



# planning, monitoring and evaluation

Department:  
Planning, Monitoring and Evaluation  
**REPUBLIC OF SOUTH AFRICA**

## Report on the Assessment of Government Evaluations

Evaluation Title:

Measuring Success in Human Settlements Development:  
An Impact Evaluation Study of the Upgrading of Informal  
Settlements Programme in selected projects in South  
Africa

Evaluation Number:

423

Evaluation Completion Date:

31 May 2011

Period of Evaluation:

Financial Year 2009/10-2011/12

Submitted:

20 March 2015 by Mike Leslie

Approved:

20 March 2015 by Mike Leslie

## Evaluation Details

Evaluation Title:	Measuring Success in Human Settlements Development: An Impact Evaluation Study of the Upgrading of Informal Settlements Programme in selected projects in South Africa
Evaluation Number:	423
Evaluation Completion Date:	31 May 2011
Created:	19 May 2014 by Mike Leslie
Submitted:	20 March 2015 by Mike Leslie
Approved:	20 March 2015 by Mike Leslie
Period of Evaluation:	Financial Year 2009/10-2011/12
Known Cost:	
Known Cost Type:	No Data
Initiated By:	National Department of Human Settlements
Initiated By Internal:	Yes
Undertaken By:	National Department of Human Settlements, Development Impact Evaluation (DIME)- World Bank and Vari Consulting
Undertaken By Internal:	No

## Assessors

Mike Leslie	mikel@pdg.co.za
-------------	-----------------

## Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
UISP Impact Evaluation Cover Page.pdf	Evaluation report	Mike Leslie	19 May 2014
South Africa UISP Impact Evaluation Concept_Note.doc	Terms of Reference (ToR) for the evaluation	Mike Leslie	07 January 2015
research protocol 15 feb 2010[1].doc	Any other relevant documentation pertaining to the evaluation process	Mike Leslie	07 January 2015
ADVERTISED TOR ..REVISED RSA Housing IE data collection - FINAL TOR08 09 2009.doc	Terms of Reference (ToR) for the evaluation	Mike Leslie	07 January 2015
UISP Impact Eval Results Presentation 2011.pdf	Presentations of evaluation findings and recommendations	Mike Leslie	07 January 2015
Free State HH questionnaire.pdf	Any other relevant documentation pertaining to the evaluation process	Mike Leslie	07 January 2015
LIMPOPO HH questionnaire.pdf	Any other relevant documentation pertaining to the evaluation process	Mike Leslie	07 January 2015
MOU.pdf	A letter of engagement or contractual agreement with the service provider	Mike Leslie	18 February 2015

Research Protocol - Final.pdf	Any other relevant documentation pertaining to the evaluation process	Mike Leslie	18 February 2015
Assessment Report at Moderation.pdf	Assessment Report at Moderation	Nishendra Moodley	17 March 2015

## Quality Assessment Summary

The evaluation received a score of 3.61 out of 5 on the quality assessment scale indicating an evaluation that is more than adequate. Overall, this evaluation serves as an innovative example of an impact evaluation in the human settlements sector conducted by an internal evaluation team with the assistance of external expertise.

The planning & design phase of the evaluation was the strongest as considerable time and consultation was put into the development of a concept note that became the internal evaluation proposal and was particularly detailed and well-thought through as it pertained to the evaluation design and sampling. Further, the planned use of the exercise for capacity building within DHS stood out and contributed to the phase score of 4.11.

In the course of the evaluation the biggest threats to the process stemmed from issues with the survey company, described as 'out of their depth', collecting less data than originally intended. Despite this, observation of ethical protocols and strong consultation on account of the internal nature of the evaluation contributed to a phase score of 3.36.

The report itself scored highly due to the accessible manner in which it was written with a clear explanation of the analysis and ways of interpreting the data, including acknowledging limitations. This phase scored 3.96 and serves as a good example of the kind of evaluation report that is both technically rich yet can be comprehended by those without a background in statistics.

The final phase of follow-up, use and learning marks a low point with a score of 2.65. The lack of accessibility of the document, unfamiliarity with the results of the evaluation by a key stakeholder, and its tangential contributions to improvements and changes did a serious disservice to what was otherwise a good practice evaluation with an exemplary capacity development component (reduced to a score of 3.25 on account of the learning shortcomings of the last phase).

Coordination and alignment stands out as an overarching consideration where the evaluation fared particularly well, in part because of the embedded nature of the DHS staff.

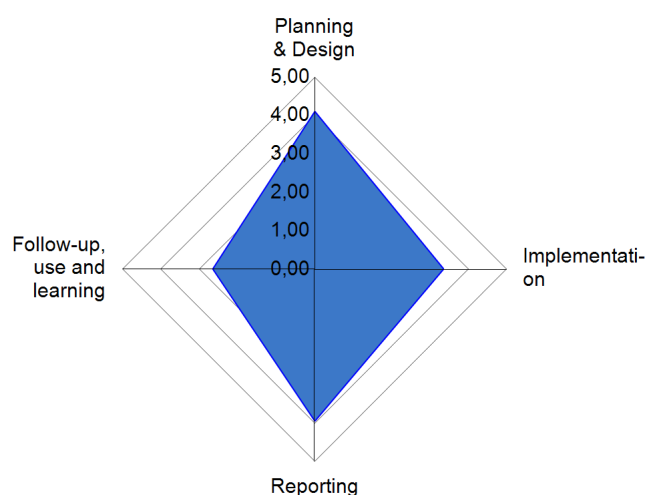
Overall, this was an innovative impact evaluation that displays some best practice characteristics. It would have scored higher had there been better utilisation of the evaluation by DHS or management within the UISP. Nevertheless, it still serves as an evaluation of above adequate quality that is unique for the sector and content area.

## Quality Assessment Scores

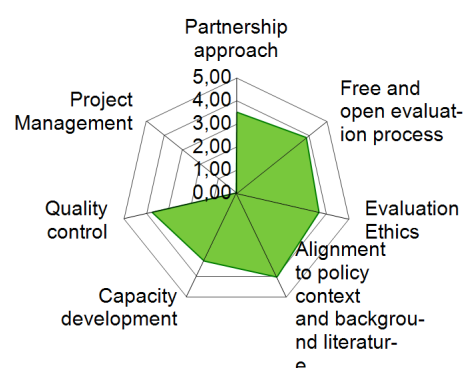
Phase of Evaluation	Score
Planning & Design	4,11
Implementation	3,36
Reporting	3,96
Follow-up, use and learning	2,65
Total	3,61

Overarching Consideration	Score
Partnership approach	3,50
Free and open evaluation process	3,87
Evaluation Ethics	3,67
Alignment to policy context and background literature	4,04
Capacity development	3,25
Quality control	3,76
Project Management	
Total	3,61

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
Planning & Design	Quality of the TOR	4,00
Planning & Design	Adequacy of resourcing	4,09
Planning & Design	Alignment to policy context and background literature	5,00
Planning & Design	Appropriateness of the evaluation design and methodology	4,06
Planning & Design	Project management (Planning phase)	4,00
Implementation	Evaluation ethics and independence	4,00
Implementation	Participation and M&E skills development	3,70
Implementation	Methodological integrity	3,04
Implementation	Project management (Implementation phase)	3,00
Reporting	Completeness of the evaluation report	3,92
Reporting	Accessibility of content	4,73
Reporting	Robustness of findings	4,36
Reporting	Strength of conclusions	3,36
Reporting	Suitability of recommendations	3,77
Reporting	Acknowledgement of ethical considerations	3,31
Reporting	Project management (Reporting phase)	3,00
Follow-up, use and learning	Resource utilisation	2,40
Follow-up, use and learning	Evaluation use	2,71
<b>Total</b>	<b>Total</b>	<b>3,61</b>

## Planning & Design

---

### Quality of the TOR

Standard:	The evaluation was guided by a well-structured and complete TOR or a well-structured and complete internal evaluation proposal (e.g. Background, Purpose, Evaluation Questions, Design & Methodology, Deliverables & Timeframes, Resource requirements, Intended Audience & Utilisation, etc).
Comment and Analysis:	The evaluation was guided by a well-structured and complete concept note which served as an internal evaluation proposal. The concept note included the following sections: Background; Principal Hypotheses and Research Questions; Methodology; Sample Selection, Data and Power Calculations; Impact Evaluation Team and Preliminary Timeline. The document dedicates considerable space to providing a background to informal settlement upgrading and the Upgrading Informal Settlement Programme (UISP) before motivating for the evaluation design and sight selection. The concept note envisions two evaluations: a retrospective impact evaluation which is currently under quality assessment; and a prospective impact evaluation to be conducted in the future including the data collected as part of the retrospective impact evaluation.
Rating:	4: The evaluation was guided by a well-structured and complete TOR or internal evaluation proposal of a good standard
Moderation:	Reconsider
Moderation Comment:	The rating is fine. However, I cannot easily understand the last line of comment.
Approval:	Accepted
Standard:	The purpose of the evaluation stated in the TOR (or an internal evaluation proposal) was clear and e
Comment and Analysis:	The concept note sets out the impact evaluation as one of the projects of the World Bank partnership with the Department of Human Settlements. Specifically, this project is set to "rigorously estimate the impact of the Informal Settlement Upgrading on the welfare of beneficiaries, including the effects on health and safety, employment, consumption and productive activities, and the empowerment of local communities." It distinguishes between the retrospective and prospective evaluation designs and indicates that the evaluation under quality assessment "will identify the causal impact of alternative approaches to informal settlement upgrading."
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted
Standard:	The approach and type of evaluation was suited to the purpose and scope of the evaluation TOR (or an internal evaluation proposal)
Comment and Analysis:	The retrospective impact evaluation design was specifically tailored to the project sites with considerable time and effort put into the identification of projects within the UISP programme that would allow for the kind of retrospective quasi-experimental design providing a counterfactual to support causal claims with a particular emphasis on internal validity. This project exemplifies best practice in terms of the approach and type of the evaluation being suited to the purpose and scope.
Rating:	5: The approach and type of the evaluation was perfectly matched to the purpose and scope of the evaluation TOR and there was a reinforcing complementarity between the sections
Moderation:	Accepted
Approval:	Accepted

Standard:	The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs
Comment and Analysis:	The fact that the evaluation is conducted internally with international assistance is a unique indication of the administration's information needs and implicitly the National Department of Human Settlements (then Housing) and the Provincial Departments are identified as the users of the evaluation and their needs pertaining to determinations of the effects of the current policies and implementation thereof are stated. Specifically, the concept note identifies that this evaluation is but the start of a process that could allow Provincial Departments to "determine whether these complementary [housing] strategies may be put in place to accelerate the rate of housing development and disbursement of funds."
Rating:	4: The TOR identified the intended users of the evaluation and differentiated between their information needs well
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation questions in the TOR (or an internal evaluation proposal) were clearly stated and ap
Comment and Analysis:	There are not explicit evaluation questions set out in the concept note so much as there are clear indications of what 'impacts' the evaluation will seek to determine whether the intervention is the cause of. These include specific human settlements outcomes as they pertain to: Health; Household Economy; Social and Economic Networks; Household Characteristics and Quality; Settlement/Community characteristics; and Satisfaction. Further, these are differentiated by the phase for the type of intervention within the existing UISP programme so that the outcomes are distinguished between in situ upgrading, rollover, relocation, etc.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Standard:	Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation
Comment and Analysis:	The development of the ToR was a collaborative process between the World Bank and the project team at the National Department of Human Settlements. There was an extended period of consultation across national, provincial and local government stakeholders in the identification of the sites for inclusion. This would have been considered excellent, except that it is unclear to what extent other housing institutions were consulted.
Rating:	4: A wider range of stakeholders (i.e. beyond government stakeholders) were meaningfully involved in the scoping of the TOR and choosing the purpose of the evaluation
Moderation:	Reconsider
Moderation Comment:	The comment sounds like it could have just as easily been a 3. This depends on extent to which the NDHS is a stakeholder or client, and how many other key stakeholders there. I suspect more.
Approval:	Accepted

## Adequacy of resourcing

Standard:	The evaluation was adequately resourced in terms of time and budget allocated
Comment and Analysis:	The concept note indicates that the evaluation was originally allocated a 25 month period from April 2008 (Evaluation Design and Site Selection, including the Concept Note) through April 2010 when the Evaluation Report for the retrospective evaluation is expected to be finished. The project was well-resourced in terms of the time available.
Rating:	4: The evaluation was well resourced in terms of the time and budget allocated (i.e. there was some room for flexibility)
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation was adequately resourced in terms of original budget
Comment and Analysis:	The evaluation benefited from the budget and staff allocated for technical assistance from the World Bank, funds set aside specifically for contracting a survey firm (funded by DHS), as well as the staffing contingent of DHS that was included as the project team. The evaluation appeared to be well-resourced in terms of funding as a significant amount of the costs were covered through the partnership arrangement and the inclusion of DHS full-time staff as part of the evaluation team.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	The team conducting the evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The evaluation was very well-resourced in terms of the staffing and skills set. From the side of the World Bank, a team of M&E specialists and economists were appointed to work with senior management in DHS' Monitoring and Evaluation Chief Directorate. The staffing and skills set would have been excellent, had the survey firm, Vari Consulting not been described as lacking and out of their depth.
Rating:	4: The evaluation was well resourced in terms of staffing and skills sets
Moderation:	Accepted
Approval:	Accepted

Standard:	Where appropriate, the evaluation planned to incorporate an element of capacity building of partners
Comment and Analysis:	The evaluation was conducted as part of a technical assistance initiative organised between the World Bank and the South African Government. As such, the entire initiative was considered to be a capacity building exercise for DHS staff to support the growth of a strong impact evaluation practice within the National Department. In terms of planning, this evaluation exemplifies how a rigorous impact evaluation can be planned in partnership with external specialist and delivered in a meaningful, developmental way that grows the skills sets of civil servants and answers important questions about the efficacy of government interventions and policy.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted



## Alignment to policy context and background literature

Standard:	There was evidence that a review of the relevant policy and programme environments had been conducted
Comment and Analysis:	The concept note provides a clear and concise overview of the policy and programme environment as part of the Background section. This demonstrates familiarity with the policy and programme environments and that a thorough review was used in planning the research. Specifically, Breaking New Ground and the provisions pertaining to the 'Informal Settlement Upgrading Program' are briefly analysed and used to motivate the evaluation. It is clear that an extensive review of policy and programme arrangements was considered in planning the evaluation and this is an excellent example of such.
Rating:	5: 5
Moderation:	Reconsider
Moderation Comment:	I am not sure whether you are sufficiently justifying a 5.
Approval:	Accepted

Standard:	There was evidence of a review of appropriate literature having been conducted and used in planning
Comment and Analysis:	There was a clear and concise review of international literature on the topic of informal settlement upgrading and this was used to inform the research, particularly the evaluation design. The Background section of the Concept Note captures this review well and exemplifies best practice both in what is presented and how it is applied to the evaluation design, projects selected and the methodology planned.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

## Appropriateness of the evaluation design and methodology

Standard:	There was explicit reference to the intervention logic or the theory of change of the evaluand in the planning of the evaluation
Comment and Analysis:	There was a reference to the intervention logic of the evaluand as the Concept Note made explicit that it did not explore the process and fidelity to the theory of change so much as it sought to describe and attribute the outcomes of the intervention. It was made clear that this design specifically considered it, but was not theory-based.
Rating:	3: There was explicit reference to the intervention logic or the theory of change of the evaluand in the TOR or the Inception Report
Moderation:	Accepted
Approval:	Accepted

Standard:	Key stakeholders were consulted on the design and methodology of the evaluation
Comment and Analysis:	The evaluation was jointly planned between the World Bank team and staff from DHS. These staff were in regular contact with provincial and local government counterparts as part of the planning of the design and methodology to identify the best sites for the kind of evaluation design envisioned and they therefore were therefore well consulted on the evaluation design and methods.
Rating:	4: 4
Moderation:	Reconsider
Moderation Comment:	Same as 1.1.6
Approval:	Accepted

Standard:	The planned methodology was appropriate to the questions being asked
Comment and Analysis:	The planned methodology was well considered and certainly appropriate to meeting the objectives of the evaluation. Given practical considerations, one might consider this for best practice had the questions themselves been made explicit, rather than implicit to the purpose and objectives of the study.
Rating:	4: The planned methodology was well suited to the questions being asked and considered the data available
Moderation:	Reconsider
Moderation Comment:	if implicit, should this be a 4.
Approval:	Accepted

Standard:	The sampling planned was appropriate and adequate given the focus and purpose of evaluation
Comment and Analysis:	The planned sampling was tailored to the specific projects and the different sites as examples of variations of the intervention/control groups. This Concept Note provides one of the few examples of where the sampling size per project, and differentiated per site, with clear motivation, is provided at the inception stage. It exemplifies best practice as an internal evaluation proposal.
Rating:	5: The sampling planned was creative and ideal for the focus, purpose and context of the evaluation
Moderation:	Accepted
Approval:	Accepted

Standard:	There was a planned process for using the findings of the evaluation
Comment and Analysis:	The evaluation under quality assessment was meant to be the first phase in a broader process that would culminate with the conclusion of another prospective evaluation using data obtained. There was therefore a clear intention to use the data, although the process for using the findings and results of the evaluation was not necessarily as clear.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

## Project management (Planning phase)

Standard:	The inception phase was used to develop a common agreement on how the evaluation would be implemented
Comment and Analysis:	According to the project plan in the Concept Note, as corroborated by interviews, a period of five months was set aside just to allow for the joint development of the evaluation design between the World Bank and DHS team. In terms of the evaluation team, it is clear this phase was used to an excellent standard in terms of establishing a common agreement for evaluation implementation. However, given that there were later challenges with the survey firm appointed for data collection, the 'inception' element of the data collection appointment may not have been as well executed, resulting in an overall rating of 'good'.
Rating:	4: The inception phase was used to good effect to achieve a common agreement and understanding of how the evaluation would be implemented
Moderation:	Reconsider
Moderation Comment:	Is 4 not too high, given the later challenges?
Approval:	Accepted

## Implementation

---

### Evaluation ethics and independence

Standard:	Where data was gathered in contexts where ethical sensitivity is high, informed consent, assurances of confidentiality and appropriate clearance were achieved; e.g. through an ethics review board, in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance
Comment and Analysis:	Prior to undertaking data collection, a research protocol was developed which specifically indicates that because of the health related questions and some of the data sought that the research would be approved via the University of Limpopo's ethics clearance board and this did occur. Further, the research protocol document indicates how each subject is assigned a unique identification number to allow for confidentiality and that the records linking the names and IDs to the unique identification numbers were kept in a secure storage facility.
Rating:	5: There was clear evidence that best practice ethical protocols in the sector were observed in all data collection instances including: informed consent agreements; confidentiality; documenting and storing data notes, recordings or transcripts; and ethics review board approvals
Moderation:	Accepted
Approval:	Accepted

Standard:	Where external, the evaluation team was able to work without significant interference and given access to existing data and information sources
Comment and Analysis:	Evaluation team was internal.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation team was impartial and there was no evidence of conflict of interest
Comment and Analysis:	There was an inevitable degree of partiality given the fact that this was conducted internally. However, the presence of international experts from the World Bank, including one embedded colleague who oversaw the entire process locally, provided evidence that the team would have been sufficiently impartial and mitigated the risk of conflict of interest given the circumstances.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

### Participation and M&E skills development

Standard:	Key stakeholders were involved in the evaluation through a formalised mechanism or institutional arrangement
Comment and Analysis:	There was an existing institutional arrangement but this was more ad hoc than it was a scheduled engagement of the Steering Committee on the project.
Rating:	3: Key stakeholders were involved in the evaluation through a formalised mechanism or institutional arrangement (e.g. a steering committee or reference group)
Moderation:	Accepted
Approval:	Accepted

Standard:	Where appropriate, an element of capacity building of partners responsible for the evaluand and evaluators was incorporated into the evaluation process
Comment and Analysis:	This project stands out because 6-7 staff in the Department of Human Settlements held responsibilities for the evaluation from the conception through to the follow-up use and learning. Although there was not necessarily a structured programme for the capacity building element, hands-on assignment and tasks including undertaking the analysis for the project itself provides a good example of how this occurred. However, this capacitation appeared to focus mostly on the M&E unit rather than the UISP programme staff and was not clearly structured.
Rating:	4: Structured capacity building of evaluators and partners responsible for the evaluand was incorporated into the evaluation process
Moderation:	Accepted
Approval:	Accepted

## Methodological integrity

Standard:	The methods employed in the process of the evaluation were consistent with those planned
Comment and Analysis:	Due to some challenges with the service provider responsible for the survey component, not all of the intended data collection methods were applied consistent with those planned, particularly with regards to the collection of the anthropometric data, including faecal tests, and it was reported that the data collected at Chris Hani Extensions 1-3 in Gauteng did not occur as originally envisioned in the Concept Note where Winnie Mandela was identified.
Rating:	2: 2
Moderation:	Reconsider
Moderation Comment:	I agree with the rating, but does this not also raise ethical questions, where it received a 5?
Approval:	Accepted

Standard:	The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Comment and Analysis:	The evaluation relied primarily on surveys at the three sites (and sub-sites) using a household questionnaire structured into 14 modules to address the range of outcome areas identified in the internal evaluation proposal. This was appropriate given the scope of the evaluation.
Rating:	3: The methods employed in the process of the evaluation were consistent with those planned and implemented adequately
Moderation:	Accepted
Approval:	Accepted

Standard:	A pilot of basic data collection instrumentation occurred prior to undertaking data collection and it was used to inform the research process
Comment and Analysis:	The data analysis approach and methods employed are explained well and documented in Section 5. Interpreting Results in the report. This section makes clear the model specifications statistical techniques employed. Regression analysis was employed through four models (two Intention to Treat Models and two Treatment on the Treated models) which both took into consideration control variables in producing indications of impact for the sites in Limpopo and Free State, while the results in the case of Gauteng were descriptive.
Rating:	4: All components of the data collection instrumentation were piloted which led to some improvements in the data collection instrumentation or affirmation of the instruments
Moderation:	Accepted
Approval:	Accepted

Standard:	Data collection was not compromised by fieldwork-level problems or unplanned diversions from original
Comment and Analysis:	Data collection was compromised by fieldwork-level problems as a result of challenges with the survey service provider pertaining to anthropometric data and with regards to how comprehensive the final data collection of the Chris Hani extension was compared to what was planned. Thus, there were problems, but these did not detract too seriously from the original intentions as the evaluation was still able to rely on data collected in Limpopo and Free State sites in an innovative impact evaluation.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Standard:	Data was collected from key stakeholders (e.g. implementers, governance structures, indirectly affected stakeholders) as data sources
Comment and Analysis:	The executive summary captures the key components well. However it lacks a clear structure and doesn't differentiate between the different components of the report, thus providing a summary of the analysis of the report more than an executive summary of the evaluation as a whole.
Rating:	3: Data was collected from key stakeholders (e.g. Implementers, governance structures, indirectly affected stakeholders) as data sources
Moderation:	Accepted
Approval:	Accepted

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	The broader legislative and programme context is presented as part of the literature review while a separate section is dedicated to introducing each of the three sites identified for data collection. This section explains the differences in the nature of the intervention and how this lends itself to the impact evaluation design.
Rating:	5: Beneficiaries were thoroughly and representatively included as the primary source of data amongst multiple sources of data and information (or if based on secondary data, includes data from beneficiaries and beneficiaries consulted on emerging findings and provide meaningful input to recommendations)
Moderation:	Reconsider
Moderation Comment:	A RATING OF 5 is not sufficiently motivated.
Approval:	Accepted

Standard:	Key stakeholders were significantly engaged as part of the methodology
Comment and Analysis:	This evaluation report did not include representatives of key stakeholders as a data source for collection as part of the methodology but instead elected to go directly to citizen level, as the key beneficiaries.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

Standard:	The methodology included engaging beneficiaries appropriately as a key source of data and informatio
Comment and Analysis:	The methodology comprised beneficiaries exclusively and used a thorough 14 module questionnaire of households to provide a measure of outcomes at the different sites. As such, beneficiaries were the key source of data for the evaluation and they were thoroughly engaged in this regard.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

## Project management (Implementation phase)

Standard:	The evaluation was conducted without significant shifts to scheduled project milestones and timefram
Comment and Analysis:	The evaluation saw significant shifts to the timeframes and project milestones for the project. This was particularly the case with the service provider responsible for data collection. However, because the evaluation was conducted internally relying on the availability and scheduling of the DHS staff in addition to their existing responsibilities, this was not considered to be particularly problematic.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Standard:	The steering committee, technical working group and service provider worked together adequately to facilitate achievement of the objectives of the evaluation
Comment and Analysis:	A paragraph on page 7 is dedicated to providing a motivation for the study while only one specific evaluation objective, as compared to a question is presented. While the rationale for this objective is clear, it does not serve as a particularly strong example of where the kind of information (and depth of information) sought is particularly well explained.
Rating:	3: The steering committee, technical working group and service provider worked together adequately to facilitate achievement of the objectives of the evaluation
Moderation:	Accepted
Approval:	Accepted

# Reporting

---

## Completeness of the evaluation report

Standard:	The scope or focus of the evaluation is apparent in the report
Comment and Analysis:	The scope and focus of the evaluation is apparent within the evaluation report. Section 3. Project Sites and Section 4. Methodology make clear exactly what is within the scope of this evaluation and what the focus of the evaluation is.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	A detailed methodology is outlined in the relevant section of the report to the point that a reader
Comment and Analysis:	Sections 4 and 5 provide a detailed methodology and explanation of how to interpret the results based on the statistical models and analysis employed. The sections are clear, detailed and accessible even across a range of experience levels, serving as a best practice example.
Rating:	5: 5
Moderation:	Reconsider
Moderation Comment:	"Explanation" misspelt.
Approval:	Accepted

Standard:	Acknowledgement of limitations of all aspects of the methodology and findings are clearly and succinct
Comment and Analysis:	The robustness checks and caveats section following the findings are an acknowledgement of limitations and aspects of the methodology and findings that precede it. Further, where there are differences in terms of the design of the evaluation at a given site, as in Gauteng, differences are noted implicitly as well.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

Standard:	Key findings are presented in a clear way; they are made distinct from uncertain or speculative find
Comment and Analysis:	Key findings are presented clearly and distinctly and there is not any evidence of unused data in the body of the report. The report does a good job of making clear findings with supporting graphs, visuals and explanations as to what those results mean.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions and recommendations are clear and succinctly articulated
Comment and Analysis:	Section 10 sets out 8 clear recommendations before Section 11 which clearly and succinctly concludes the quality assessment providing a good example of how conclusions and recommendations should be presented.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

## Accessibility of content

Standard:	The final evaluation report is user-friendly, written in accessible language and adequate for publication (e.g. adequate layout and consistent formatting; complete sentences and no widespread grammatical or typographical errors; consistency of style and writing conventions; levels of formality; references complete and consistent with cited references in reference list and vice versa; etc.)
Comment and Analysis:	The data analysis appears to have been very well executed as questionnaire data is systematically presented and analyzed following an explanation of how to interpret it as well as the models used for analysis.
Rating:	4: The final report is well written, accessible to the common reader and ready for publication with only minor spelling, grammar or formatting mistakes
Moderation:	Accepted
Approval:	Accepted

Standard:	The final evaluation report is user-friendly, written in accessible language and its content follows
Comment and Analysis:	The final evaluation report is user-friendly and very accessible given the technical nature of the design and the analyses employed. The document does not assume that readers are necessarily well versed in impact evaluation methods and analysis and takes time to carefully and simply explain these things while following a clear logic for the report which serves as an example of best practice.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	Figures, tables and appropriate conventions are used in presentation of data (e.g. use of appropriate statistical language; reporting of p-values where appropriate; not reporting statistically insignificant findings as significant; clarifying disaggregation categories in constructing percentages; not using quantitative language in reporting qualitative data, etc.) and are readily discernible to a reader familiar with data presentation conventions
Comment and Analysis:	The report presents all of the data relevant to the purpose and objective of the assessment and analyses it in some depth for both the Free State and Limpopo sites. Limited analysis and findings are presented from the Gauteng site on account of the intention to use the data in future evaluations.
Rating:	4: Figures, tables and conventions are well used for a variety of types of data presentations and supporting explanations make them accessible to readers
Moderation:	Accepted
Approval:	Accepted



Standard:	Quality of writing and presentation is adequate for publication including: adequate layout and consi
Comment and Analysis:	Quality of writing and presentation is adequate for publication and the document presents a good example of how a report should be written and presented for publication.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	Appropriate conventions are used in presentation of data (e.g. use of appropriate statistical langua
Comment and Analysis:	Appropriate conventions are used in the presentation of data and care is taken to make sure that only appropriate statistical language is used, with explanations of the statistical conventions to assist readers unfamiliar with the type of evaluation design and analysis. This is an excellent example of considered and appropriate use of conventions in presenting data.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

Standard:	The use of figures and tables is such that it supports communication and comprehension of results; a
Comment and Analysis:	The evaluation makes excellent use of graphs and tables to communicate the findings and results as these are well supported beginning with explanations of the methodology through to the robustness checks and caveats. An explanation of the different models and explanation of how impact can be interpreted for binary and continuous variables goes further to assist all readers in comprehending and marks this as an excellent example of this standard.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted

## Robustness of findings

Standard:	Data analysis appears to have been executed to an adequate standard
Comment and Analysis:	The report appears free of significant methodological and analytical flaws. It is clear to emphasize when to take care in interpreting data, it qualifies findings as appropriate and dedicates a section on robustness checks and caveats to address limitations. Further, feedback from the interviewees indicated that the report was scrutinized by senior World Bank experts and one can confidently state that it is free of any significant methodological or analytic flaws.
Rating:	5: Data analysis is thorough and well executed to an exceptional standard for all datasets
Moderation:	Accepted
Approval:	Accepted

Standard:	Findings are supported by available evidence
Comment and Analysis:	All of the findings presented are supported by data evidence that, in most instances, is based on a comparison of control versus treatment group data, which was subjected to four types of modelling. Since the findings are site based, one can be confident that all findings included are supported by the available evidence.
Rating:	5: 5
Moderation:	Accepted
Approval:	Accepted
Standard:	Findings are supported by evidence which is sufficiently and appropriately analysed to support the argument, integrating sources of data
Comment and Analysis:	The conclusions section focuses on the results of the evaluation and because of how concise the section is, does not appear to consider other empirical or analytic work beyond an implicit recognition set out in the literature review. However, the recommendations section does appear to consider and draw back to the literature to take into account other empirical and analytic work, particularly as it pertains to policy considerations around backyarders.
Rating:	3: The evidence gathered is analysed to support the argument to an adequate standard and integrates sources of data
Moderation:	Accepted
Approval:	Accepted
Standard:	There is appropriate recognition and exploration of the possibility of alternative interpretations
Comment and Analysis:	The conclusions do address the original evaluation purpose although this and the original objective of the evaluation, to "rigorously measure the impact of the UISP on local communities across a broad range of indicators" are addressed generally for the Free State and Limpopo sites.
Rating:	3: There is appropriate recognition of the possibility of alternative interpretations
Moderation:	Accepted
Approval:	Accepted
Standard:	The report appears free of significant methodological and analytic flaws
Comment and Analysis:	As the evaluation explained at the outset that it sought to measure intended and unintended impacts of the UISP regardless of fidelity to the programme theory this is standard is considered not applicable in this instance.
Rating:	: N/A
Moderation:	Accepted
Approval:	Accepted
Standard:	There is appropriate recognition of the possibility of alternative interpretations
Comment and Analysis:	The use of the four different modelling approaches provides for different ways of analyzing the data and interpreting it. A section that specifically explains how these models were used and applied in presenting the findings, as well as a later section that acknowledges some caveats to the analysis, is a good example of this. Emphasis as to areas where care must be taken in interpreting the findings also provide an indication of the possibility of alternative interpretations or evidence.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	Limitations of all aspects of the methodology and findings are clearly articulated (e.g. limitations of scope or evaluation design, recommendation for additional research, data collection challenges, etc)
Comment and Analysis:	As the report went through a series of drafts that included consultations and engagement with experts in the World Bank, as well as those experts associated with the NUSP programme, there is evidence that sectoral partners and experts were well consulted in the development of recommendations.
Rating:	4: Limitations of all aspects of the methodology and findings are clearly articulated and distinguish between different kinds of limitations
Moderation:	Accepted
Approval:	Accepted

## Strength of conclusions

Standard:	Conclusions are derived from evidence
Comment and Analysis:	DHS staff and programme officials were embedded within the evaluation team and other departmental stakeholders were consulted throughout the process providing strong evidence of the involvement of relevant officials and stakeholders. However, either through scheduling availability challenges or issues of ownership of the evaluation within the department more broadly, this seems to fall short of excellence in the shaping of recommendations.
Rating:	4: Conclusions are derived from evidence and well supported by multiple sources of data that has been well analysed
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions are derived from evidence
Comment and Analysis:	Conclusions are clearly derived from evidence and make concise references back to the key findings around programme impacts as presented and analyzed per the provincial sites of assessment. Conclusions are clear and to the point, albeit very concise.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions address the original evaluation purpose and questions
Comment and Analysis:	The recommendations are jointly targeted at the implementers and policy makers, which is a bit broad. They could have been more specifically targeted as comments such as "Incremental upgrading should take special care to..." and "the UISP can consider how best to integrate..." are rather broad, rather than identifying who or what structure needs to do these things. Further, recommendations around financial awareness of households and capacity building of implementers is made without any indication given to affordability. Involvement of key stakeholders would imply these are acceptable and feasible recommendations.
Rating:	3: Conclusions adequately address the original evaluation purpose and questions
Moderation:	Accepted
Approval:	Accepted

Standard:	Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	The relevant limitations of the evaluation design are noted up front and it is positioned as the first in a series of impact evaluations that are necessary to provide a credible assessment of the UISP's impact overtime throughout the country, rather than in the specifics sites included as the focus of this study. The section on recommendations also reiterates these limitations and indicates "While the evaluation cannot begin to answer a number of key questions and, in fact, generates more questions than answers, the following points can serve as a guideline for further debate".
Rating:	4: Conclusions are drawn with an explicit reference to, and provide a clear judgement on, the intervention logic or theory of change
Moderation:	Accepted
Approval:	Accepted

## Suitability of recommendations

Standard:	Recommendations are made in consultation with relevant government officials, stakeholders and sectoral experts
Comment and Analysis:	The full report does not document the procedures to ensure confidentiality and secure informed consent. However, given the ethical protocol followed during implementation, this is surely implied, but falls short of the standard of explicit acknowledgement in the report.
Rating:	2: Recommendations are made with indirect or partial consultation of government officials, stakeholders and sectoral experts
Moderation:	Accepted
Approval:	Accepted

Standard:	Recommendations are useful- they are relevant, specific, feasible, affordable and acceptable
Comment and Analysis:	The results of the evaluation were presented to stakeholders within the national and provincial departments at the conclusion of the evaluation. In addition, lead evaluation staff from DHS traveled to Washington DC to present the findings of the evaluation to the World Bank as well.
Rating:	4: Recommendations are well-formulated for use- they begin to differentiate by user and are relevant to the current policy context, specifically targetted, feasible to implement, affordable and acceptable to key stakeholders
Moderation:	Accepted
Approval:	Accepted

Standard:	Recommendations are relevant to the current policy context
Comment and Analysis:	Eight recommendations are made which "restricts the discussion to results that have clear value to implementers and policy-makers alike". The recommendations are clearly relevant to the policy context around the upgrading of informal settlements and against the backdrop of BNG. One of the best examples of which is recommendation 6 that formalisation of backyarders be pursued as a policy position rather than that of eradication.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

## Acknowledgement of ethical considerations

Standard:	Peer review of the draft evaluation report occurred prior to finalisation of the evaluation report
Comment and Analysis:	Review of the draft evaluation report by staff at the World Bank removed from the evaluation team occurred prior to finalisation of the report, serving as an internal to government equivalent of peer review.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted
Standard:	The full report documents procedures intended to ensure confidentiality and to secure informed consent where necessary (in some cases this is not needed - e.g. evaluation synthesis - in which case N/A should be recorded)
Comment and Analysis:	The evaluation was completed without any acknowledged changed to the project budget, even for the service provider who was reportedly operating at a loss by the time of the final site survey. Keeping in mind that the evaluation team was internal and therefore paid as part of the operational costs of the department, for a project that ran for more than a year longer than envisioned, it does become difficult to judge in terms of an 'agreed budget'. Although the project was delayed this was not reported as resulting in any cost overruns and it is therefore understood as being of an adequate standard.
Rating:	3: The full report documents some procedures intended to ensure confidentiality and to secure informed consent where necessary
Moderation:	Reconsider
Moderation Comment:	there is an argument to treat this as n/a given that it was internally conducted. As such the main costs would be internally borne, and successive delays or extensions, imply that the original 'budget' was not sufficient, as it can only be understood in time.
Approval:	Accepted
Standard:	There are no risks to participants or institutions in disseminating the evaluation report on a public website
Comment and Analysis:	The report is not publicly available via the DHS website and an electronic copy could not be obtained until contact was made with a representative from the World Bank. Thus, while there is availability online, this is not reference or acknowledged by the DHS anywhere, nor is it easily found by undertaking an online search. As such this undermines the evaluation and is marked as falling below this standard in the South African context.
Rating:	2: There are some risks to either participants or institutions in disseminating a summary version of the evaluation report on a public website
Moderation:	Accepted
Approval:	Accepted
Standard:	There are no risks to participants in disseminating the original report on a public website
Comment and Analysis:	There are no apparent risks to participants in disseminating the original report on a public website and the World Bank has done so already, but the report is not currently easily available from DHS.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

Standard:	There are no unfair risks to institutions in disseminating the original report on a public website
Comment and Analysis:	Similar to the above, there are no unfair risks to institutions in disseminating the original report as it has already been made public by the World Bank, but this is not easily accessible from the side of DHS at this time.
Rating:	4: 4
Moderation:	Accepted
Approval:	Accepted

## Project management (Reporting phase)

Standard:	A project closure meeting that reflected on the challenges and strengths of the evaluation process o
Comment and Analysis:	Given the fact that the project was conducted internally and part of an on-going engagement between the World Bank and DHS, there wasn't a project closure meeting per se, as the item was addressed as a matter of on-going operations. That said, there were some informal reflections expressed, particularly as it pertained to data, and the need to identify comparable control and treatment sites in other provinces for other impact evaluations. Given this, the standard can be considered met to an adequate degree even if there was not a specifically identified project closure meeting.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted

## Follow-up, use and learning

---

### Resource utilisation

Standard:	The evaluation was completed within the planned timeframes
Comment and Analysis:	The evaluation took much longer to complete than originally intended. This was due to challenges with the service provider during data collection as well as repeated delays on account of the evaluation team comprising DHS officials having other core responsibilities in addition to this project. Despite the space the project had because of the nature of the evaluation team arrangement, this still marks as a significant extension on the original timeframes set out in the concept note whereby the final baseline report was expected to be completed by May 2010 and it was actually concluded in May 2011.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

Standard:	The evaluation was completed within the planned timeframes and budget
Comment and Analysis:	As one of the first evaluation's domestically, and one of the few internationally, to provide a measure of the household impact of informal settlement upgrading through an innovative comparison of different incremental upgrades at sites in Limpopo and Free State, the evaluation is surely of conceptual value in understanding what the results of this initiative are for beneficiaries in these two sites.
Rating:	4: The evaluation was completed within the planned timeframes and budget and allowed for additional value to be achieved
Moderation:	Accepted
Approval:	Accepted

### Evaluation use

Standard:	Results of the evaluation have been presented to relevant stakeholders
Comment and Analysis:	Although this would require engagement with additional stakeholders to conclusively determine, the lack of access to, reference and use of the report within DHS itself provides an indication that its contribution has been at best residual over the medium to long term, in line with the above.
Rating:	2: Results of the evaluation have been presented to stakeholders involved in the management of the evaluation project only but not more broadly
Moderation:	Accepted
Approval:	Accepted

Standard:	A reflective process has been undertaken by the steering committee (if no steering committee exists)
Comment and Analysis:	There was not a clear reflective process undertaken as this occurred in a rather ad hoc manner, rather than as part of any deliberate reflective exercise. This is particularly unfortunate that the evaluation was initially pitched as the first part in a later impact evaluation that does not appear to have occurred.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

Standard:	A reflective process has been undertaken by the steering committee with the service provider (if no steering committee exists then by the evaluation management team or the involved department officials) to reflect on what could be done to strengthen future evaluations
Comment and Analysis:	As the evaluation team was itself comprised of staff from the DHS there was certainly a strong skills development element amongst the evaluators themselves who through the practical experience of the evaluation and related trainings on SPSS and Excel were capacitated in the process. It is hard to imagine a process where skills development could be better integrated into the evaluation for an internal evaluation team. This is a best practice example.
Rating:	5: A reflective process has been undertaken and reflections on how to strengthen future evaluations have been documented, including identification of scarce skills areas, competences in need of improvement and/or relevant capacity building opportunities
Moderation:	Accepted
Approval:	Accepted
Standard:	The evaluation study is seen by interviewed stakeholders as having added significant symbolic value to the policy or programme (e.g. raised its profile)
Comment and Analysis:	There was not a formal peer review so much as a review by senior specialists in the World Bank on the evaluation design and methodology. While not a true peer review, this was a review outside of the core evaluation project team which provided sufficient quality assurance be considered adequate in light of the standard.
Rating:	3: The evaluation study is seen by interviewed stakeholders as being of symbolic value to the policy or programme
Moderation:	Accepted
Approval:	Accepted
Standard:	The evaluation study is seen by interviewed stakeholders as having added significant symbolic value
Comment and Analysis:	It is debatable the extent to which this evaluation was of symbolic importance as it pertained to the UISP. There is very little evidence to suggest that it alone raised its profile, but it does seem to have acted in concert with other initiatives and efforts which were identifying some of the same issues that needed to be addressed related to the informal settlement upgrading (e.g. NUSP, Outcome 8 reporting, etc). It certainly raised the profile of the informal settlement upgrading programme within the M&E directorate and the staff responsible for monitoring it, but the fact that senior management responsible for UISP did not claim to be in possession of a copy of the report is certainly problematic.
Rating:	3: 3
Moderation:	Accepted
Approval:	Accepted
Standard:	The evaluation study is of conceptual value in understanding what has happened and possibly in shaping future policy and practice
Comment and Analysis:	A pilot of data collection instrumentation was reported to have occurred as per the standard practice for World Bank related evaluations. However, this was not acknowledged or documented within the evaluation report.
Rating:	3: The evaluation study is of conceptual value in understanding what has happened and possibly in shaping policy and practice
Moderation:	Accepted
Approval:	Accepted



Standard:	Development of a draft improvement plan has been started, but not completed, based on the findings a
Comment and Analysis:	There is no evidence that any kind of improvement plan or follow-up related to this evaluation was initiated or undertaken. However, at the close of the last term this evaluation included a suite of assessments which were being considered as informing planning for the new term of government and one can therefore consider that it indirectly, and in concert with other initiatives, has fed into improvement planning relating to informal settlement upgrading.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

Standard:	There is clear evidence of instrumental use - that the recommendations of the evaluation were implem
Comment and Analysis:	There is limited evidence of instrumental use of this evaluation. Within the UISP programme specifically this evaluation has not been considered or engaged with by some senior management despite its potential value and application. However, when considering it amongst the suite of other work completed one could say that some of the recommendations, particularly integrating of economic opportunities into the upgrading process, the approach to backyarders and additional research on informal settlements, has found expression, just not as a result of this evaluation alone.
Rating:	2: 2
Moderation:	Accepted
Approval:	Accepted

## References

---

Department of Human Settlements. 2011. Measuring success in Human Settlements development: An impact evaluation study of the Upgrading Informal Settlements Programme in selected project in South Africa. Accessed from the World Bank website: [www.microdata.worldbank.org/](http://www.microdata.worldbank.org/)

Department of Human Settlements. 2009. Memorandum of Understanding between the Department of Human Settlements and the International Bank for Reconstruction and Development concerning the Impact Evaluation Study on the Upgrading of Informal Settlements Programme. Unpublished, dated 11 August 2009.

Martinez, S. 2010. Research protocol BUILDING A BRIGHTER FUTURE: THE EFFECT OF INFORMAL SETTLEMENT UPGRADING ON HEALTH IN SOUTH AFRICA. Unpublished, version 15 February 2010.

Department of Human Settlements. 2009. Impact Evaluation Concept Note- BUILDING A BRIGHTER FUTURE: THE EFFECT OF INFORMAL SETTLEMENT UPGRADING ON HEALTH IN SOUTH AFRICA. Unpublished, Revised draft dated 1 May 2009.

Department of Human Settlements. 2011. Measuring Success in Human Settlements Development Presentation. DIME Seminar, 7 July 2011.

World Bank. 2011. DIME Brief- Upgrading Informal Settlements in South Africa. Accessed [www.worldbank.org/dime](http://www.worldbank.org/dime)

## List of Interviewees

---

Johan Wallis, Chief Director: Programme Implementation Facilitation (DHS), Telephonic Interview, 22 January 2015.

Mulalo Muthige, Director: Impact Evaluation (DHS), Telephonic Interview, 2 February 2015.

Aidan Coville, World Bank DHS Technical Support, Telephonic Interview, 7 January 2015.