual to qualitative evidence synthesis findings: Introduction to the series. *Implementation Science*, 13(Suppl. 1), 2.



Kavya Setu

A Multidisciplinary Open Access, Peer-Reviewed Refereed Journal

Impact Factor: 7.2

ISSN No: 3049-4176

15. Nasheeda, A., Abdullah, H. B., Krauss, S. E., & Ahmed, N. B. (2019). A narrative systematic review of life skills education: Effectiveness, research gaps and priorities. *International Journal of Adolescence and Youth*, 24(3), 362–379.
16. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
17. Sen, A. (1999). *Development as freedom*. Oxford University Press.
18. Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45.