

IOM REGIONAL DATA HUB FOR THE EAST AND HORN OF AFRICA

USING RETROSPECTIVE SURVEY MEASUREMENT IN ASSESSING MIGRANT REINTEGRATION

Evidence from IOM Programmes in Ethiopia, Somalia, and Sudan *

Dr James Dennison,** March 2022

* *The research upon which this article is based was funded by the European Union under the EU-IOM Joint Initiative for Migrant Protection and Reintegration (Horn of Africa).*

** *Researcher at the University of East Anglia and part-time Professor at the European University Institute's Migration Policy Centre.*



This study was produced with the financial support of the European Union. The contents of this report are the sole responsibility of the authors and do not necessarily reflect the views of the International Organization for Migration (IOM), nor those of the European Union. Neither IOM nor the European Union make claims – expressed or implied – on the completeness, accuracy and suitability of the information provided through this report. Names and boundaries do not imply official endorsement or acceptance by the International Organization for Migration (IOM), nor by the European Union.

Publisher: International Organization for Migration
IOM Regional Office for the East and Horn of Africa
Postal Address: P.O. Box 55040-00200, Nairobi, Kenya
Visiting Address: Sri Aurobindo Avenue, Off Mzima Springs,
Lavington, Nairobi, Kenya
Email: Regional Data Hub (RDH) rdhronairobi@iom.int
Website: <https://eastandhornofafrica.iom.int/regional-data-hub>

This publication was issued without formal editing by IOM.

Design: We2 – www.we2.co

Required Citation: International Organization for Migration (IOM), 2023. *Using Retrospective Survey Measurement in Assessing Migrant Reintegration: Evidence from IOM programmes in Ethiopia, Somalia, and Sudan*. IOM. Regional Data Hub for the East and Horn of Africa.

© IOM 2023



Some rights reserved. This work is made available under the [Creative Commons Attribution-NonCommercial-NoDerivs 3.0 IGO License](https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode) (CC BY-NC-ND 3.0 IGO).*

For further specifications please see the [Copyright and Terms of Use](#).

This publication should not be used, published or redistributed for purposes primarily intended for or directed towards commercial advantage or monetary compensation, with the exception of educational purposes, e.g. to be included in textbooks.

Permissions: Requests for commercial use or further rights and licensing should be submitted to publications@iom.int.

* <https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>

TABLE OF CONTENTS

Introduction	1
Usage of Retrospective Survey Measurement	2
Application to Returning Migrant Reintegration Surveys	3
Empirical Testing	4
Conclusion and Recommendations	5

INTRODUCTION

How can practitioners and policymakers effectively measure the impact of migration policy programmes? More specifically, how can *ex ante*¹ baselines be measured in a context of transitory populations, developing countries, and sometimes reactive and hastily assembled policy interventions? This brief theoretically and empirically considers the extent to which retrospective baseline measurement can act as an accurate and, in some circumstances, optimal solution to these challenges. It is based on a research conducted in the context of the “IMPACT” evaluation² of the EU-IOM Joint Initiative for Migrant Protection and Reintegration (Horn of Africa), in which a retrospective baseline measurement was introduced in 2021 for both practical and methodological purposes.

Retrospective survey measurement, also known as “retrospective enumeration”, is characterised by its use of recall survey questions in which an individual is asked to remember and measure a past state - such as their economic, social, or psychological situation - to produce baseline estimates that can be used in the calculation of over-time change. Scientifically, understanding the extent to which retrospective or recall questions can be used to measure baseline scores is of interest because of debates about the extent to which such measures are overly subject to biases regarding memory, consistency, survival, and social desirability.³ More profoundly, it raises questions about human ability to understand one’s own past and world accurately and objectively, or whether such memories and understandings are modified over time to serve cognitive needs and motivations in the present. More broadly, comparing retrospective measurement to other forms of baseline

measurement, such as for example contemporaneous, “prospective”⁴ baseline measurement as part of panel data,⁵ allows us to consider under what conditions each of the approaches represents the best or “least bad” available approach. **The nature of migration means that traditional panel and longitudinal approaches are not always possible or are particularly liable to their own biases such as panel fatigue,⁶ conditioning,⁷ and especially attrition.⁸** Substantively, understanding the actual effects of interventions aimed at improving reintegration outcomes is of overwhelming practical importance for advocacy organisations, governments, communicators, policymakers, and those working in politics who want to know what is likely to be a sustainable, effective, and value-for-money migration policy framework.

- 1 Refers to the baseline being conducted before the event in consideration, for example before the reception of reintegration assistance.
- 2 Further information on the IMPACT study is available on the webpage dedicated to this initiative: <https://eastandhornofafrica.iom.int/impact-study>. Last accessed on 10 February 2021.
- 3 Response bias that influences a participant to choose responses that reflect what they believe is more socially desirable or acceptable rather than their true thoughts and feelings.
- 4 In prospective studies, individuals are followed over time and data about them is collected as their characteristics or circumstances change.
- 5 Panel datasets are developed by asking the same questions to the same individuals across time.
- 6 Refers to diminishing quality of data gathered from the same individual due to fatigue with the data collection efforts, for example if she or he is expected to stay in the panel for too long a duration (i.e. for too many waves) of data collection.
- 7 Repetitive interviewing may introduce panel conditioning, a state in which interviewees change their beliefs or behaviour just by being exposed to and answering a variety of questions over time.
- 8 Attrition occurs when participants leave from a study.

USAGE OF RETROSPECTIVE SURVEY MEASUREMENT

Retrospective survey measurement has been used in assessing the effects of public policy responses to a broad range of social phenomena, such as the COVID-19 pandemic, socio-economic changes, and migration, particularly when there is difficulty in tracking individuals over time. However, despite its practical benefits, it has also been argued to result in potential forms of measurement error, such as (i) susceptibility to survival bias;⁹ (ii) high cognitive demands on respondents, notably in terms of remembering topics that are or were unimportant to them and difficulty in accurately dating life events, reducing confidence in the temporal ordering criterion for causal inference,¹⁰ or negatively affecting the accuracy of responses to more volatile or fast-moving events; and (iii) biased responses to avoid cognitive dissonance with contemporaneous self-assessments, narratives, or self-understandings or to maintain consistency with contemporaneous social norms and values, a form of social desirability bias.

Both high cognitive demands on respondents and avoidance of cognitive dissonance constitute a form of *consistency bias* that would overestimate similarity between the retrospective baseline measurement and the contemporaneous “endline” (or “ex-post” or “post-treatment”) measurement. Aside from this potential for *underestimating* difference, several biases would also exaggerate the *positivity* with which respondents retrospectively assess their past, notably “rosy retrospection”, the related “euphoric recall”, and egocentric bias, whereby, amongst other things, individuals overestimate their own outcomes in hindsight (see also “hindsight bias”¹¹). Alternatively, there may be a *negative* effect of retrospective measurement if the treatment leads individuals to reappraise *ex post* their subjective assessment of their situation *ex ante*, i.e. their standards are changed by the treatment so that they judge their previous situation more negatively than they would have at the time, a form of systematic measurement error whereby the treatment affects the baseline.

However, empirical evidence remains too mixed to rule out retrospective measurement, certainly in comparison to other approaches, with researchers repeatedly finding only a weak and uncertain influence of recollection. That said, other studies have shown that bias increases with more cognitively demanding questions and the passage of time and that researchers should aim to isolate those questions about which people can remember from those that they cannot. Researchers have suggested that accuracy can be improved by focussing on more recent events – with two and five years both given as cut-off points after which accuracy diminishes – those of longer duration, linking questions to other high salience life events (e.g. marriage, having children, return migration), asking multiple members of the same household the same retrospective questions, and having interviewers record the quality of the interview. **Overall, the general conclusions of recent studies are that (i) the current presumption against the use of long-term recall questions in field surveys ignores a potentially rich source of data, and (ii) both prospective and retrospective data have merits and drawbacks so ideally both should be used.**

9 This is also the case for panel data, although attrition in the latter case increases over time rather than the inverse.

10 An estimation strategy that allows researchers to draw causal conclusions based on the data collected.

11 The tendency to overestimate one’s ability to predict the outcome of an event, after learning the actual outcome.

APPLICATION TO RETURNING MIGRANT REINTEGRATION SURVEYS

The EU-IOM Joint Initiative for Migrant Protection and Reintegration (henceforth, the EU-IOM Joint Initiative) was launched in December 2016 with circa €0.5 billion in funding from the EU Emergency Trust Fund for Africa (EUTF)¹², conceiving of reintegration in terms of not just economic indicators but also social and psychosocial ones at the individual, community and structural level across 26 countries. The “IMPACT” study, which starting in 2019 evaluates the EU-IOM Joint Initiative in Ethiopia, Somalia, and Sudan, uses the Reintegration Sustainability Survey (RSS)¹³ and its related individual scoring system to measure reintegration outcomes. Measurements are taken *ex ante* the treatment (a ‘baseline’ RSS interview, conducted within three months after the

migrant has returned in the country of origin) and *ex post* (an ‘endline’ RSS interview, generally conducted at least one year after the migrant has returned, and often after reintegration assistance was provided by the EU-IOM Joint Initiative). Since 2021, the RSS questionnaire used in the context of the IMPACT study was adapted to include both retrospective baseline and contemporaneous baseline questions for both practical and methodological considerations, relating respectively to the COVID-19 pandemic, political instability in the target countries, mounting data collection costs, and to testing the potential utility of retrospective questions in future reintegration programme impact assessments and other monitoring and evaluation activities.

Table 1. Example of contemporaneous/prospective and retrospective questions in the RSS questionnaire used for the IMPACT study

CONTEMPORANEOUS / PROSPECTIVE QUESTION	RETROSPECTIVE QUESTION	RESPONSE OPTIONS (Common to both questions)
How often are you invited or do you participate in social activities (celebrations, weddings, other events) within your community?	If I asked you the same question 3 months after you returned, what would you have responded? Try to think if the frequency of being invited changed or not since then, if this helps you recall.	Very often Often Sometimes Rarely Never I do not wish to answer

The theoretical considerations outlined in the previous section lead to several hypotheses about the effect that retrospective measurement may have on responses to RSS questions. It is plausible that returning migrant respondents asked retrospective reintegration questions are more *positive* in self-appraisal of their initial post-return situation than respondents asked contemporaneously (*consistency* and *rosy retrospective* biases), more *negative* (*systematic measurement error*), or that there is no difference (particularly likely in the case that the baseline period is recent and memorable). It is

also plausible that that difference in RSS reintegration scores between those asked retrospectively and those asked contemporaneously is contingent on one’s memory, which can be self-assessed, of the baseline period. Finally, the difference in migrant RSS reintegration scores between those asked retrospectively and those asked contemporaneously may be contingent on the duration of time between the time of questioning and the time of return and – given the cognitive demands of recall – a lower education level of the respondent.

12 <https://migrationjointinitiative.org/about-eu-iom-joint-initiative> accessed on 10 February 2021.

13 See MEASURE report, Samuel Hall and IOM, Setting Standards for an Integrated Approach to Reintegration, summary report commissioned by IOM and funded by the United Kingdom Department for International Development (Geneva, IOM, 2017).

EMPIRICAL TESTING

Data continues to be collected for the IMPACT study at the time of writing, including a full ‘calibration’ group of non-migrants that did not receive IOM reintegration assistance. Such data will continue to be collected until the end of 2022. However, the 1774 observations collected by the IMPACT study by early 2022, 1095 of whom were measured retrospectively, can already be used to test the above hypotheses. Although this survey included many dozen variables, for the purposes of this article, only ten independent variables and four dependent variables are utilised (the composite RSS reintegration score and the three ‘dimensional’ RSS scores measuring economic, social and psychosocial reintegration outcomes). The independent variables measure sex, age, origin country, education, regularity of current psychological problems, interview type (face-to-face or phone), and self-assessed ease of memory of the baseline period.¹⁴

When we compare the mean RSS scores between those whose baseline was measured contemporaneously and those for whom it was measured retrospectively, we see practically no difference between the two groups. Notably this is the case in terms of the overall RSS score and the three ‘dimensional’ economic, social and psychological scores. However, this may be due to the non-random allocation of measurement type, which can be partially resolved via regression analyses that control for the above independent variables. When we do so, we can see that retrospective (as opposed to contemporaneous) measurement has a statistically significant negative effect of 0.04 on the composite RSS score (the score scale is between 0 and 1), a considerable effect which is larger than, for example, education.¹⁵ However, **the effect of retrospective measurement more than halves and becomes non-statistically significant when the 20 per cent of those measured retrospectively who reported finding it not easy to remember the time-period are removed from the analysis.** This suggests that the negative effect observed

in the full sample is largely a result of memory bias. The same is the case across the three other ‘dimensional’ RSS scores. By contrast, neither days since the baseline nor education level have any statistically significant effect on the effect of retrospective measurement on the comprehensive RSS, suggesting that the observed memory bias does not increase over time or decrease with higher education.

Given the clear importance of self-reported retrospective memory in determining the effect of retrospective measurement on one’s RSS score, it is worth considering further the determinants of self-reported retrospective memory, again using regression analysis. To do so, a dichotomous variable¹⁶ of retrospective memory is produced (0 for those saying it is easy to remember 3 months after return and 1 for those *either* stating that it is difficult to do so or neither difficult nor easy) and the same socio-demographic and measurement control variables as the earlier regressions are included. In the case of this new analysis, notably, none of sex, education, age, or days since the baseline measurement influence one’s memory. Instead, we see that **having declared to be experiencing symptoms associated with common mental disorders at the time of responding, and doing the interview over the phone, rather than face-to-face increase the chance of self-assessed memory difficulties.**

14 “When asked about things related to the past (3 months after return), how easy or difficult was it to remember your situation back then and answer the questions?”

15 However, if standard errors are clustered by country, the negative effect of retrospective measurement is not statistically significant despite being of the same magnitude.

16 Categorical variable with two categories or levels, otherwise also known as ‘binary’ or ‘dummy’ variable.

CONCLUSION AND RECOMMENDATIONS

As such, because retrospective measurement has significant practical and logistical advantages over panel approaches – notably in terms of efficiency, subject as panel approaches are to attrition and other biases that are especially problematic in dealing with transient populations – this article recommends that researchers ideally utilise both forms of baseline measurement in impact evaluations and other monitoring and evaluation initiatives conducted in the context of returning migrant reintegration. When using retrospective measurement, however, it is vital to:

1. Collect data on self-reported ease-of-memory of the time being measured.
2. Test for variation in the results according to self-reported memory.
3. Because phone interviews are found to increase self-reported problems in memory, it is recommended that face-to-face interviews are especially prioritised when using retrospective baseline measurement.

Future research should utilise the IMPACT study's growing body of data, including the future inclusion of a comparison group, to produce genuine randomised control trials from which the effect of retrospective measurement can be further tested. Furthermore, given the seeming centrality of memory bias to problems of retrospective measurement – and the inability of this article to find socio-demographic determinants of it – it is vital to further investigate the causes of variation in this memory variable, particularly regarding those practical issues that researchers may be able to influence and so diminish memory bias.

For further information, please contact the IOM Regional Data Hub for the East and Horn of Africa at rdhronairobi@iom.int, Attn: Davide Bruscoli.

IOM REGIONAL DATA HUB FOR THE EAST AND HORN OF AFRICA

✉ rdhronairobi@iom.int

International Organization for Migration
IOM Regional Office for the East and Horn of Africa
Sri Aurobindo Avenue,
Off Mzima Springs, Lavington, Nairobi, Kenya



■
EU-IOM
Joint Initiative for
Migrant Protection
and Reintegration
■

