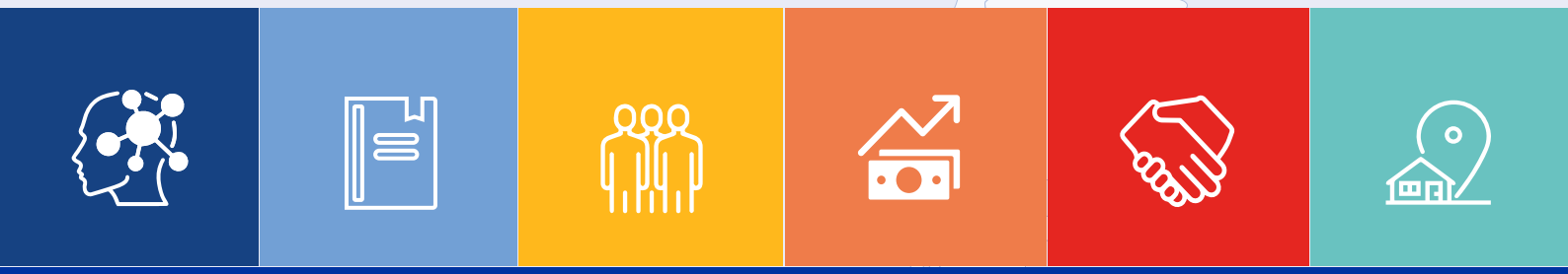
MULTIDIMENSIONAL INTEGRATION MEASUREMENT TOOLKIT:

OFFERING A MULTIDIMENSIONAL APPROACH TO MEASURE MIGRANT INTEGRATION OUTCOMES

A PRACTICAL GUIDE FOR THE APPLICATION OF THE TOOLKIT



IOM
UN MIGRATION

The opinions expressed in the report are those of the authors and do not necessarily reflect the views of the International Organization for Migration (IOM). The designations employed and the presentation of material throughout the report do not imply expression of any opinion whatsoever on the part of IOM concerning the legal status of any country, territory, city or area, or of its authorities, or concerning its frontiers or boundaries.

IOM is committed to the principle that humane and orderly migration benefits migrants and society. As an intergovernmental organization, IOM acts with its partners in the international community to: assist in meeting the operational challenges of migration; advance understanding of migration issues; encourage social and economic development through migration; and uphold the human dignity and well-being of migrants.

This document was produced with the financial assistance of the IOM Development Fund, the Immigration Policy Lab and the Geneva Science-Policy Interface.

For any comments and/or suggestions, please contact the Integration and Migrant Training Unit (intmt@iom.int).

Publisher: International Organization for Migration
17 route des Morillons
1211 Geneva 19
P.O. Box 17
Switzerland
Tel.: +41 22 717 91 11
Fax: +41 22 798 61 50
Email: hq@iom.int
Internet: www.iom.int

This publication was issued without formal editing by IOM.

Required citation: International Organization for Migration (IOM), 2023. *Multidimensional Integration Measurement Toolkit: Offering A Multidimensional Approach to Measure Migrant Integration Outcomes*. IOM, Geneva.

ISBN 978-92-9268-563-8 (PDF)

© IOM 2023

Some rights reserved. This work is made available under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 IGO License (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

PUB2023/041/R*

MULTIDIMENSIONAL INTEGRATION MEASUREMENT TOOLKIT:

OFFERING A MULTIDIMENSIONAL APPROACH TO MEASURE MIGRANT INTEGRATION OUTCOMES

A PRACTICAL GUIDE FOR THE APPLICATION OF THE TOOLKIT

DISC INITIATIVE
The Joint Global Initiative on Diversity, Inclusion and Social Cohesion

IOM UN MIGRATION | IOM DEVELOPMENT FUND

ipl immigration policy lab
Stanford | ETH Zurich

GENEVA SCIENCE-POLICY INTERFACE

ACKNOWLEDGEMENTS

As migrants settle into their host communities, governments and organizations working on the ground have been challenged with devising plans to integrate them into society. At the same time, there is limited knowledge on the broader integration experience of migrants. This highlights the need for empirical tools that measure their assets, knowledge, and capacities to establish themselves in the various spheres of community life from a multidimensional understanding of integration. In line with the 2030 Agenda for Sustainable Development's overarching principle of "leaving no one behind" and the Global Compact for Safe, Orderly and Regular Migration, Objective 1 "C*ollect and utilize accurate and disaggregated data as a basis for evidence-based policies*", the development and implementation of mechanisms to capture data in integration and migration management is of paramount importance to inform policy and practice.

Therefore, this Practical Guide presents the Multidimensional Integration Measurement Toolkit, which aims to develop the capacities of practitioners to systematically assess beneficiaries and enhance understanding of integration by showcasing how the toolkit was built, piloted, and can be used in different contexts. Through this Practical Guide on measuring migrant integration, IOM makes strides towards achieving the Global Compact for Migration Objective 16: "*Empower migrants and societies to realize full inclusion and social cohesion*" as well as the institutional objectives outlined in the Migration Governance Framework (MiGOF) that ensure migration and related policies are formulated using evidence and whole-of-government approaches (Principle 2) and in advancing the socioeconomic well-being of migrants and society (Objective 1). That being said, the toolkit's versatility and multidimensional lens can be leveraged beyond integration programmeming to include other programmatic contexts that cater to various migrant populations.

I encourage IOM missions and partner organizations to benefit from this guide as it serves as a valuable resource for those interested in capturing the level of integration of migrant beneficiaries. The document provides concrete guidance for applying the toolkit and key recommendations on how it can be used to gather robust data that are comparable across diverse contexts.

This guide would not have been possible without the tireless efforts of IOM's Joint Global Initiative on Diversity, Inclusion and Social Cohesion (DISC) team comprised of Jobst Koehler (Senior Integration and Migrant Training Specialist), Ace Dela Cruz (Programme Coordinator), and Morgane Reina (Project Assistant) from the Integration and Migrant Training Unit, Niklas Harder (Consultant/Academic), and the contributions from the Immigration Policy Lab (ETH Zurich/Stanford University). The development of this guide also relied heavily on the support and inputs from various IOM colleagues from different divisions, units, country offices, and missions, particularly IOM Peru, IOM Brazil, and IOM Dominican Republic where the toolkit has been piloted. I also gratefully acknowledge the valuable input of our colleagues of the IOM Regional Office in Buenos Aires and the Office of the Inspector General.

Through a systematic measurement of migrant integration, we can improve IOM's direct assistance services to migrants, as well as shape policy discussions with both government and diverse stakeholders to work towards long-term, effective integration prospects for all migrant populations.

Monica Goracci
Director of Programme Support and Migration Management
IOM

TABLE OF CONTENTS

ACKNOWLEDGEMENTS III

EXECUTIVE SUMMARY VII

1. INTRODUCTION 1
   1.1. Understanding and measuring migrant integration outcomes 2
   1.2. Developing a toolkit to measure integration outcomes 2
   1.3. Overview and structure of the practical guide 3

2. MULTIDIMENSIONAL NTEGRATION MEASUREMENT TOOLKIT 5
   2.1. A core survey instrument: the IPL integration index 6
   2.2. Structure and components of the multidimensional integration measurement toolkit 8
      2.2.1. Introduction and consent 9
      2.2.2. Demographic questions 10
      2.2.3. IPL integration index: survey questions 11
      2.2.4. Additional and targeted questions 12
   2.3. Scoring system 13
      2.3.1. Basics of the scoring rule 13
      2.3.2. Detailed scoring 13
      2.3.3. Rescaling 15

3. APPLICATION AND ADAPTATION OF MULTIDIMENSIONAL INTEGRATION MEASUREMENT TOOLKIT 17
   3.1. A systematic approach to the application of the toolkit 18
   3.2. Design of the study: defining the purpose and measurement context 18
      3.2.1. Basic concepts in study design 19
      3.2.2. Basic survey designs 21
      3.2.3. Basic evaluation designs 23
   3.3. Local adaptation of the index 26
      3.3.1. IPL-12 versus IPL-24 26
      3.3.2. Standard adaptations 26
      3.3.3. Dropping or adding questions 29
   3.4. Pilot testing and data collection 30
      3.4.1. Hiring of a survey firm: roles and responsibilities 30
      3.4.2. Training of enumerators 31
      3.4.3. Data collection 32
   3.5. Data analysis and interpretation 33
      3.5.1. What to look for in the data 33
      3.5.2. First steps of the analysis 35
      3.5.3. Methods 35

TABLE OF CONTENTS

4. PILOTING A TOOLKIT THAT OFFERS A MULTIDIMENSIONAL PERSPECTIVE ON MIGRANT INTEGRATION IN THE LATIN AMERICAN AND THE CARIBBEAN REGION 37
   - 4.1. Context of piloting the measurement tool in the Latin American and the Caribbean region 38
     - 4.1.1. Methodology of piloting the toolkit 38
     - 4.1.2. Overview of the studies in Brazil, the Dominican Republic and Peru 40
   - 4.2. Case studies and country reports 41

5. CONCLUSION AND FINAL CONSIDERATIONS 45
   - 5.1. Towards a multidimensional measurement approach of integration outcomes 46

6. ANNEXES 49
   - 6.1. Checklist to plan the application of IPL integration index 50
   - 6.2. Modes of data collection – an overview 53

7. REFERENCES 57

EXECUTIVE SUMMARY

Migrant integration or the degree to which migrants can build a successful and fulfilling life in their country of residence is a major issue for policymakers and practitioners. This is especially true for the Latin America and Caribbean (LAC) region, as the displacement of over five million Venezuelans and a large-scale movement from Central America to South American countries has presented a major challenge for governments and hosting communities. In order to meet this challenge and facilitate migrant integration, it is important to understand where help is needed most and how well different interventions work.

This guide presents the Multidimensional Integration Measurement Toolkit and the Immigration Policy Lab (IPL) Integration Index which serves as the core survey instrument of the toolkit developed by IOM and the IPL at ETH Zurich. The index is a pragmatic and versatile survey instrument to measure migrant integration across six key dimensions: social, political, navigational, economic, psychological and linguistic dimensions. While the guide does not suggest that the index is the only way to measure integration, the index nevertheless presents questions that are versatile enough to be asked to migrants and non-migrants alike and designed to produce valid results in nearly every context. The collected responses can also be used to study integration outcomes in different dimensions or to compute an index score that can be used to compare migrant integration across various contexts. These different measures of migrant integration can then be used to better understand the needs of migrant populations or to evaluate the effectiveness of different interventions, thereby informing policy and practice towards the full inclusion of migrants.

To facilitate the use of the broader measurement toolkit, the report details how the index was piloted and applied in Brazil, the Dominican Republic and Peru. Based on the experiences in these three studies, the report offers important lessons learned and a detailed step-by-step guide on how to plan and implement similar research studies. The report further presents detailed guidelines on how to apply the measure, organize data collections, evaluate support programmes, and how to use the response data. The report aims to be a valuable resource for IOM missions and external partners who want to use surveys and survey data to better understand the communities they serve or the impact of their support programmes.

1. INTRODUCTION

A better understanding of the intersectional nature of the various domains of integration, and how they interact and unfold over time, is critical for designing and implementing interventions and policies with real impacts. In that sense, the International Organization for Migration (IOM), in partnership with the Immigration Policy Lab at ETH Zurich, has developed a multidimensional approach to measure integration outcomes of migrants and accompanying materials. The approach defines integration as the degree to which migrants have the knowledge and capacity to build a successful, fulfilling life in the host society.

Many Venezuelans traveling through the continent do so by foot carrying all their children and possessions. Caminantes, or walkers, trek along major highways and through difficult terrain. They must go through mountainous areas where temperatures drop below zero and through scorching hot areas where water is scarce. Many make this journey with just a light jacket, rubber flip flops and a small backpack with the most essential items they manage to carry. September 2019, Colombia. © IOM 2022 /Muse MOHAMED

1.1. Understanding and measuring migrant integration outcomes

Migrant integration, albeit a complex issue, has always been an important part of the migration phenomenon. In an increasingly globalized world, the growth in the absolute number of migrants over the past 50 years and the diversification of migrants' origins, socioeconomic backgrounds and reasons for migrating have led to more social, cultural, ethnic and religious diversity in receiving societies. The unprecedented growth in the number of migrants over the last years poses questions on how to fully utilize migrants' existing assets and resources, as well as to capture their broader settlement perspectives and integration experiences in society.

However, little is known about migrants' broader integration experience and their perspectives of arrival and settlement in different facets of community life. Existing surveys and recent assessment studies on migrants have shown that the focus has been mostly on employment, service access and to a lesser extent, on levels of discrimination and social belonging.

Oftentimes, researchers and policymakers assess "integration" from a purely "service access" or economic perspective, without looking at the social and relational aspects that migrants have towards their communities. Many of the existing integration frameworks, assessment tools and studies fail to consider migrants' knowledge and ability to establish themselves in new communities – a process that includes psychological, linguistic, social and political or civic dimensions, among others (Harder et al., 2018; United Kingdom Home Office, 2019).

Against this background, there remains limited data on integration that could inform the development and implementation of policies and interventions to address these issues. This is further aggravated by the lack of empirical tools to capture the multidimensionality of the integration process, which impedes progress towards designing and implementing more effective interventions.

1.2. Developing a toolkit to measure integration outcomes

A better understanding of the intersectional nature of the various domains of integration, and how they interact and unfold over time, is critical for designing and implementing interventions and policies with real impacts. **In that sense, the International Organization for Migration (IOM), in partnership with the Immigration Policy Lab (IPL) at ETH Zurich, developed a multidimensional approach to measure integration outcomes of migrants and accompanying materials. The approach defines** *integration as the degree to which migrants have the knowledge and capacity to build a successful, fulfilling life in the host society.* **This definition recognizes the importance of migrants' existing knowledge and capacity.**

Central to this approach is the IPL Integration Index – a pragmatic survey instrument – that captures six key dimensions of migrant integration: economic, social, political, psychological, linguistic and navigational dimensions. The proposed approach is easy to replicate and can be deployed in multiple contexts, including in large-scale population surveys, project-linked assessment studies and other data-driven methodologies to improve understanding of the integration process and its links to broader development outcomes.

Using this measurement tool can address important data gaps, considering that migrants are often "statistically" invisible in regular national data collection instruments (e.g. population censuses and labour force surveys). In addition to addressing these data gaps, replicating this measurement tool could also generate comparable data on migrant integration across different contexts, which are often lacking in lower- and middle-income countries. Furthermore, this would not only contribute to a better understanding of contexts and trends and improve the design of interventions, but also help articulate the contribution of IOM interventions to broader strategic outcomes in national action plans or policies.

This guide details how the index was piloted in three countries in the LAC region: Brazil, the Dominican Republic and Peru, and draws lessons from this data-collection exercise using the index.

MEASURING INTEGRATION OUTCOME VERSUS INTEGRATION POLICY

Measuring integration outcomes, as opposed to measuring integration policies, is critical as it describes the integration process as experienced by migrants in certain areas of society. It can describe the results of integration policies, which is relevant for capturing the extent and nature of the integration process. Measuring these outcomes can yield insights into the quality of the integration process and where there are gaps and opportunities in interventions and policies. This is oftentimes overlooked by actors who are mainly interested in measuring integration policies.

Please see Migration Governance Indicators for measuring policies on migration.

1.3. Overview and structure of the practical guide

This practical guide contributes to strengthening the capacities of practitioners to measure integration outcomes by providing detailed and in-depth guidance for the use of the Multidimensional Integration Measurement Toolkit. The guide seeks to present the toolkit's components, key features, offering concrete guidance and recommendations that build upon the lessons learned during the pilot of the toolkit in the Latin America and Caribbean region. It provides some basic operational considerations that can allow for its replicability and scalability in different contexts. Furthermore, the practical guidance offers insights on where and how the IPL Integration Index can be applied in different measurement approaches and programmatic contexts.

This document is structured as follows:

CHAPTER 2

Multidimensional integration measurement toolkit
This chapter describes the key components of the multidimensional integration measurement approach with IPL Integration Index as the core survey instrument.

CHAPTER 3

Application and adaptation of multidimensional integration measurement toolkit
This chapter provides basic considerations in rolling out the index, including a systematic step-by-step approach in locally adapting the toolkit, designing the study, onboarding and piloting, data collection and analysis and interpretation.

CHAPTER 4

Piloting a multidimensional integration measurement toolkit in the Latin America and the Caribbean (LAC) region
This chapter describes the multistep operational process of adapting the index and toolkit and the different methodological approaches and findings from the studies conducted in Brazil, the Dominican Republic and Peru.

CHAPTER 5

Conclusion and final considerations
This chapter provides some practical and strategic considerations and recommendations in terms of programming and policy discussions based on the experience of applying the toolkit in the LAC region and on how it can be further scaled and replicated in other measurement and geographical contexts.

2. MULTIDIMENSIONAL INTEGRATION MEASUREMENT TOOLKIT

"The opportunity to collaborate on a multidimensional approach to measure migrant integration outcomes exemplifies the importance of bringing in a research perspective in monitoring and evaluation activities to generate robust data and shape new solutions. With its data-driven mission, our team at the Immigration Policy Lab is thrilled to leverage our expertise and cutting-edge analytical tools to improve policy and practice on a global sphere and we are eager to continue working with IOM to advocate for evidence-based policymaking on migrant integration."

***Dominik Hangartner**, Professor of Public Policy at ETH Zurich and Co-Director of IPL*

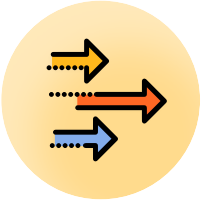
Jose Quero (in white jacket) left the Bolivarian Republic of Venezuela with four friends three days ago, leaving behind his three children in the middle of the night, something he described as an "impossible decision." September 2019, Colombia. © IOM 2019 / Muse MOHAMED.

2.1. A core survey instrument: the IPL integration index

Most actors in the space of public policy agree that migrant integration is as crucial a goal for receiving societies as it is for migrants. Despite this consensus, there is a lively debate on what integration is and how it can be measured. In stark contrast to this debate, many empirical studies use different ad hoc or stopgap measures of migrant integration. Frequently, only migrants' employment status or language proficiency are used to operationalize the complex process of migrant integration. Not only does this inadequately represent the multifaceted concept of integration, but the many different measures also make it hard to systematically aggregate knowledge from different studies on migrant integration.

The IPL Integration Index – which is the core survey instrument of the Multidimensional Integration Measurement Toolkit – was originally developed and published in 2018. It was designed to measure the six most important dimensions of migrant integration, while at the same time being general enough to be applied in as many contexts as possible. The goal was to create a single survey measure of migrant integration that covered different aspects of integration and could produce comparable data on different dimensions of the migrant integration process. Over time, the data generated by the survey tool could create a broad empirical foundation for evidence-based policies to foster migrant integration.

To create the index, the research team first consulted the literature and organized an expert workshop to collect the most important dimensions of integration and possible survey questions. To narrow the set of possible dimensions and questions down, the research team then made two crucial conceptual decisions.



First, to be independent from local normative debates, all dimensions should have a clear directionality. This means that on all included dimensions, higher values – all else being equal – usually mean better integration. For example, speaking a language better, earning a higher wage, or feeling more at home usually also means better integration. This is not the case for cultural values. While a given society might be more progressive or more conservative, being too progressive or too conservative will typically also lead to processes of disintegration or disassociation. Here, a measure of cultural integration would first have to define a set of cultural practices and values that are generally accepted in the respective society. This would have made the index very dependent on local normative debates and thus would have limited comparability.

Second, integration was understood as a process of acquiring new capacities and resources (e.g. learning a new language, finding new friends or learning about a new political system). This differentiates the index's definition of integration from assimilation. Measures of assimilation also measure the process of giving up capacities and resources, such as no longer speaking a language or practicing a certain religion.

After these considerations, the research team at IPL ended up with six dimensions of integration (see below) and tested more than 100 survey questions to measure them. In testing the questions, the research team looked for questions that could be posed to migrants and non-migrants alike, showed a good distribution of answers between different migrants, would not exclude certain populations by design (e.g. questions about the workplace would exclude the unemployed and homemakers), and could be used in as many contexts as possible.

The index's focus on broadly applicable survey items and a multidimensional understanding of integration makes it a good and practical measure of integration outcomes in many situations. For very specific studies, however, the index might be too broad. Studies that focus on very particular migrant populations or research questions should therefore consider complementing the index with questions that are designed to capture the unique features of the population at hand. The same is true for evaluations of very particular support programmes or policies.

Integration dimensions

PSYCHOLOGICAL
Captures respondents' feeling of connection with host country, their wish to continue living there and their sense of belonging.

LINGUISTIC
Captures respondents' assessment of their ability to read, speak, write and understand the dominant language of their host country or region.

SOCIAL
Captures social ties and interactions with nationals in the host country, as well as bridging social capital as evidenced by participation in organizations with nationals.

ECONOMIC
Captures income, employment, satisfaction with employment situation and the ability to meet different levels of unexpected expenses.

POLITICAL
Captures understanding of the important political issues facing the host country and the degree to which respondents engage in discussion and political knowledge.

NAVIGATIONAL
Captures the ability to manage basic needs in the host country, such as seeing a doctor, addressing legal problems and searching for jobs.

RESOURCES

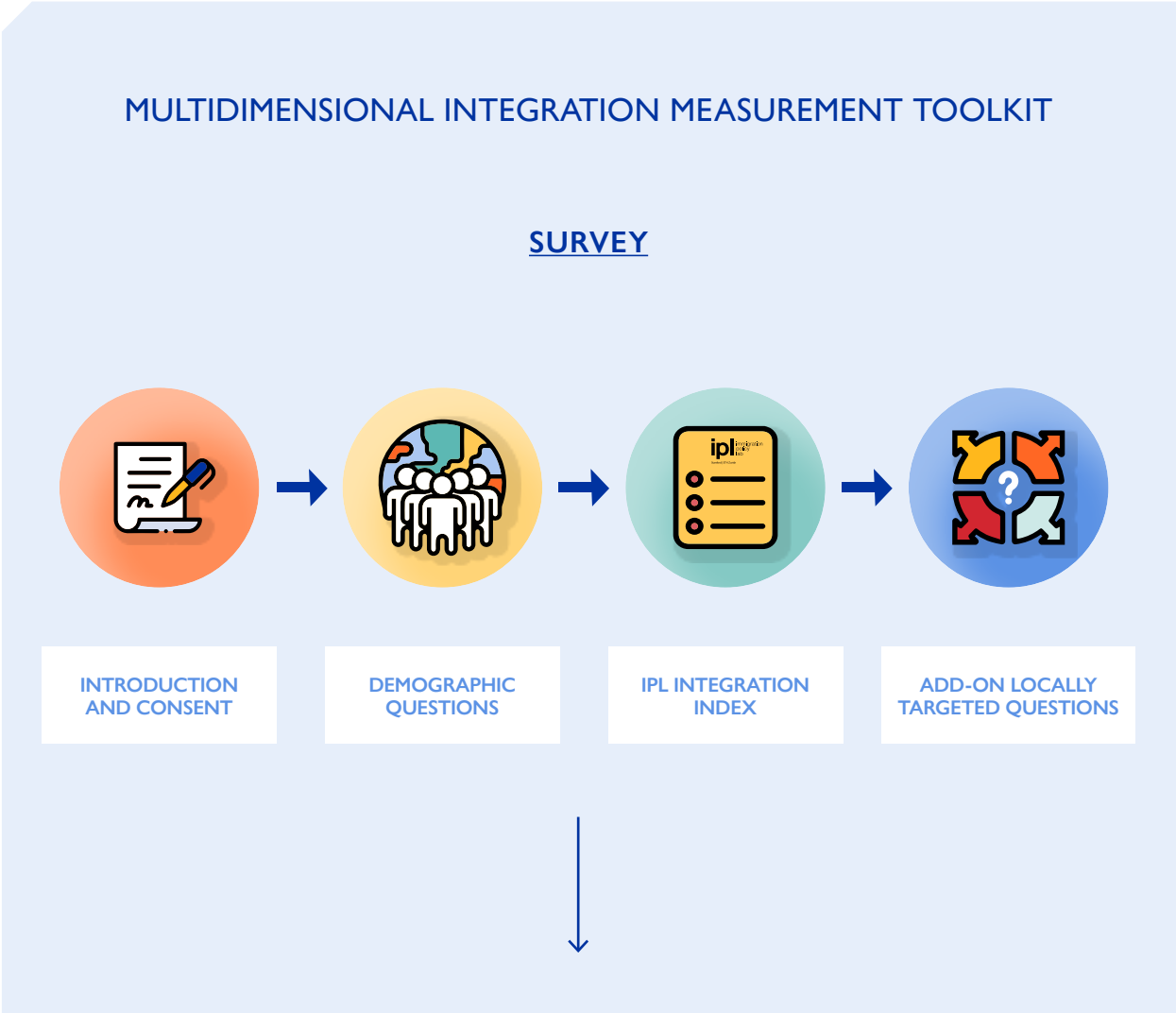
To understand more about the IPL Integration Index, please access the following documents:

- The IPL Integration Index website
- Multidimensional measure of immigrant integration (The academic paper published in the *Proceedings of the National Academy of Sciences* (PNAS), an American peer-reviewed multidisciplinary scientific journal).
- Supplementary information for "Multidimensional measure of immigrant integration".

2.2. Structure and components of the multidimensional integration measurement toolkit

The Multidimensional Integration Measurement Toolkit is made up of two components: the survey and the accompanying materials. Though not mutually exclusive, there are four parts that constitute the **survey component** of the Multidimensional Integration Measurement Toolkit. In the **introduction and consent section,** respondents are informed, in a language they understand, what the purpose of the study is, with whom their data would be shared, what data would be shared and their right to refuse the data collection. They are communicated that in case they share their personal data and/or do not participate in the survey, there will not be any implications for their legal status and access to services. Only respondents who give their informed consent are interviewed. In the second part, respondents are asked questions regarding basic **demographics** such as age, gender and household size. The **IPL Integration Index** makes up the core module, which is the third part of the overall survey tool. The fourth part includes **additional variables** of specific interest to the context at hand. It is important to start with the introduction and consent section, while the order of the other three parts is not as important.

In addition to the survey component, the toolkit includes a wide range of accompanying materials which support its design and implementation. The below graphic further details how each of the materials feed into the application of the toolkit. By encompassing both the survey and accompanying materials, the Multidimensional Integration Measurement Toolkit offers practitioners a comprehensive approach to capture robust data on beneficiaries' integration outcomes.



ACCOMPANYING MATERIALS

The research paper "Multidimensional measure of immigrant integration" and the accompanying appendix present the conceptual foundations and process of constructing the IPL Integration Index.

The Study Guide provides specific instructions on methodological proceduresassociated with using the IPL Integration Index questionnaire. The guide has been developed to help ensure methodological consistency.

The Check-list to Identify Projects for Evaluation provides guidance to identify projects or interventions whose design is suitable for an assessment.

The Enumerators Guide provides guidance on how the interviews should be conducted in case they are not self-administered.

The Scoring System allows to calculate the integration score based on the score of each question, computed between 1 and 5 points. This document refers to the survey developed for Brazil in collaboration with IOM headquarters and IOM Brazil.

The Form for Pilot Findings and Feedback presents some thoughts on what to consider after conducting a pilot test as well as best practices.

The Survey Modes present an overview of the different modes of survey data collection.

The Structure and Application of the Index presents the different dimensions and an overview of the questions.

The survey is available in multiple languages: English, Spanish, Portuguese and French.

Code to analyse IPL-12 (R, Stata) and IPL-24 (R, Stata) data.

The Survey Bank on Migrant Integration and Social Cohesion (2021) contains questionnaires developed around the world to measure migrant integration and social cohesion in different contexts.

2.2.1. Introduction and consent

Before starting the survey, it is important that all necessary information is clearly present and defined so that the respondent can make an informed, voluntary decision about what they agree to. It is crucial to consider the age, culture, ethnicity and literacy level of respondents, to mitigate the risk of misinterpretation or alienation.

The introductory script should contain the following, but is not limited to:

- Information about the institution commissioning the survey
- Research purpose
- How their data will be used
- Duration of subject's participation
- Contact information of the research team
- A statement confirming their participation is voluntary and their identity will be kept confidential.

It is crucial that respondents are informed of the above points, in a language they understand, emphasizing that their refusal to share their personal data and to participate in the survey will not have any implications for their legal status and/or access to services. Only respondents who give their informed consent are interviewed.

A checklist to help support the process of attaining informed consent is available here (internal link only).

Sample introductory script for a telephonic survey

Hello! My name is.... and I am calling on behalf of <insert name of the institution.> We are carrying out a research on migrants' integration in <insert place> and we would like to learn about your experience. We got your contact details from <insert institution/programme>.

Participation in the survey is entirely voluntary. The survey will take around <insert time>. Would you have time to speak with me now?

[If Yes: Proceed with survey]
[If No:] Would it be possible to call you back some other time?
[If Yes:] Which day and time would be good for you?
[Record date and time in a separate file. Then, please close the survey here.]
[If No: Thank respondent and exit survey.]

All of your answers will remain strictly confidential. Your name will not be linked or included with the results; however, your age, household income, and location will be recorded with your responses. Your name will not be linked to any information you share with us during the interview. You can skip or refuse to answer some questions or stop the survey at any point in time. There are no right or wrong answers. Your participation will not affect your participation in any IOM programmes. If you should have any questions or concerns regarding the use of your responses or your participation in the survey, you can contact <insert contact>.

By answering this survey, you agree that we can use your answers for this research for statistical analysis.

Do you agree to take part in this survey? [Yes or No]

2.2.2. Demographic questions

Depending on the purpose of the study and its target audience, below are some possible demographic questions that could be incorporated as part of the survey. Note that these questions are not included in the measurement index.

Questions:

- Gender
- Age
- Nationality
- Household size
- Location
- Race and ethnicity
- Length of stay in the host country
- Education level
- Immigration status
- Legal documentation

These variables will provide additional insights into characteristics of the surveyed respondents and show how integration outcomes differ between groups. For example, asking about migrants' education levels can provide insight into how integration levels vary between those with more education and those with only little education. Based on these questions, demographic vulnerabilities can also be identified (e.g. elderly, Lesbian Gay Bisexual Transgender Queer Intersex Asexual and more, irregular migrant, NEET – not in education, employment or training) which can provide more awareness into how these people fare in comparison to other subgroups in various integration dimensions and who might need more support for example.

2.2.3. IPL integration index: survey questions

There are two versions of the index: the short form (IPL-12) which proposes a reduced set of questions, with two questions, for each dimension of integration, and the long form (IPL-24) which allows a more precise measurement of each dimension with up to four concepts per dimension. The added precision, however, comes at the cost of survey time. Previous surveys have shown that answering the IPL-24 takes two to three times longer than answering the IPL-12. The IPL-24 should therefore only be used when the additional questions are of substantial interest to the research team.

DIMENSION	IPL-12	IPL-24
NAVIGATIONAL	Seeing a doctor (access to health).	Seeing a doctor (access to health).
	Searching for job listings (access to employment).	Searching for job listings (access to employment).
		Getting help with legal problems (access to justice).
		Seeking medical help for chronic back pain.
		Sending mail.
		Paying taxes.
		Alcohol consumption when driving.
ECONOMIC	Monthly household income.	Monthly household income.
		Indication of necessary but extra expenses that respondents can pay.
	Occupation in the last 4 weeks.	Paid work in the last 4 weeks.
		If yes, description of type of work.
		If yes, formal or informal.
		If not, description of occupation/reason not to work.
		Satisfaction with employment situation.
SOCIAL	Phone contacts and conversations with locals.	Frequency of sharing meals with locals
		Phone contacts and conversations with locals.
		Participation in groups or associations (in person or virtual): job, religious, hobbies, social or political cause, another volunteer organization.
	Frequency of sharing meals with locals.	How many locals in these groups.
		Frequency of providing favors to locals.

DIMENSION	IPL-12	IPL-24
LINGUISTIC	Skills in "host country's language": reading.	Skills in "host country's language": reading.
		Skills in "host country's language": speaking.
	Skills in "host country's language": speaking.	Skills in "host country's language": writing.
		Skills in "host country's language": listening.
PSYCHOLOGICAL	Connection to "host country".	Connection to "host country"
		Frequency of feeling as an outsider.
	Frequency of feeling as an outsider.	Where want to live in the future.
		Frequency feeling isolated in the local society.
POLITICAL	Understanding of political issues.	Understanding of political issues.
		Understanding of public policy and current events regarding political parties.
	Frequency of discussion of political issues with others.	Frequency of discussion of political issues with others.
		Level of engagement in political process, including activist movements.
HOUSEHOLD	Number of people living in a household.	Number of people living in a household.

2.2.4 Additional and targeted questions

Most surveys include questions on basic demographics such as age, gender, or marital status. To better understand potential differences between such groups, further questions on the respondents' experiences in a country are often helpful. Hence, depending on the respective context, it is advisable to include further questions on the respondents' life. Such questions can include questions on experience with discrimination, prior migration, family ties in the host country, aspirations, or neighbourhood characteristics. Every additional question increases the time it takes to answer the survey and thus the costs for researchers and survey respondents. Additional questions should therefore only be included if there is a strong theory or expectation about how answers to these questions will help to better understand the index scores.

A good overview of potentially relevant and pre-tested questions can be found in the Global Survey Bank on Migrant Integration and Social Cohesion (developed by DISC and IOM's Global Migration Data Analysis Centre) or other surveys on migrant integration.

2.3. Scoring system

All questions of the IPL Integration Index have a substantial meaning that can be evaluated without further data manipulation. For example, learning about a respondent's income or feeling of connectedness can be informative without recoding the data. However, if the goal is to compare migrant integration over different contexts and dimensions, a single, unified integration score can be useful. To compute this score, substantive answers like "I find it somewhat difficult to see a doctor" are translated into a scoring system that creates aggregate values between 0 and 1. It is important to keep in mind that while a value of 0 signals very low levels of integration, it is not necessary to reach a value of 1 to speak of high levels of integration. All values above 0.75 can be considered as a signal for high levels of integration and in many cases even native citizens will not reach a score of 1 but score somewhere between 0.75 and 1. The sections below will go into more detail on how the scoring system works. For a full guide, see the supplementary information for the journal article that introduced the IPL Integration Index (Harder et al., 2018). For further information, refer to the Annexes.

2.3.1. Basics of the scoring rule

The scoring system starts with assigning a value between 1 and 5 to every possible answer. Here, 1 is assigned to the answers that suggest the lowest integration and 5 to the answers that suggest the highest integration.

For example, respondents who judge their ability to "read and understand the main points in simple newspaper articles on familiar subjects" as "very well" are assigned the value 5 on this question. Respondents who judge their ability as "not well at all" are assigned the value 1. After assigning a value to every answer, these values are now summed up. This is relatively straightforward for the IPL-12 and a bit more complicated for the IPL-24. The IPL-12 has exactly twelve questions and respondents can score a value between 1 to 5 on each of the questions. Therefore, the sum of all scores must lie somewhere between 12 (12 x 1) and 60 (12 x 5). The IPL-24 has 24 scored concepts. Consequently, the sum of all scores must lie between 24 (24 x 1) and 120 (24 x 5). To make the two scores more comparable, the final score is then rescaled to the range between 0 and 1.

2.3.2. Detailed scoring

The scoring for all questions on psychological and linguistic integration are straightforward and the respective values can be taken directly from the annex (please see coded values of the Brazilian questionnaire). For the IPL-12, the scoring for navigational, social, and political integration is similarly straightforward. Here, only the coding for the IPL-24 is more complex.

Compared to all other questions, assigning scores to the respondents' household income is a more complex task. First, the "equivalized household income" needs to be computed. Second, the computed equivalized household income needs to be set in relation to the respective country's median equivalized household income.

Computing the equivalized household income

As a household grows, its needs grow as well. However, the growth in needs is usually not proportional to the growth of the household. Compared to a single person household, a three-person household would usually not pay three times the rent or utilities. To account for this, researchers have developed a number of different ways to adjust household income by household size (OECD, n.d.). The IPL Integration Index follows this approach and instead of using just the household income, uses the "equivalized household income".

There are different ways to adjust or equivalize household income. Most of these adjustment procedures need additional information, such as the age of all household members. This information is not always available or is costly to collect. A more lightweight and straightforward way to adjust household income is to divide household income by the square root of the household size. According to this approach, a four-person household would need about twice the income of a single-person household to live in similar economic prosperity. The index uses this approach to estimate the respondents' "equivalized household income".

Computing the "equivalized household income" requires values for household income and household size. While household size is usually measured directly, the question for household income uses income brackets[1] as answer options. To compute the equivalized household income, the index suggests taking the middle value from each income bracket. This value can then be divided by the square root of the respondents' household size. This procedure yields an estimate of the respondents' equivalized household income.

Scores in relation to the median equivalized household income

It is unreasonable to assume that a respondent's household should be among the richest households to be considered "integrated" or economically successful. Instead, the index assigns maximal economic integration to every household with an equivalized household income that is substantially higher than the median equivalized household income in the respective country. More specifically, the index assigns the following scores:

The index assigns a score of 5 to all respondents with an equivalized household income that is more than a third of the median equivalized household income higher than the respective country's median equivalized household income.

A score of 4 is assigned to all respondents with an equivalized household income higher than the median equivalized household income but not more than a third of the median equivalized household income.

A score of 3 is assigned to all respondents with an equivalized household income that is higher than two thirds of the median equivalized household income but not higher than the median equivalized household income.

A score of 2 is assigned to all respondents with an equivalized household income that is higher than a third of the median equivalized household income but not higher than two thirds of the median equivalized household income.

A score of 1 is assigned to all respondents with an equivalized household income that is lower or equal to a third of the median equivalized household income.

Example: understanding the scoring of household income question

To better understand the scoring of the household income question, consider this example. Instead of the yearly household income, the study in Brazil used the monthly household income. Based on available research, the study team estimated the median equivalized household income to be around 1,327 Brazilian reais (R$). Based on this value, the categories for the final coding are R$ 0 to R$ 442 = 1; R$ 443 to R$ 885 = 2; R$ 886 to R$ 1,327 = 3; R$ 1,327 to R$ 1,769 = 4; R$ 1,770 and higher = 5.

Now assume a participant answers that they are a family of four with a household income of R$ 1,400. In the survey, this family would have selected the income bracket R$ 1,101 to R$ 1,500. This answer would be recoded as the value that falls in the middle of the interval, R$ 1,300.5, and divided by the square root of four: R$ 1,300.5/√(4) = 650.25. This value falls into the second scoring interval R$ 443 to R$ 885 and is assigned a score of 2.

[1] Typically, deciles of the respective country's income distribution.

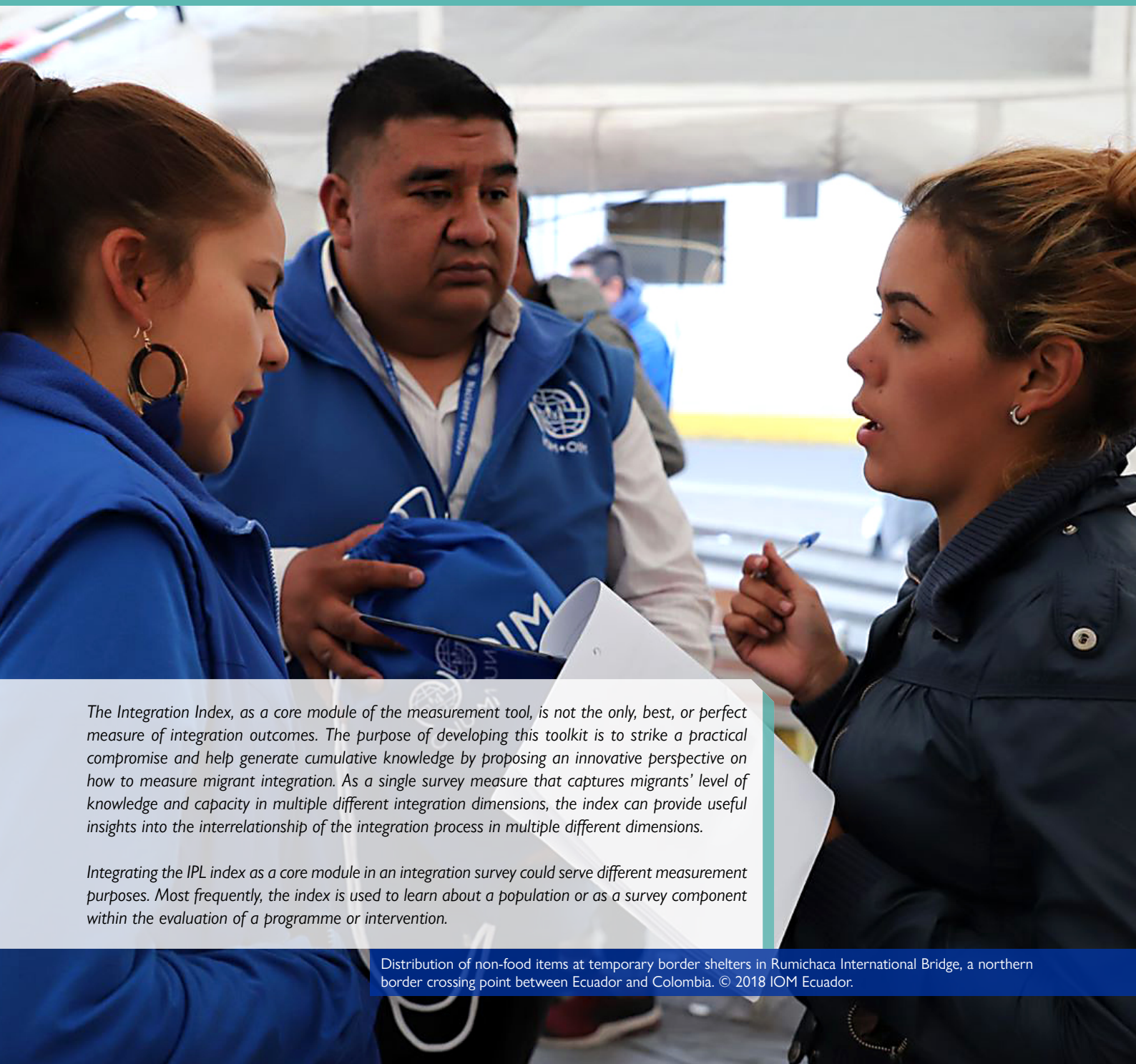
2.3.3. Rescaling

After scoring all answers and adding up the scores, every respondent should have a total score somewhere between 12 and 60 (in case of the IPL-12) or between 24 and 120 (in case of the IPL-24). The total score then needs to be rescaled to values between 0 and 1. This can be achieved by applying a simple formula using Excel or any other statistical software.

> To rescale IPL-12 values, compute (x-12)/48. To rescale IPL-24 values, compute (x-24)/96.

When rescaling data, it is important to handle missing values correctly. If a respondent did not answer one of the 12 (or 24) questions, it is impossible to compute the exact index score. However, something similar can be computed if the formula for rescaling is changed. The full form formula for rescaling is: (x – smallest possible value) / (highest possible value – smallest possible value). If the answer to one or more questions is missing, the smallest and highest possible values change. For example, if the answer to a question in the IPL-12 is missing, the formula changes to: (x – 11) / (55 – 11) = (x-11)/44. This way, the data can be rescaled without adding additional errors in the rescaling process. Still, due to the missing answer the rescaled score reflects less information than the exact index score. Simulations with the index data from different datasets show that the overall score changes little if one or two questions are removed. If more than two questions are removed, however, the final scores differ significantly (see section G.3 in the Supplementary Information of the PNAS paper).

3. APPLICATION AND ADAPTATION OF MULTIDIMENSIONAL INTEGRATION MEASUREMENT TOOLKIT



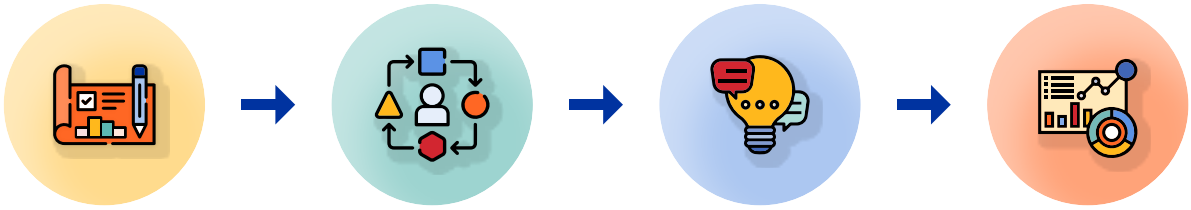
The Integration Index, as a core module of the measurement tool, is not the only, best, or perfect measure of integration outcomes. The purpose of developing this toolkit is to strike a practical compromise and help generate cumulative knowledge by proposing an innovative perspective on how to measure migrant integration. As a single survey measure that captures migrants' level of knowledge and capacity in multiple different integration dimensions, the index can provide useful insights into the interrelationship of the integration process in multiple different dimensions.

Integrating the IPL index as a core module in an integration survey could serve different measurement purposes. Most frequently, the index is used to learn about a population or as a survey component within the evaluation of a programme or intervention.

Distribution of non-food items at temporary border shelters in Rumichaca International Bridge, a northern border crossing point between Ecuador and Colombia. © 2018 IOM Ecuador.

3.1. A systematic approach to the application of the toolkit

While the first chapter describes how the core survey instrument – the IPL integration index – works, this chapter will provide a step-by-step guide on how to design and implement a multidimensional measurement approach to integration outcomes. The latter involves four steps:



DESIGN OF THE STUDY

Helps define the purpose for which the Multidimensional Integration Measurement Toolkit can be used and proposes a range of survey designs in accordance with this purpose. This section also presents different survey modes according to the survey designs.

LOCAL ADAPTATION

Provides instruction on how the index can be adapted to the local context, to ensure that the questions are relevant and understandable to the respondents.

PILOT TESTING AND DATA COLLECTION

Provides methodological tips and a presentation of research methods in order to conduct systematic and coherent tests and data collection activities.

ANALYSIS AND INTERPRETATION

Provides practical guidance on what to look for in the collected data as well as some methods to analyse and interpret them.

RESOURCES

For more information on how the toolkit has been applied and what needs were addressed in other project contexts in IOM, please refer to the country case studies in **Chapter 4** where the country reports can also be downloaded.

For more comprehensive information on developing empirical studies in line with IOM standards, please refer to the IOM Monitoring and Evaluation Guidelines.

3.2. Design of the study: defining the purpose and measurement context

This guide proposes to use the IPL Integration Index as a core module in a multidimensional integration survey. As a single survey measure that captures migrants' level of knowledge and capacity in multiple different integration dimensions, the index can provide useful insights into the interrelationship of the integration process in multiple different dimensions.

For example, the objective of a vocational training programme for migrants might be to improve their employment outcomes. Yet such employment-oriented intervention can also have secondary effects on other dimensions of integration such as social or linguistic integration (e.g. improving connections with host communities or improved communication). Conversely, the level of integration in those other dimensions can also influence the employment outcomes that the programme tries to achieve.

Similarly, the index could also be used as part of baseline assessment for developing an integration intervention. Adding the index to a broader beneficiary survey could yield insight into the levels of integration across the different dimensions – whether economic, social, navigational or linguistic. The findings of such a survey could help with prioritizing which integration dimension is to be targeted by an intervention or developing more complex interventions that target multiple dimensions (e.g. a vocational training programme with a strong language-training component).

Integrating the IPL index as a core module in an integration survey could serve different measurement purposes. Most frequently, the index is used to learn about a population or to learn about the effect or impact of a programme or an intervention. Both purposes are not mutually exclusive but conducting a sound programme evaluation is significantly more demanding.

For example, while in Brazil and the Dominican Republic the index was used to learn about the integration outcomes of cash-based intervention (CBI) project beneficiaries, in Peru it was used within the context of a causal impact evaluation of the effects of the CBI itself on their beneficiaries' integration outcomes. Please see Chapter 4 for more information on how the index was used.

This section will discuss different survey designs and research designs and introduce basic concepts in study design.

3.2.1. Basic concepts in study design

Target population

The first thing to think about when designing a study is the target **population**. This is the group the study wants to learn about.

Here, it is important to think carefully about who the study wants to learn about. For example, a study on "migrant beneficiaries" in country A could target only migrants who are currently beneficiaries of a given programme. However, the same study could also include all migrants who ever benefited from a programme, including migrants who have since left the country. Or the study could survey all migrants who would qualify for the benefits programme. There are no right or wrong answers to these questions, as answers depend a lot on the specific context. However, carefully considering such questions will help to better understand the scope, strengths, and limitations of the study at hand.

Sample

In most cases, it is not feasible to survey the whole population of interest. Instead, many studies only survey a **sample** of the population. Answers from this sample are then used to generalize to the whole population. The quality of such generalizations critically depends on the way the sample was generated. The most reliable generalizations are possible if every person in the population had the same probability to be included in the sample. Such samples are usually called **random samples**. In most cases, generating a random sample is only possible if the sample is defined using some type of register that includes the whole population. However, even in cases which do not allow for a random sample, it is useful to consider and to limit the potential factors influencing the probability to include certain groups from the population in the sample. For example, web surveys can only be accessed by respondents with Internet access and the technological know-how to operate computers or smartphones. Surveys during working hours are less likely to be answered by employed members of the population. The other way round, surveys conducted in the afternoon or evening are less likely to be answered by members of the population who work night shifts. When such factors biasing the sample are well understood, they can be considered when analysing and interpreting the data, such as making adjustments during statistical analysis. The best way to do this depends on many different factors. It is thus recommended to collaborate or consult with external survey researchers to statistically adjust sample bias.

Survey mode

The kind of sample that can be realized often depends on the **survey mode**. The way survey responses are collected is called the "survey mode". In general, there are four main ways to collect survey responses. Respondents can either (1) be interviewed face-to-face by an enumerator, (2) be interviewed through a phone call by a trained interviewer, (3) fill out a paper-based survey that was mailed or handed to them, or (4) fill out an online survey on their phone or computer. These main survey modes can also be mixed. Computer-assisted personal interviewing, for example, is a technique in which surveys are answered on an electronic device, but with the help of a face-to-face interviewer. The choice between survey modes is often determined by the type of available contact information (i.e. addresses, phone numbers, or email addresses) but also comes with a trade-off between cost and data quality. Due to the health restrictions during the COVID-19 pandemic, for instance, research conducted in 2020 and 2021 was conducted via phone surveys. A first step to make this decision is to consider typical benefits and drawbacks of the different survey modes:

FACE-TO-FACE SURVEYS	Benefits of face-to-face surveys	Generally understood to deliver the most reliable data.
		A trained interviewer can make sure that the survey is answered by the desired respondent, that all questions are understood correctly, and can encourage answers to all questions.
		The survey can last longer, which allows you to include more questions.
	Drawbacks of face-to-face surveys	A face-to-face interview often takes longer than answering phone or online surveys.
		Hiring and training interviewers is costly.
		Interviewers need to travel to conduct interviews. This increases the time in the field and the cost of the survey.
PHONE SURVEYS	Benefits of phone surveys	Generally understood to deliver reliable data.
		A trained interviewer can make sure that all questions are understood correctly, and can encourage answers to all questions.
		Interviewers do not need to travel.
		Interviews can be held from a lab with translators on site.
	Drawbacks of phone surveys	A single interview often takes longer than in online surveys.
		Hiring and training interviewers is costly.
		Interviewers cannot confirm whether they are talking to the right person.
		Only respondents with phones can be interviewed.
		The survey time has to be shorter than for an in-person interview.
WEB SURVEYS	Benefits of web surveys	A cost-efficient way to interview many respondents in a short time.
		Benefits from all advantages of digital survey forms (i.e. randomization, attention checks).
	Drawbacks of web surveys	Data quality can be questionable.
		Hard to confirm whether the survey is answered by the right person.
		Only literate respondents with Internet access and a certain level of technical knowledge can be interviewed.

PAPER-BASED SURVEYS	Benefits of self-administered paper-based surveys	Generally understood to deliver reliable data.
		Respondents can answer at their own time and choose when to answer.
		Previewing all questions can increase respondents' trust in the survey.
		No need to employ and train interviewers.
	Drawbacks of self-administered paper-based surveys	Printing and distributing the survey can be costly or unreliable. The same is true for receiving the filled-out surveys.
		Offering the survey in different languages further increases print and distribution costs.
		Hard to confirm whether the survey is answered by the right person.
		Randomizations are hard to implement.
		Responses need to be digitized by the research team.
		You need to have the respondents' address and be able to send them paper-based surveys and/or to find a way to hand them out to the intended population.

RESOURCES

For more information on the benefits of the different modes of survey data collection, please refer to Survey modes.

3.2.2. Basic survey designs

Cross-sectional surveys

Gathering information from a population of interest at single point in time is often the least complex study design. This design requires drawing a single sample from the population of interest and surveying it once. Using the IPL Integration Index in such a study will give you a "snapshot" like understanding of the level of integration and correlates of integration in your population of interest.

Typical insight from cross-sectional surveys:

- Data from the individual index questions deliver insights into important aspects of integration. For example, what is the respondents' most frequent employment status, how well can they speak the local language, or how high or low are the Integration Index scores on the individual dimensions?
- Index scores can be used to compare the integration level of the surveyed population of interest with integration levels from other studies. For example, if the study at hand was done on migrants from country A in country B and another study was done on migrants from country A in country C, the index score can be used to compare and contextualize migrant integration in countries B and C.
- Both answers to the individual questions and the index score can be compared over different groups in your survey. For example, it can be tested if the answers differ by gender, race, or age.

Requirements for cross-sectional surveys:

- Field access to a sample of at least 100 respondents.
- If you want to compare different groups, each group should have at least 50 respondents. Note that if you want to cross-tabulate two groups (for example gender and race), each cell should have at least 50 respondents.

Limitations of cross-sectional surveys:

- A cross-sectional survey allows no insights into the development of integration over time, or if a certain programme has helped foster integration among programme participants. Therefore, they are unable to determine cause and effect.
- In a cross-sectional survey, there may be differences between cohorts, which can lead to potential biases.

Longitudinal surveys

Repeated data collection over a longer period of time offers more insights than a cross-sectional survey. Such longitudinal surveys can deliver insights into how integration changes over time and how it is affected by different other factors. There are different types of longitudinal surveys. For example, panel or cohort surveys. Cohort studies repeatedly survey people identified by specific characteristics in a defined time period (e.g. a student cohort), whereas panel studies repeatedly survey a sample that aims to represent the whole population.

Typical insight from longitudinal surveys:

- All insights that can be generated from one-off surveys but for more than one point in time.
- Longitudinal data show how overall integration or answers to individual questions change over time for the same respondent or sample.
- Longitudinal data show how the change in integration differs between groups or how the change in the index score compares to data from other studies.

Requirements for longitudinal surveys:

- Repeated field access to the same sample of at least 100 respondents.
- If you want to compare different groups, each group should have at least 50 respondents. Note that if you want to cross-tabulate two groups (for example gender and race), each cell should have at least 50 respondents.
- For longitudinal surveys, mak sure you collect the right contact information for the respondents (primary contacts) and ask them to provide the contact details of someone else in their household or someone who would know how to reach them in case they change phone numbers (secondary contacts).

Limitations of longitudinal surveys:

- Panel surveys are usually seen as the "gold standard", even though contacting the same migrants is often challenging, as they move, change their phone number (if any), and so on. Always try to aim for a panel survey if you can.
- Longitudinal surveys are a requirement for sound programme evaluation. However, whether data from such surveys can be used for programme evaluations depends on several other factors (see below).
- Longitudinal surveys depend on repeated answers from the same respondents. It is extremely important to motivate respondents to answer every wave of a panel survey. Attrition (respondents not answering to all waves of a panel survey) severely limits the value of panel surveys.
- Because longitudinal surveys involve multiple variables and repeated answers, they tend to take longer and cost more money and resources.

RESOURCES

Where to learn more about surveys:

There are several handbooks that can give you a good introduction to surveys in social research; a place to start could be *Surveys In Social Research* by David De Vaus.

There are also some good online resources. For example, Evidence in Governance and Politics (EGAP, 2020) offers concise guides on how to design surveys for randomized controlled trials:

10 Things to Know About Survey Design.
10 Things to Know About Survey Implementation.

3.2.3. Basic evaluation designs

The survey designs mentioned above are fundamental to evaluate programmes or interventions. Asking respondents about their experience with a given programme or about their opinion on the programme is a first step towards programme evaluation. **The designs presented below, however, seek to answer the question how a programme changed a respondent or, more formally, how a respondent compares to their hypothetical self, had they not taken the programme.** The first step towards such an evaluation is to define an outcome of interest. If a programme is designed to ease labour-market integration, employment could be a possible outcome of interest. For language classes, language ability would be an obvious outcome. The IPL-12 and IPL-24 indices both offer scales for such particular outcomes of interest (i.e. the scales for the different dimensions of integration) as well as a measure of overall integration (i.e. the index score). Once an outcome of interest is defined, there are different ways to estimate the effect a programme has had on that outcome.

A description of the different techniques are found in the following table.[2]

Methodology	Description	Who is in the comparison group	Required assumption	Required assumptions	Drawbacks
Pre-post	Measure how programme participants improved (or changed) over time.	Programme participants themselves – before participating in the programme.	The programme was the only factor influencing any changes in the measured outcome over time.	Before and after data for programme participant.	The difference between before and after data is influenced by everything that happened between the two surveys (e.g. economic upturns or downturns, high/low seasons, political events). Thus, the estimated programme effect is very likely biased.
Simple difference	Measure difference between programme participants and non-participants after the programme is completed.	Individuals who didn't participate in the programme (for any reason), but for whom data were collected after the programme.	Non-participants are identical to participants except for programme participation, and were equally likely to enter the programme before it started.	"After" data of the before-and-after scenario for programme participants and non-participants.	The estimated effect is influenced by all factors that explain why some respondents took part in a programme and others did not (e.g. age, prior participation, employment). Thus, the estimated programme effect is very likely biased.
Difference-in-differences	Measure improvement (change) over time of programme participants relative to the improvement (change) of non-participants.	Individuals who didn't participate in the programme (for any reason), but for whom data were collected both before and after the programme.	If the programme didn't exist, the two groups would have had identical trajectories over this period.	Before and after data for both participants and non-participants.	1. Usually requires more coordination and targeted sampling than the other designs. 2. Requires advanced econometric methods. 3. The estimated effects can still be biased by other factors that are correlated with programme participation and vary over time.

[2] This table was adapted from the Chapter 4 of IOM M&E Guide (p 139–140), which was adapted from the J-PAL Impact Evaluation Methods table.

Methodology	Description	Who is in the comparison group	Required assumption	Required assumptions	Drawbacks
Multivariate regression	Individuals who received treatment are compared with those who did not, and other factors that might explain differences in the outcomes are "controlled" for.	Individuals who didn't participate in the programme (for any reason), but for whom data were collected both before and after the programme. In this case, data are not comprised of just indicators of outcomes, but other "explanatory" variables as well.	The factors that were excluded (because they are unobservable and/or have not been measured) do not bias results because they are either uncorrelated with the outcome or do not differ between participants and non-participants.	Outcomes as well as "control variables" for both participants and non-participants.	The risk of missing-variable bias (i.e. not all variables that could have changed the outcome are included in the regression model) is a primary concern with this approach. On the other hand, if many variables are included in the model, sample size and statistical power become an important concern. An advanced understanding of econometric methods is neccessary to run the analyses and to understand all potential trade-offs and pitfalls.
Statistical matching	Individuals in control group are compared to similar individuals in experimental group.	Exact matching: for each participant, at least one non-participant who is identical on selected characteristics. Propensity score matching: Non-participants who have a mix of characteristics, which predict that they would be as likely to participate as participants.	The factors that were excluded (because they are unobservable and/or have not been measured) do not bias results, because they are either uncorrelated with the outcome or do not differ between participants and non-participants.	Outcomes, as well as "variables for matching" for both participants and non-participants.	The approach crucially depends on the availability of good matches (i.e. non-participants who are very similar to participants). An advanced understanding of econometric methods is neccessary to match observation, run the analyses, and to understand all potential trade-offs and pitfalls.
Regression discontinuity design	Individuals are ranked based on specific, measurable criteria. There is some cut-off that determines whether an individual is eligible to participate. Participants are then compared to non-participants and the eligibility criterion is controlled for.	Individuals who are close to the cut-off, but fall on the "wrong" side of that cut-off, and therefore do not get the programme.	After controlling for the criteria (and other measures of choice), the remaining differences between individuals directly below and directly above the cut-off score are not statistically significant and will not bias the results. A necessary but sufficient requirement for this to hold is that the cut-off criteria are strictly adhered to.	Outcomes, as well as measures on criteria (and any other controls).	1. Relies heavily on the existence of a sharp eligibility threshold. 2. Requires many respondents who are very close to the threshold. 3. Requires advanced econometric methods.
Instrumental variables	Participation can be predicted by an incidental (almost random) factor, or "instrumental" variable, that is uncorrelated with the outcome, other than the fact that it predicts participation (and participation affects the outcome).	Individuals who, because of this close-to-random factor, are predicted not to participate and (possibly as a result) did not participate.	If it weren't for the instrumental variable's ability to predict participation, this "instrument" would otherwise have no effect on or be uncorrelated with the outcome.	Outcomes, the "instrument," and other control variables.	Finding a factor that predicts participation but does not at all affect the outcome of interest is the most challenging part of this approach. An advanced understanding of the case at hand and of econometric methods is neccessary to evaluate the quality of the instrument.

Methodology	Description	Who is in the comparison group	Required assumption	Required assumptions	Drawbacks
Randomized evaluation	Experimental method for measuring a causal relationship between two variables.	Participants are randomly assigned to the control groups.	Randomization "worked." That is, the two groups are statistically identical (on observed and unobserved factors).	Outcome data for control and experimental groups; control variables can help absorb variance and improve "power".	1.Usually requires a lot of coordination and very targeted sampling. 2.Requires advanced econometric methods.

RESOURCES

There are many other designs that combine features of the presented designs or were tailored to address context-dependent specifics. For example, the study in Peru used a **stepped-wedge trial.** This is basically a randomized controlled trial in which the treatment group is not treated all at once. Instead, randomly selected subgroups of the treatment group are treated successively in well-controlled intervals. This design can be used if a lack of capacity prevents the research team from treating the whole treatment group all at once, if the whole sample should eventually receive the treatment.

Please see Chapter 4 for more information on how the index was piloted in Peru.

ADDITIONAL RESOURCES

BetterEvaluation
n.d. Compare results to the counterfactual.

Gertler, P.J., S. Martinez, P. Premand, L.B. Rawlings and C.M.J. Vermeersch
2011 *Impact Evaluation in Practice*. World Bank, Washington, D.C.

International Fund for Agricultural Development (IFAD)
2002 Annex D: Methods for monitoring and evaluation. In: *Managing for Impact in Rural Development*: A Guide for Project M&E. Rome.
2014 Republic of India, Impact Evaluation of Jharjkhand – Chhattisgarh Tribal Development Programme (JCTDP). Approach Paper. Independent Office of Evaluation.
2016 *Evaluation Manual*. Second edition. Independent Office of Evaluation of IFAD, Rome.

Leeuw, F. and J. Vaessen
2009 *Impact Evaluations and Development: NONIE Guidance on Impact Evaluation*. World Bank, Washington, D.C.

United Nations Evaluation Group (UNEG)
2013 Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management. Guidance document.

US Department of Health and Human Services, Centers for Disease Control and Prevention (CDC)
2018 Data collection methods for program evaluation: Observation. Evaluation Brief no. 16.

White, H. and D. Raitzer
2017 *Impact Evaluation of Development Interventions: A Practical Guide*. Asian Development Bank, Metro Manila.

White, H. and S. Sabarwal
2014 Quasi-experimental design and methods. Methodological Briefs, Impact Evaluation No. 8. UNICEF, Florence.

White, H., S. Sabarwal and T. de Hoop
2014 Randomized controlled trials (RCTs). Methodological Briefs, Impact Evaluation No. 7. UNICEF, Florence.

3.3. Local adaptation of the index

The questions of both indices (12 and 24) were developed to be deployed among different target populations around the world. Despite this high level of generality, some elements need to be adapted whenever the index is used in a new location. Most studies will have to review and change the host country name, local language, income brackets, and employment categories. While making these changes, it is important that the index score should not be biased by changes to these elements of the survey. In addition to that, there might also be cases in which questions need to be adapted or removed for reasons not foreseen by the original research team. This section will go over the different aspects that need to be considered before using the index in a new context.

3.3.1. IPL-12 versus IPL-24

The first question that needs to be decided is whether to use the short form, IPL-12, or the long form, IPL-24. Both indices were developed with the goal to create a reliable and valid measure of integration that can be compared across different countries and migrant groups. To do this, both indices produce a single measure of integration that falls somewhere between 0 and 1. As the IPL-12 is a subset of the longer IPL-24, it is not surprising that the resulting scores are highly correlated (see section E in the supplementary information of the PNAS paper). Considering the high correlation between the two indices and that the median response time of the IPL-24 (seven to eight minutes) is more than twice as long as the median response time of the IPL-12 (two to three minutes), it is usually sufficient to use the IPL-12 if the goal of the study is to work with the final index score. However, if a certain integration dimension is of special interest and should be measured with more precision, the additional IPL-24 questions from the respective dimension can be added to the study. In this case, the additional questions can be used to evaluate integration on their respective dimension but should not be used to compute the overall index score.

3.3.2. Standard adaptations

There are six elements that might need adaptation every time the index is used in a new context. This chapter will discuss each in detail.

Country name

All questions on psychological and political integration refer to the country a respondent lives in (e.g. How connected do you feel with the United States?). Here, the country name needs to be changed to the country the respondents currently live in. In some cases, it might not be appropriate or feasible to use a country name (in the case of autonomous regions or breakaway republics, for example). To find an appropriate alternative, it is important to consider the aim of the questions. Questions on psychological integration should measure a respondent's sense of belonging to the larger political entity they live in. Questions on political integration should measure a respondent's knowledge about the political system that holds political power over the individual. Considering these aims, local research teams can find an appropriate alternative to using a country name.

Referring to the local non-migrant population

All questions on social integration refer to the rooted local population by their nationality (e.g. In the last 12 months, how often did you eat dinner with Americans who are not part of your family?). The decision to just refer to a nationality was made due to a trade-off between simplicity and precision. In the example above, "American citizens", "Americans born in the United States", or "rooted Americans" would have been some of many possible alternatives. However, tests showed that more complex terms are a lot more challenging to understand for many respondents and that the interpretation of the term "Americans" does not differ a lot from the interpretation of the other terms. Hence, using the nationality of the country a respondent lives in is sufficient in most cases. However, in regions in which nationality carries a very specific meaning, alternative phrases should be considered. A good alternative phrase should relate to the rooted local population and should be easy to understand and interpret by all respondents.

Local language

All questions on linguistic integration refer to the local official language. For example, the IPL-12 questions on linguistic integration focus on a respondent's ability to read and speak in that language. Reading was deemed important as migrants are very likely to receive documents and legal information in writing. Speaking was deemed important as it captures how well migrants can articulate their needs and because speaking is often an important step to transferring professional skills to a new context. In cases where there are two or more official languages, the following is an important consideration that can be taken into account to guide the decision regarding which language to use in the survey. The language used in the survey should be the language migrants will most likely need to understand important legal information, to speak to authorities, and/or to succeed in the workplace.

Income

Answers to the question on household income are scored in relation to the local household income. The original index uses the median gross annual household income. In theory, the median net annual household income, or the median gross/net monthly household income could also be used. The decision regarding which income figure to use should depend on two factors. First, which figure are respondents more likely to know? To answer this question,

it helps to know how income is usually discussed in the respective country. Would respondents talk about their salary in terms of gross or net salary? When discussing a new job, would they rather know their monthly or their annual income? Second, how accessible are the data regarding nationally representative medians for the respective type of income? To make a decision on which income type to use, research teams should find a balance between these factors. Most importantly, the type of income used for the questions needs to be the same type of income that is used as a reference for scoring. This means, if the question asks for a respondent's net monthly income, the scores should be assigned in relation to the median net monthly income.

Employment

The second question on economic integration asks for a respondent's employment status over the last four weeks and lists several options. Some of them might be implausible in a particular context and could be removed from the question. Likewise, there might be plausible options that are not on the list. Such options can be added to the list, if they can be assigned to one of the three values that are used for scoring. A score of 1 is assigned if a respondent is unemployed and has no other daily activity that keeps the individual from working. A score of 3 is assigned if a respondent is in unpaid work (e.g. domestic care work), is in school or another educational programme, or if a respondent is not employed due to sickness, disability, or old age. A score of 5 is assigned if a respondent is employed or earns an income in a different way (e.g. in military service).

Household

The term "household" is ambiguous and can be interpreted differently from context to context. For the index, the term should refer to a group of individuals who share family links and who reside at the same place and routinely share their living expenses and other financial responsibilities. To account for the fact that especially recent migrants often live in larger shelters, the original index describes this concept to respondents as "everyone with whom you share an apartment or house and with whom you are also related by birth, marriage, partnership, or adoption." For every new study, the research team should review if this description makes sense in the respective context or if the description should be slightly altered and should make sure the definition used is understood by the respondents.

What not to change

In some cases, the research team might be certain that responses to a particular question will show little to no variation. For example, it is very unlikely that migrants who just arrived days or weeks ago will be employed. In these cases, it is tempting to omit or modify the question. However, even if there is no variation in the answers, not being employed is still crucial information to understand the level of integration. Furthermore, the question is still necessary to compute the standardized index score.

A similar issue came up during the survey in Brazil. As shared meals are very common in Brazil, it was assumed that the responses to the respective question would all cluster at the top of the scale. However, if this is an accurate reflection of how migrants are integrated in Brazilian society, and if it can be assumed that these shared meals are accompanied by conversations, the question still measures an important part of social integration that should be reflected by the index. Here, ad hoc changes to individual questions make it harder to compare index scores between surveys and will likely introduce new measurement error.

3.3.3. Dropping or adding questions

In some situations, a project team might consider omitting or adding questions to the IPL Integration Index. This section will briefly discuss both possibilities.

Generally, omitting questions from the IPL-12 is discouraged as it limits the comparability of the index score and because it limits the research team's understanding of the respective dimension of integration. However, there might be situations in which a question does more harm than good. If a research team fears that a question might make respondents suspicious, less likely to answer other questions truthfully, or that a question is too personal or sensitive for most respondents, the overall data quality is likely to improve if that particular question is omitted.

For example, in politically oppressive countries or circumstances, many respondents might see questions on political integration as too sensitive or that answering such questions might even put respondents at risk of persecution. In such cases, it is possible to omit a question or a whole dimension (two questions) from the IPL-12. As mentioned above, simulations with IPL-12/24 data from different datasets show that the overall score changes little if one or two questions are removed. If more than two questions are removed, however, the final scores differ significantly. While it is acceptable to omit a question if that question would put respondents or the study at risk, questions should not be omitted because the project team expects low or no variation in the answers. For example, the question on employment should not be omitted because a survey's target population is not allowed to work. Likewise, questions on political integration should not be omitted because a survey's target population has no right to political participation. In these cases, low index scores on economic or political integration should be seen as important information on how the institutional and political context prevents integration. Hence, it is important to include such questions even if the answers seem largely predictable.

Adding questions to the index: key considerations

The IPL Integration Index is designed to produce comparable measures of different dimensions of integration in many different situations. However, the index was not designed to capture all possible dimensions and aspects of integration. Accordingly, a project team might want to add questions to their survey that capture aspects that are of particular importance to the study at hand. This is generally unproblematic, if two issues are considered. First, answers to additional questions should not be considered when calculating the overall index score. The overall index score is used to compare integration between different groups, studies or contexts. Therefore, the index score should always be calculated from the same set of questions. Second, added questions should not ask for the same concepts as the IPL Integration Index questions. For example, a second question on income or employment will make it less likely that a respondent answers both questions. Thus, adding questions that are very similar to IPL Integration Index questions is likely to jeopardize the overall data quality. More generally, it is also advisable to only use pre-tested questions. Such questions can be found in the Global Survey Bank on Migrant Integration and Social Cohesion or other surveys on migrant integration, where the questions have been previously vetted.

3.4. Pilot testing and data collection

To test the validity and relevance of the study design and the locally adapted Integration Index, a two-step pilot testing activity is encouraged, if circumstances and resources allow. The overall goal of a pilot is to simulate the environment of the actual survey and test not only the content but also the logistics of the conduct of the survey. The pilot allows the project team to identify how well the respondents understood the questions, and if they were able to answer them with ease.

The **pre-pilot stage** serves to test the more sensitive questions and controversial questions of the index. It can also serve to test the more tricky questions to translate to the local language and adapt some of the questions to the current local realities (e.g. income). A reduced set of questions can be asked to a smaller group of respondents.

Based on the lessons learned from the pre-pilot, the subsequent **pilot stage** will test the overall questionnaire with a bigger set of respondents that mirror the sample. The aim of the pilot is to also get a sense of the response rate and the overall logistics of carrying out the data collection, in accordance with the chosen design of the study.

The interview during the pilot testing typically consists of four parts: introduction and consent, the IPL Integration Index, a set of feedback questions to respondents, and a set of feedback questions to enumerators.

In rolling out the index in the LAC region, these pilot testing initiatives gave the opportunity to change things, improve the questionnaire and optimize the logistics before collecting the actual data. There are usually changes made after a pilot. If this is not the case, the data collected during the pilot might also be used for the final data analysis. The pre-pilot and the pilot were crucial stages for establishing lessons learned to be later applied during the larger data collection phase.

RESOURCES

Please click here for a suggested template of the pilot testing.

3.4.1. Hiring of a survey firm: roles and responsibilities

In the absence of in-house support for an enumerator-administered survey, the project team should ensure that a survey firm is contracted and eventually briefed on the IOM Data Protection Principles. The roles and responsibilities of the survey firm may involve, but are not limited to the following:

- Deploy enumerators who will administer the survey.
- Translate the questionnaire(s).
- Carry out training for the enumerators to familiarize them with the process, software and survey form
- Provide materials and logistical means. In the event of telephonic, or in-person surveys, provide tablets and phones for enumerators, as well as shoulder mobile and internet costs, among others.
- Communications, outreach and follow-up with target respondents if there is no answer, ensuring anonymity of respondents and keeping in line with instructions during the training.
- Conduct interviews in accordance with the chosen software (for example: Qualtrics).
- Review all the data gathered to ensure that it is complete and accurate.
- Facilitate the distribution of incentives (if needed).
- Make referrals (if needed) according to the enumerators, referral guide that will be provided by the implementing mission.

TIPS

- Research on reputable survey firms that have previously been engaged with IOM and other United Nations (UN) agencies so they may already have some familiarity with the UN procedures and regulations on data collection and handling vulnerable populations.
- The survey firm should directly coordinate and agree a workplan for the data collection with the project team.
- The project team and survey firm should upate each other throughout all the steps of the data collection, ideally on a weekly basis.
- If possible, the project team, with or without the help of researchers, should run frequent quality checks on the data during data collection. In this case, make sure you have access to the data collected on an agreed-upon frequency and run your analysis (survey length by enumerators, percentage of "Don't know/Refuse to answer", variations in the answers, etc.)
- Ask the survey firm to send audio recordings of one interview per enumerator (without the respondent's voice) for the pilot, at the start of data collection and/or at any point in time during the data collection.
- If possible, meet the survey firm and the enumerators in person and agree on the content and mechanics of the enumerators' training.

3.4.2. Training of enumerators

In the event of an enumerator-administered survey – as opposed to a self-administered survey – a training session for enumerators is essential before any data collection activity will take place. This will ensure that enumerators are familiar with the questionnaire, and they are able to properly address questions that respondents may have, motivate them to give complete and honest answers, emphasizing cultural and gender sensitivity throughout the contact with the respondents. This training also aims to ensure that enumerators familiarize themselves with the questionnaire, survey protocols, target respondent, Qualtrics and the devices (phone and tablet) that are used to conduct the interviews.

In rolling out the index in the LAC regional context, the contracted survey firm, with the support from IOM, organized a training session for its enumerators before the pilot testing and actual data collection. IOM conducted the training, went through the questionnaire and most sensitive questions, and answered questions that arose during the training and that had not been already addressed in the enumerators' guide. It was also a good opportunity for enumerators to ask all the questions they had prior to and during the training. Given that the respondents were beneficiaries of IOM programming in vulnerable situations, the training included a section on how to approach vulnerable populations and case management/handling referrals.

RESOURCES

Please click here for a sample enumerators and referrals guide.

TIPS

- Before the training, the survey firm should be provided with an accompanying practical guide for enumerators which provides some tips for conducting the survey and clarifications.
- The training should include topics on, inter alia, the importance of obtaining consent and data security, how to command the chosen survey platform (e.g. Qualtrics), dealing with vulnerable populations and case management and referrals.
- The training and guide for the enumerators should include a list of contacts/organizations to refer respondents in need.
- Prepare a referral guide with (i) common types of problems faced by migrants (type of violation of rights, legal and migratory orientation); (ii) contacts of organizations assisting migrants in each state; (iii) frequently asked questions; and (iv) to indicate a focal point within the project team for the beneficiaries to reach out to in case of emergency.

3.4.3. Data collection

The IOM Data Protection Principles should be considered throughout the data collection process, including the storage, sharing and usage of collected data. At the outset, it is essential to agree on a timetable for data collection, according to the broader research workplan and the study design, as IOM needs to ensure the availability of cases for the survey firm to carry out the planned surveys according to the schedule.

Once the research design has been identified and the local adaptation of the index, after the pre-pilot and the pilot have been developed, the data collection can start. Below are some key recommendations to follow during the data collection phase:

- To safeguard the data collected and prevent leaks of sensitive information, it is recommended to assign each respondent a unique identifier. This ensures that the data cannot be traced back to them if the data is disclosed to other parties. Likewise, the data collected should be stored in a safe place with accurate access controls.
- To facilitate the data cleaning and the data analysis, the use of online survey tools (e.g. Qualtrics) is strongly advised to collect data. These platforms make it easy to download and transfer the collected data to statistical programmes for data analysis. In the case of paper-based surveys, the data collected need to be transferred onto a computer application first, such as Excel.
- Data need to be accurate, valid, reliable, timely, relevant and complete. Hence, data quality and data cleaning are needed once the data is collected. Data errors can occur at different stages of the design, implementation and analysis of data and need to be screened for diagnosis and treatment. Some key errors to look for when screening data are: spelling and formatting irregularities, lack of data (questions with fewer answers than surrounding questions), excessive data (duplicate entries, more answers than originally allowed), outliers/ inconsistencies (values that are so far beyond the typical distribution that they seem potentially erroneous), strange patterns that suggest cheating rather than honest answers, suspect analysis results (answers seem counter-intuitive or extremely unlikely).
- When the problematic observations have been identified, these need to be treated before the data can be analysed. Key approaches to dealing with data errors could be: leave the data unchanged (accepting the erroneous data as valid responses and make no changes), correct the data (without modification of the intention or meaning given by the respondent), delete the data or re-measure the suspect/erroneous values, if time and resources permit.

TIPS

- The survey firm and the project team involved should monitor and follow up the progress of the data collection and have continuous communication between them before, during and after the data collection, ideally on a weekly basis. Since the survey firm may not have access to the data collection programme, IOM and the research partners involved should collaborate to match the reported number of surveys by the survey firm with what is observed in the programme.

- Assign one or two staff who will monitor and check the number of completed surveys per week, as well as detecting missing forms, double responses for a respondent, unfinished surveys, survey duration and responses out of the expected range (outliers), and so on. Having this weekly check-in can reduce the amount of errors during the data collection.

- Other types of common errors can occur during data collection. For example after the survey, the enumerator can forget to press the "finish" button. Therefore, the survey is not saved as completed. It is important to alert and spread common errors found with all enumerators to reduce and prevent them.

- A good practice for quality-control methods are audio recordings of the interviews by enumerators (without the respondent's voice) during the data collection.

- The possibility of unexpected circumstances should be considered, which may require a modification and/or a prolongation of data collection.

3.5. Data analysis and interpretation

3.5.1. What to look for in the data

If IPL Integration Index data were collected as part of a carefully designed evaluation, the study design will usually guide the way the data are analysed. However, even in the absence of such a design, or in addition to such a design, there are several different ways to make sense of the data.

Using data from additional questions allows for interesting group comparisons.

Do integration levels differ by age or gender? How do race or legal status correlate with integration? For example, a survey of more than 5,000 Venezuelan migrants in Peru showed somewhat higher index scores for men (0.56) than for women (0.51).

- Overview of integration scores
  - Integration score by gender
  - Integration score by age
  - Integration score by household size
  - Integration score by income
- Educational attainment
  - Integration score by educational attainment and gender
  - Integration score by educational attainment and income
- Employment
  - Integration score by employment and gender
  - Integration score by employment and income
  - Integration score by employment situation
- Length of stay
  - Integration score by length of stay and gender
  - Integration score by length of stay and income
- Immigration status
  - Integration score by migratory status and gender
  - Integration score by migratory status and income
- Migration intentions
  - Integration score by migration intentions and gender
  - Integration score by migration intentions and income
- Discrimination
  - Integration score by level of discrimination and gender
  - Integration score by level of discrimination and income
- Negative coping strategies
  - Integration score by negative coping strategies

Comparing integration scores over the six dimensions will deliver insights into how even or uneven integration proceeds in the different dimensions. A further comparison between different groups can be even more insightful. It is very likely that employed migrants have a higher income, but do they also show higher index scores for social integration? Do migrants who migrated with their families develop a stronger sense of belonging than unaccompanied migrants?

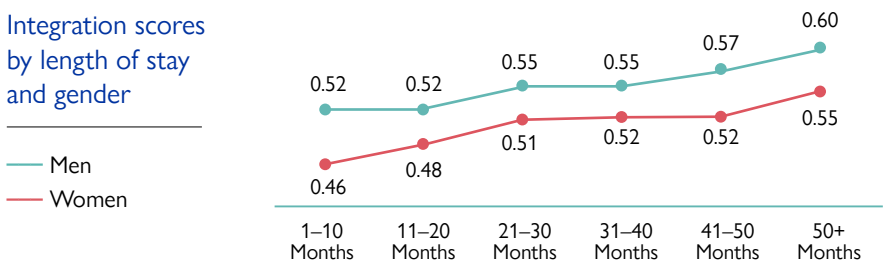
How the index can show uneven integration process: the example of Peru

For example, the survey of Venezuelan migrants in Peru showed a considerable difference between psychological and navigational integration. The average index score for psychological integration was 0.59 and the average index score for navigational integration was only 0.35. In general, the results show uneven integration processes across dimensions and gender. The index showed low scores for navigational integration for men (0.37) and women (0.35). Other dimensions, however, showed considerable differences between men and women. Social integration, for example, was much higher for men (0.53) than for women (0.41).

Repeated surveys allow group and dimension comparisons over time. Such comparisons can start with a simple question such as "does integration increase over time?" As a follow-up step, time trends can be compared between groups. Here, a research team could ask if integration increases faster among young migrants as compared to older migrants. Even more detailed are time-group-dimension comparisons. A research team could analyse if legal status influences economic integration differently than social integration.

How the index can show increasing integration outcomes over time

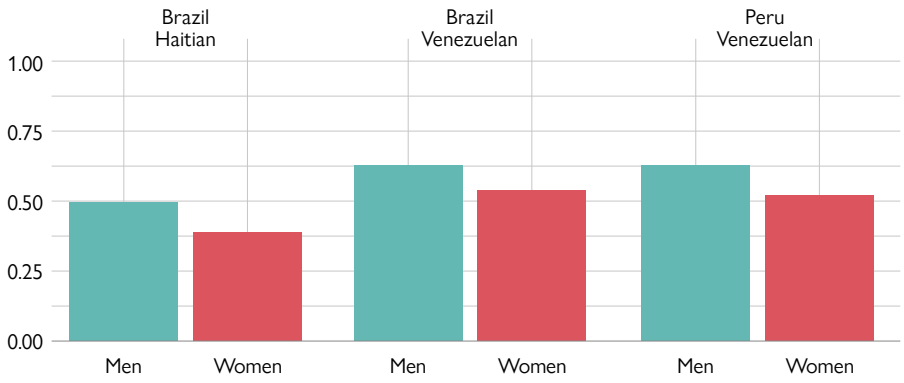
While the data from Peru showed different integration scores between men and women, integration over time was similar for both groups. The average integration index score for migrants who had been in Peru for 50 months or more was 0.08 (men) and 0.09 (women) higher as compared to recently arrived migrants.



The standardized format allows comparisons with data from other studies.

If the study at hand collected data on migrants from country A in country B and another study collected data on migrants from country A in country C, **and if the two samples are comparable**, researchers can compare the integration levels of migrants from country A over different, contexts. For example, the two populations of Venezuelan migrants in Peru and Brazil are quite different, which make comparisons difficult. The average integration score for Venezuelan men in Peru was 0.56, while it was 0.61 in Brazil. Integration index scores for women, on the other hand, were similar in both countries.

Figure 1: Integration Index scores of Haitian and Venezuelan migrants in Brazil and Peru



3.5.2. First steps of the analysis

There are many ways to analyse the IPL Integration Index data. In most cases, there are more ways to compare and cross-tabulate the data than researchers can reasonably interpret and track. It is therefore advisable to pre-define the most important or pressing questions. This selection is very context specific. However, there are a few ways to identify potentially more relevant questions.

First, comparisons between larger groups in your data are less likely to be biased by single extreme values than comparisons between smaller groups. Furthermore, findings that pertain to larger groups are more likely to be of substantial value to other researchers and practitioners than findings that pertain to very small and specific groups.

Second, what is true for large groups is also true for large differences. Large differences between groups or over time are less likely to be biased by sampling or measurement errors than small differences. Thus, it is advisable to start an analysis with comparisons that are expected to show substantial differences.

Third, to produce actionable insights from your data, focus on factors that can be observed and manipulated by local missions. For example, there is a certain correlation between education and income in many societies and in most cases, which holds for migrants and non-migrants alike. It will be very hard for a local mission to address this phenomenon directly. However, if legal status and linguistic integration are systematically related, it might be possible for a local mission to provide specialized language classes or to lobby for targeted changes in migration law.

3.5.3. Methods

Simple bar charts for group comparison and line charts for time trends are usually the first step in data analysis. Cross-tabulations can be used to analyse the intersection of two different groups (e.g. gender and age). Some exemplary R code for aggregating, comparing, and visualizing data from the IPL Integration Index can be found in the Annex, online in the Harvard Dataverse (Harder, 2018) and on the project homepage.[3] It is very unlikely that two groups have the exact same score. Measurement error and random chance will often lead to at least small differences even when comparing very similar groups. Different statistical tests can be used to discern random differences from systematic (i.e. not random) differences. In most cases, a simple t-test will be enough to check if group differences are likely to be systematic (statistically significant) or due to chance. The mathematical routines behind a t-test are implemented in statistical software such as R or STATA and can also be computed in Microsoft Excel. There are many dedicated online resources on how to apply and interpret these techniques. For example, the Evidence in Governance and Politics (EGAP, n.d.) network offers a number of short methods guides that can help you design or understand your data analysis better. Here you can find guides such as "10 Things to Know About Reading a Regression Table" or "10 Things to Know About Hypothesis Testing". Furthermore, several guides have been translated to French and Spanish.

[3] www.integrationindex.org

4. PILOTING A TOOLKIT THAT OFFERS A MULTIDIMENSIONAL PERSPECTIVE ON MIGRANT INTEGRATION IN THE LATIN AMERICAN AND THE CARIBBEAN REGION

Over the past decade, the Latin America and Caribbean region has experienced the largest intraregional migration movements in recent history, mainly due to the displacement of over five million Venezuelans, as well as the rise in large-scale movements from Central America to South American countries (IOM, 2022a). As migrants settle into their host communities, governments and organizations working on the ground have been challenged with devising plans to help integrate them into society. In addition to there being little knowledge on the broader integration experience of migrants, there is also a lack of empirical tools that measure their assets, knowledge, and capacities to establish themselves in the various spheres of community life.

IOM and its academic partner, the Immigration Policy Lab (IPL) at ETH Zurich, came together to roll out this project, initiated in 2020. In 2018, IPL developed the Integration Index, a measurement tool which takes a multidimensional approach. The measurement tool captures six key dimensions to facilitate the integration of migrants into new environments: psychological, navigational, economic, social, linguistic, and political integration.

Source: IOM, 2019 Return and Reintegration Key Highlights, 2020.
Note: This map is for illustration purposes only. The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the International Organization for Migration.

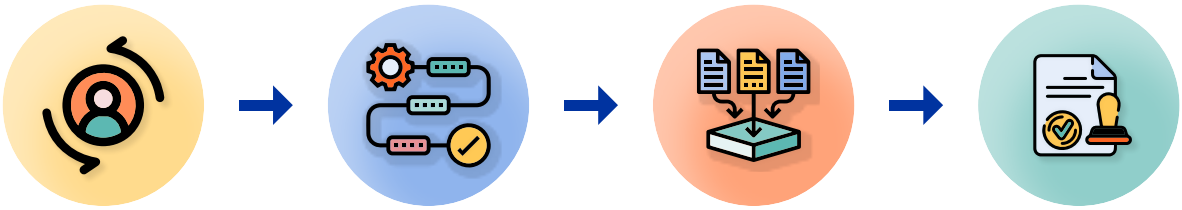
4.1. Context of piloting the measurement tool in the Latin American and the Caribbean region

Migration in the Latin America and Caribbean (LAC) region has been traditionally marked by extraregional and intraregional flows. However, in recent years, the trend in the region has increasingly changed in terms of the direction of migration flows, intensity and composition of migrants. More and more, intraregional and mixed-migration flows (labour migrants, irregular migrants, asylum-seekers, refugees and other types of migrants) have become dominant within the region. One clear case is the presence of 7 million Venezuelan migrants globally (as of September 2022), of which nearly 80 per cent are in the Latin America and Caribbean region (R4V, 2022). This shift, which is exacerbated by the larger influx of migrants within the LAC region, has adversely affected a number of factors including host communities' capacities, resources and perceptions, as well as migrants' broader integration experience.

While governments in the region have notably progressed with their respective national strategies and plans to respond to these changes, there remains limited comparative data to understand under what conditions migrants are able to establish themselves as new members in their host communities and which of the existing migration policies and interventions need to be further adjusted and strengthened. Considering the limited comparative data and lack of empirical tools, the IPL Integration Index offers new prospects on how to examine integration outcomes and offer new approaches, evidence, and partnerships to facilitate successful inclusion of migrants in their new communities.

4.1.1. Methodology of piloting the toolkit

The development and application of the toolkit followed a multistep operational process that built upon existing resources and knowledge base on integration, involving key internal and external stakeholders in the overall implementation of the project.



SURVEY DEVELOPMENT AND ADAPTATION

- Launch and stakeholder consultations
- Review of existing literature and tools
- Local adaptation of the IPL Integration Index
- Development of methodology and practical guidance materials
- Translation and coding of surveys.

PILOT TESTING

- Conducting of two-step pilot testing activities
- Incorporation and analysis of feedback from respondents and enumerators
- Iterative process of question writing, empirical testing and data collection
- Finalization of questionnaires and scoring rule.

ACTUAL DATA COLLECTION AND PROJECT-LINKED EVALUATION

- Conduct of population-based surveys in Brazil and the Dominican Republic
- Roll-out of a causal impact evaluation in Peru
- Data analysis and interpretation
- Development of country reports and extraction of programmatic and policy recommendations.

VALIDATION AND CAPACITY-BUILDING INITIATIVES

- Conduct of validation workshops with government and other relevant partners
- Technical assistance to stakeholders for application and replication of the survey.

At the outset of the project, IOM successfully forged a partnership with the Immigration Policy Lab, a research group with branches at the University of Stanford and ETH Zurich. Through this cooperation, IPL's existing Integration Index was utilized as the basis for the development of the multidimensional toolkit that grew from this initiative and was applied in the Latin American and the Caribbean region.

Prior to the roll-out of the toolkit in the LAC region, the index had already gone through an iterative process of question writing, empirical testing, and statistical validation tests through IPL's channels. This validation process involved several rounds of major revisions based on workshops with experts, pilot surveys of various immigrant samples, and testing of over 200 questions through close to 4,000 interviews, which were then statistically interpreted and analysed. However, up until that point, the validation and testing process had been mainly concentrated in North America and Europe, leaving out other contexts such as the LAC region that were also facing large migratory flows.

In an effort to expand the toolkit's reach and continue testing its validity given the limited integration data that had been collected in the LAC context, the IOM offices in Peru, the Dominican Republic and Brazil were engaged to lead the roll-out of the index among their migrant beneficiaries.

In close coordination with the implementing missions and partners, an extensive review and mapping of existing international and national surveys, IOM tools, datasets and relevant academic literature were conducted at the outset to adapt the original IPL Integration Index to specific country contexts.

To test the validity and effectiveness of the measurement tool in the LAC context, a two-step pilot testing activity was implemented. The first step, known as the "pre-pilot," tested out the more sensitive questions of the index. Based on the findings and lessons learned from the pre-pilot phase, some questions were omitted, adjusted and further improved for the next phase of data collection activities. The following pilot phase was another trial run of the entire index before it was applied to larger-scale data collection activities. This also aimed at testing the logistics of carrying out the interviews, assessing the comfort level, ease and general experience of migrants in completing the questionnaires and overall relevance of the questions. In both the pre-pilot and pilot, a list of feedback questions (internal link only) (see Annex) was incorporated for both respondents and enumerators to answer at the end of the survey. All in all, a total of six pilot exercises (two in each of the three countries) were done.

Feedback of migrant respondents from pilot testing activities

The feedback from the two-phased pilot surveys and the results of the feedback questions from both the enumerators and respondents did attest that the toolkit was applicable to the LAC context, with many expressing their satisfaction with the survey, as evidenced through the pilot reports of implementing missions. The pilot testing initiatives allowed the team to adjust the survey questions (in terms of wording, choices, translation, enumeration, etc.) based on the respondents' feedback and general understanding of the questions. In some cases where respondents felt uncomfortable answering certain questions, these were then omitted from the final questionnaire. This is the case for example of certain political questions.

With the exception of some questions regarding political issues, the majority of the respondents had overall positive feedback about the questions and thought that the questions were useful, simple, and reflective of their personal situations. Most of the respondents understood all the questions and the purpose of the survey.

While it is not the intended goal of the pilot, many of the respondents expressed their appreciation for having been given the chance to be interviewed and to share their experiences, giving them a strong feeling that IOM cared about their situations and would look for solutions to address their needs.

Following an extensive review and analysis of the report findings from each pilot and feedback obtained from enumerators, respondents, missions and research partners involved, it was concluded that the toolkit was effective and valid for generating integration outcomes and measuring gender differences across all adult migrant groups. The index was then finalized and ready for larger-scale data collection and evaluation exercises. The project team used Qualtrics to conduct telephonic surveys in all three implementing missions.

The index was then used to carry out larger-scale data collection activities, applied in two measurement contexts among IOM project beneficiaries – the first being a causal impact evaluation in Peru and the second use being a population-based survey methodology in Brazil and the Dominican Republic. The causal impact evaluation of the effects of cash-based intervention on migrant integration outcomes in Peru – which entails applying the index in several interview waves among IOM Peru's cash assistance beneficiaries – is considered the first of its kind within IOM cash-bassed intervention (CBI) programming.

Within the framework of the project, over 10,000 surveys have been conducted.

4.1.2. Overview of the studies in Brazil, the Dominican Republic and Peru

The below table shows an overview of the methodology of the assessment exercise in the three implementing missions using the toolkit.

Overall objective

The overarching goal of the project is to contribute to a more robust and comparative evidence base for policy development and programme innovation to facilitate effective migrant integration.

The resulting toolkit aims to serve as a core integration measurement instrument which can eventually be replicated, scaled and adapted in different contexts by service providers and governments.

	BRAZIL	DOMINICAN REPUBLIC	PERU
Specific objectives in the application of the toolkit	• To capture the multidimensional integration outcomes and experiences of migrants according to key integration dimensions and variables of interest. • To help inform and tailor future programming, better target future beneficiaries, and influence broader migration policy considerations.		• To evaluate the effects of cash-based intervention on migrant integration outcomes of beneficiaries.
Respondents	Beneficiaries of IOM cash-based interventions	Beneficiaries of R4V (CBI, direct assistance, seed capital for economic recovery)	Beneficiaries of IOM cash-based interventions
Age and nationality	18+ years old, Haitian and Venezuelan migrants	18+ years old, Venezuelan migrants	18+ years old, Venezuelan refugees and migrants
Sampling frame and strategy, survey mode	• Use of existing IOM databases using random sampling until the target number is reached • Telephonic interviews conducted on Qualtrics		
Survey frequency	One-off survey to obtain descriptive and general understanding of integration outcomes		Several survey waves for a causal impact evaluation
No. Surveys conducted	900	900	8 500

All the logistical activities related to the data collection were carried out remotely given travel restrictions and physical distancing requirements brought about by the COVID-19 pandemic. All the surveys were conducted over the telephone, including the coordination between different partners throughout the data collection process.

The project team and enumerators ensured that the participation of migrants in the survey was voluntary, informed, confidential and would not affect their participation in any of IOM's interventions. All steps in the data collection, including the sharing and transfer process are in line with the IOM Data Protection Principles, IOM-IPL Local Cooperation Agreement and Mission-level Service Agreements with survey firms that includes a data sharing provision (see the IOM Data Protection Principles) to ensure data security and confidentiality.

As part of the implementation, the project teams of the implementing missions and IOM headquarters organized capacity-building activities in order to introduce and socialize the IPL Integration Index and provide assistance to stakeholders interested in the results and potential use of the toolkit. At the end of the data collection, analysis and report-writing periods, each mission organized a validation and capacity-building event to present the main findings that emerged from the piloting and launch the research report. In addition, the workshops are a unique opportunity to initiate or strengthen the dialogue between IOM and stakeholders such as representatives of the federal government, local government, civil society and academia, to debate on migrant integration through a multidimensional and inter-organizational lens. These activities highlighted the importance of evidence-based policies on integration of migrants and provided a space to discuss the toolkit's future prospects concerning its role in informing policies and interventions on integrating migrants into their communities.

4.2. Case studies and country reports

Case study: capturing integration outcomes of CBI beneficiaries in Brazil

INTEGRAÇÃO DE MIGRANTES VENEZUELANOS E HAITIANOS BENEFICIÁRIOS DO PROGRAMA DE ASSISTÊNCIA FINANCEIRA DA OIM NO BRASIL

PRINCIPAIS ACHADOS DA APLICAÇÃO DO ÍNDICE DE INTEGRAÇÃO DE MIGRANTES DA OIM/IPL

RESUMO DO ESTUDO

2022

OIM
ONU MIGRAÇÃO

In Brazil, the study was implemented in five phases: pre-pilot, pilot, actual data collection, report writing and validation workshop. The target group was composed of adults (age 18+), who are refugees, and migrants from the Bolivarian Republic of Venezuela and Haiti who were beneficiaries of CBI. Special attention was also given to migrant women.

For the pre-pilot phase, the respondents' list was composed of migrants from IOM staff and a local NGO that IOM works with (Coração da África). There was a balanced gender distribution between male and female respondents. The main problems identified in this stage were related to the economic and political questions, which made some respondents feel uncomfortable. The findings of the pre-pilot resulted in a series of changes made to the questionnaire, while the introduction and consent section was also enhanced. The entire section regarding the respondents' political integration and knowledge was replaced with questions about their sense of understanding of the political landscape and how often they discussed political concerns with others.

The second stage was the pilot phase, in which 103 nationals from the Bolivarian Republic of Venezuela and Haiti were interviewed. While there was overall satisfaction with the survey, some respondents still expressed discomfort with specific questions about politics, monthly income and unexpected expenses, and chose not to answer these questions. Another challenge faced during the pilot survey was the language barrier with Haitians, who spoke mostly Haitian Creole and could not understand French to the extent that the project team had anticipated.

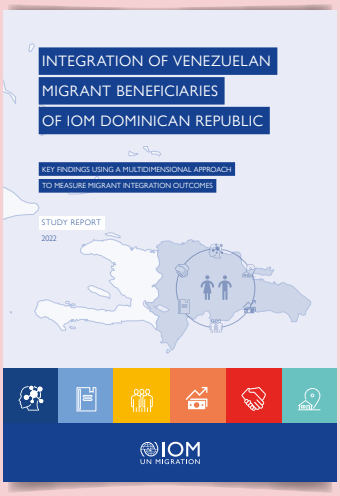
The questionnaire was translated into Portuguese for Haitians who had lived longer in Brazil. Haitian respondents could thus choose between French or Portuguese. The refusal rate for Haitians then decreased. The questions about politics, income and expenses were kept but the wording was adjusted to stress that the respondents' answers would not affect their eligibility to IOM's assistance. Nevertheless, these questions still had a higher rate of "refuse to answer" replies.

After this two-step validation process, a one-off population-based survey was conducted with 798 respondents from the Bolivarian Republic of Venezuela and Haiti for the actual data collection, reaching a total of 901 refugees and migrants surveyed (103 for the pilot phase + 798 for the final data collection). For both stages, IOM Brazil hired the survey firm Kampo Brazil and delivered two training sessions for their enumerators, one before the pilot testing and another before the actual data collection. Also, Kampo offered a financial incentive of a R$ 20 credit top-up to respondents' prepaid phones, a well-received initiative among respondents. The actual data collection took four weeks to be successfully completed.

After the completion of the data collection and the data analysis led by IPL, the project team prepared the study brief. It included relevant results and elaborated some key considerations as the integration scores that emerged from the application of the Multidimensional Integration Measurement Toolkit can help to understand where integration gaps and barriers exist. Identifying these gaps can better map and target vulnerable populations, opening the door to more tailored programming, policies and interventions. Finally, the project team in Brazil organized a validation and capacity-building workshop to launch the study brief and presented the main findings of the research. In addition to the dissemination of the results, the webinar carried out a panel discussion on "Best practices and opportunities in integration of migrants: intersectoral experiences" with representatives from the federal government, local government, civil society, and academia from Brazil.

Source: Brazil Study Brief.

Case study: capturing the integration outcomes of IOM project beneficiaries in the Dominican Republic



Applying the measurement toolkit in Dominican Republic followed a very similar process as in Brazil. The study consisted of a one-off, population-based survey that was also implemented in five phases: pre-pilot, pilot, actual data collection, report writing and validation workshop. The target group was composed of adults (age 18+), refugees and migrants from the Bolivarian Republic of Venezuela, who were beneficiaries of the IOM CBI and who were assisted in the framework of projects such as Inclusive Cities, Communities of Solidarity and the regional programme Bolivarian Republic of Venezuela's Refugee and Migrant Response Plan, during the years 2020–2021. Here again, special attention was brought to gender disaggregation.

The pre-pilot phase was conducted in house, with eighteen respondents, 90 per cent of whom were beneficiaries of IOM programming. There was almost a balanced gender distribution between male (8) and female (10) respondents. The main problems identified in this stage were related to the political questions, which made some respondents uncomfortable. The findings of the pre-pilot resulted in a series of changes made to the questionnaire, while the introduction and consent section was also enhanced. The entire section regarding the respondents' political integration and knowledge was replaced with questions about their sense of understanding of the political landscape and how often they discussed political concerns with others.

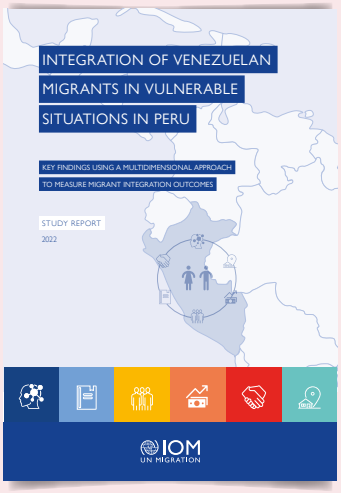
For the second and the third stages of the study, IOM Dominican Republic hired a survey firm named Sigma Dos. During the pilot phase, 64 nationals from the Bolivarian Republic of Venezuela were interviewed. While there was overall satisfaction with the survey, some respondents still expressed discomfort with specific questions about politics, choosing not to answer these questions. After the completion of the pilot phase, a few adjustments were made to the questionnaire, to ensure that respondents were not feeling obligated to answer certain questions.

The actual data collection was conducted through phone surveys with 816 Venezuelans. It took approximately two months to complete. For both stages, IOM Dominican Republic delivered two training sessions for the firm's enumerators, one before the pilot testing and another before the actual data collection. After this last stage of data collection, data analysis and interpretation were conducted by IPL and the report writing by IOM Dominican Republic and IOM Headquarters.

The last phase of the application of the toolkit also culminated with the organization of a validation workshop to disseminate the key findings and launch the study report.

Source: Dominican Republic Study Report.

Case study: Peru's impact evaluation of CBI using the measurement tool



In Peru, as mentioned, the toolkit was used to evaluate the effects of CBI on Venezuelan migrant integration outcomes. Pre-pilot and pilot phases were also conducted before the actual data collection. The target group was made of adults (age 18+), refugees and migrants from the Bolivarian Republic of Venezuela, who were beneficiaries of CBI Multipurpose. This causal impact evaluation followed an entirely different methodological approach.

The research, led by IPL and IOM Peru, followed five stages. The first one consisted of a vulnerability assessment survey to help determine if the individual qualified to receive the assistance. The second step was the first interview (ITV1), to measure the pre-treatment outcomes for each beneficiary. The third step consisted in the transfer of the cash-based assistance (the treatment) to the beneficiaries, which occurred between one and five weeks after ITV1. The fourth step was a second interview (ITV2), a post-assistance monitoring interview, about four weeks after the beneficiaries received the CBI. At last, the fifth step consisted of a third interview (ITV3) which was conducted twelve weeks after the beneficiaries received the assistance and sought to measure post-treatment outcomes. Relevant outcomes for the causal impact evaluation were collected in ITV1 and ITV3. ITV2 included only auxiliary variables.

The data collection was conducted between 17 August 2021 and 2 September 2022, and the dataset used for the analysis consisted of 42 cohorts of beneficiaries. Two types of statistical analyses were conducted:

1. *Before-after analysis: comparing beneficiaries' answers before and after treatment, i.e. before and after receiving the cash assistance.*
2. *Between analysis: comparing beneficiaries already treated with beneficiaries not yet treated in the same week.*

After the completion of the data collection, the project team prepared a first study report with a snapshot of the migrant integration outcomes before the participants received any kind of treatment (ITV1). A second study report, focusing on the results of the two waves of interviews (ITV1 and ITV3), presented the main findings regarding the causal impact evaluation. Both reports presented relevant results and elaborated several key considerations after the application of the Multidimensional Integration Measurement Toolkit in Peru. Again, the results help to understand where integration gaps and barriers exist to better map and target vulnerable populations, while opening the door to more tailored programming, policies, and interventions on integration.

Source: Peru Study Report.

5. CONCLUSION AND FINAL CONSIDERATIONS

Alternative communication and artistic workshops for migrant and hosting community population in the city of Guayaquil, Ecuador. © IOM 2021/Migranautas.

5.1. Towards a multidimensional measurement approach of integration outcomes

With the ever-increasing amount of regional and global migration, as well as growing sociopolitical, economic and environmental instabilities in the world, achieving peaceful relations between groups has become one of the most pressing issues of our time. In this sense, it is essential to understand the populations we serve – in terms of their needs and what they can contribute to the community (assets, resources, knowledge). Working to understand the broader experience of migrants when settling in their new communities has never been more important.

The measurement toolkit has been designed to provide IOM practitioners and external partners with a more in-depth understanding of migrant integration outcomes and offer another method to inform decision-making. The measurement toolkit helps draw on lessons learned from experience to provide feedback into the planning, design and implementation of future interventions and improve service delivery.

It is essential to emphasize the relevance of the multidimensionality of integration as it should be viewed from a holistic point of view. This encompasses access to resources and opportunities as well as psychosocial belonging within their own networks and host communities. It is anticipated that adopting this structured approach to measuring integration outcomes using all various dimensions of integration – as captured in this toolkit – will contribute to a greater understanding of the complex interlinkages between all of the dimensions. It is important to stress that the way the integration dimensions are presented should not suggest a hierarchy: no dimension is positioned as more important than any other.

Therefore, this guide proposes a survey-based index that captures key aspects of integration with only a small number of questions so that it can be used at low cost and facilitate comparability. A few concluding considerations to bear in mind when applying the toolkit are the following:

- **Applying the toolkit in programme settings beyond integration.** Just because the tool measures integration outcomes does not mean it is only applicable to projects or interventions that deal with integration. The toolkit can applied beyond, especially in contexts where there is already vast existing operational data in the target population of interest.

- **Cross-comparison of generated data.** The collected data can be used to compare results across contexts and over time. While there may be some differences in the observed characteristics among migrants, this would allow project teams and organizations to generate larger bases of evidence on integration.

- **Moving from one-off surveys to impact evaluations.** Aside from one-off population surveys (that only provides simplistic descriptive analysis or overview of the population at hand), it is important to look at the effects of programmes or interventions on the integration of beneficiaries.

- **Linking the toolkit with institutional monitoring and evaluation guidelines.** It is important that the toolkit is read alongside IOM monitoring and evaluation guidelines and other institutional assessment checklists and templates. The guide only provides the basic minimum requirements for evaluation, and it is important that processes and operations align with the standards set by the organization.

- **Incorporation of other field-tested instruments.** As the measurement tool does not claim to be a one-size-fits all tool, it is important to look into other complementary resources such as the Global Survey Bank on Integration and Social Cohesion, which provides insightful perspective on migrant integration outcomes. Depending on the context, this would allow practitioners to generate more useful and contextually relevant findings.

- **Capacity-building initiatives.** It is also just as important to develop the capacities of practitioners to conduct measurement exercises not just in the practical side of using and adapting the toolkit, but more so on their understanding of the integration process and the linkages of integration dimensions.

- **Linking outcomes to inform policy and practice**. The application of the measurement toolkit, as well as other survey tools, allow practitioners to generate scientific evidence to better design interventions in ways that are supportive of migrants' integration.

- **Gender-sensitive and rights-based approach.** Lessons learned from the application of the measurement toolkit reinforces that gender and rights-based approaches need to be included in research, humanitarian needs assessments and design of sustainable integration interventions.

6. ANNEXES

MEASUREMENT TOOLKIT

The research "Multidimensional measure of immigrant integration" and accompanying Appendix

Check-list to Identify Projects for Evaluation

Scoring System

Survey Modes

Survey Instrument Brazil (English, Spanish, Portuguese and French)

Study Guide

Enumerators Guide

Form for Pilot Findings and Feedback

Structure and Application of the Index

Coded surveys for IPL-12 (R, Stata) and IPL-24 (R, Stata).

INSTITUTIONAL ANNEXES

IOM Monitoring and Evaluation Guidelines

IOM Procurement Manual: Procurement of Goods, Works and Services (internal link only)

IOM Workplan template (internal link only)

IOM Risk Management Plan (internal link only)

IOM Data Protection Manual

IOM Data Protection Principles

6.1. Checklist to plan the application of IPL integration index

Some basic considerations need to be thought through first before the index is applied and used to measure integration outcomes of migrants. This includes identifying the purpose of the application of the measurement tool, resources and data requirements, and timing before the study kicks off operationally. It would be best to consult key stakeholders, both internal and external, to understand the data or information needs in the local context and how they can be used to generate evidence and inform programmatic and policy approaches.

The first step in the planning phase is to define the purpose of the application of the index. The **purpose** describes the overall reason why the toolkit should be used, its objectives and what results are expected. It is essential to identify the key issues, problems and research questions that can and cannot be addressed during the application of the index. After defining the purpose, it is important to examine the operational, cost, and logistical requirements of the study to ensure that the team has the necessary human, financial and physical **resources** including before, during and after the application of the toolkit. Third, the team needs to consider the **timing**, making sure that the study is in line with a cohesive workplan and that deliverables could be met in a timely manner. For the last part of the planning, it is important to take into consideration the **data** requirements, such as the existing data, data protection principles, and data collection approaches in order to conduct the study.

Below, some considerations can be found to provide some guidance on how to plan the application of the measurement toolkit.

Key considerations:

1. Purpose

a. What is the purpose of the application of the toolkit?
b. What are the research questions that need to be addressed in the study? What do you want to learn about the population/programme?
c. What are the expected results or outcomes? And for what purpose will you use the toolkit?
d. Do the objectives of the evaluation match the needs of key stakeholders?

To do: Develop a concept note that lists down the key questions, objectives, scope and expected results and outcomes of the data collection.

2. Resources

a. Human:
   i. Who are the key stakeholders that need to be consulted? Will this be carried out in collaboration with people external to the organization?
   ii. Who constitutes the project management or coordination structure? What are their roles and responsibilities?
   iii. Who will conduct the surveys?
   iv. Who will interpret the data and develop the report?
   v. Who will be the target population?

To do: Create a project management structure and an internal terms of reference, detailing their roles and responsibilities throughout the evaluation.

b. Financial
   i. Is there enough budget for such a study?
   ii. How much budget should be allocated to an evaluation in a project?
   iii. What are the key budgetary expenses that need to be factored in (staff, enumerators, translation, logistics, report writing, survey and statistical software, survey incentives, etc.)?

To do: Create a budget plan that shows a detailed budgetary allocation for staffing, office and operational costs. It is important to allocate sufficient resources for data collection and include budgetary estimates for this type of evaluation in the intervention development phase. It is also important to consider the possibility of additional costs due to unexpected circumstances such as prolonged data collection.

c. Physical
   i. What materials, services or equipment are needed for the conduct of the survey?
   ii. What are some of the logistical or procurement-related items or services that need to be considered at the outset?

To do: Establish a procurement plan based on identified needs which include the possible services and materials that need to be procured ahead of time. It is also important to identify the procurement method and contracting requirements of the services and goods in advance.

RESOURCES

For more information, refer to Guidelines to the Differences between Individual and Service Provider Contracts (IN/73), IOM Procurement Manual: Procurement of Goods, Works and Services (IN/168 rev. 2) and Selection and Engagement of Consultants (IN/84 Rev. 1). (internal links only).

3. Timing

a. When is the survey going to be conducted? If it is linked to an intervention, will it be conducted before, during, after the intervention or cross-sectionally?
b. What is the timeframe to do planning, evaluation design and data collection and analysis? Is it realistic enough?
c. Is there a workplan in place to monitor the implementation of activities, clearly distribute tasks among the intervention team?
d. How long is the survey (according to the survey mode chosen)?

To do: Develop a workplan that is comprehensive and realistic enough, with the right people assigned to each task, and takes into account workflows and potential scheduling problems right from the start of the project. This will ensure that the outputs are delivered within the time frame and budget available.

RESOURCES

For more information, refer to: IOM Workplan template and IOM Risk Management Plan (internal links only).

4. Data

a. How will data be collected? Are the surveys done in person, via phone, online or on paper?
b. What sampling technique will be used?
c. What data already exist? Is there an existing database that has contact information of the target population?

To do: Prepare a database that contain existing information of the target population. Depending on the database and existing information, the appropriate sampling technique and survey mode can be identified.
For more information, please refer to: IOM Data Protection Manual and IOM Data Protection Principles

6.2. Modes of data collection – an overview

The table below presents an overview of the different modes of survey data collection. The main four modes of data collection are presented for general understanding and for future survey projects.

	ENUMERATOR-ADMINISTERED		SELF-ADMINISTERED	
	1 Face-to-face	2 Telephone interview	3 Paper-based	4 Internet-based
Mode description	Enumerators visit respondents in person, at their homes, an IOM programme location or a public place.	Enumerators call respondents on their phones.	Printed questionnaires are sent to respondents' homes (alternative: respondents fill out the survey at a dedicated location, on paper or on tablets).	Respondents are invited to fill out a questionnaire online on their mobile phone and/or laptop.
Feasibility	✗ Difficult due to COVID-19 (lockdown, uncertainties around travel, safety of enumerators and respondents).		✗ Ill-suited for populations with unstable housing situations (unless administered in another location e.g. administered to beneficiaries of IOM projects at IOM office(s). In this case, using a tablet rather than pen and paper would be more efficient).	
Coverage (i.e. population included)	√ Best coverage (inclusion of illiterate respondents, non-users of Internet/phone).	✗ Excludes non-users of phones. ✗ Excludes respondents who do not have phone/ Internet network. ✗ If conducted over WhatsApp, costs for the respondents (Internet credit usage).	✗ Excludes functionally illiterate respondents (e.g. never learnt to read, cognitive impairment, bad eyesight).	✗ Excludes non-users of phones/laptops. ✗ Excludes respondents who do not have Internet network. ✗ If no free access to WiFi, costs for the respondents (Internet credit usage).
Response rates	√ Higher response rates (enumerators motivate respondents).		✗ Lower response rates, but can be improved by using **incentives.**	✗ In many countries, mobile users may receive a vast amount of advertisement via SMS. In order to catch their attention and highlight that the survey is not advertisement it might be worth calling respondents once, ask them to save the enumerators' number and then send the survey link via SMS or WhatsApp.

	ENUMERATOR-ADMINISTERED		SELF-ADMINISTERED	
	1 Face-to-face	2 Telephone interview	3 Paper-based	4 Internet-based
Data quality	√ Lower item non-response. X Enumerator effect (potential bias for sensitive questions and social desirability bias). This is crucial, given the nature of some of the questions, and the fact that enumerators would most probably come from the host country.		X Higher item non-response. √ Privacy: no interviewer effects/ biases – i.e. more standardized (higher measurement quality).	
Questionnaire length	√ Longer questionnaires are possible (due to the presence of the enumerator).		X Questionnaires should not be too long.	
	√ Longer interviews (up to 45–60 min).	√ Medium/long interviews (15–30 min).	√ Short/medium questionnaires (10–20 min).	X Only relatively short questionnaires (max. 15 min), and even shorter if answered on mobile phone.
Duration of fieldwork	X Medium/long due to enumerators' travel. If paper-based, this also requires data entry – not needed if answers are entered on tablets. (In general, we recommend the use of tablets over paper whenever possible.)	√ Fast – time to reach respondent and conduct the interview.	X Long. It requires sending the questionnaires, waiting to receive them and then entering the data. If respondents enter answers on a tablet, no data entry is required so it is shorter, but respondents have to travel to the survey place.	√ Very fast – time to reach respondent and to fill the questionnaire.
Cost	X Higher cost (enumerators' salaries or IOM staff time).		√ Low cost (no enumerators' salaries or staff time/no travel costs).	
	X Higher cost (travel cost).	Interviews can be conducted through mobile phone calls or via WhatsApp calls using the survey firm or IOM office's WiFi, whichever is cheaper and more reliable.	Cost of printing, sending and providing stamp for return envelope.	√ Low cost.
Logistics	X Contract a survey institute to carry out the fieldwork (one institute for the three countries, or three different institutes, one for each country). X Hire bilingual interviewers for Dominican Republic (Spanish, French, Creole) and Brazil (Portuguese and Spanish). X Organize training sessions for the enumerators.		Organize printing (outsourced), mail cost, data entry (requires development of data entry guidelines).	√ Survey programming can be done in-house (Qualtrics/LimeSurvey/ SurveyMonkey, etc.). X Questionnaire layout should be adapted to **mobile mode** (Qualtrics already offers a mobile-friendly option).
Sampling Frame	Sampling frame with addresses of target population required (or interviews are conducted in public spaces where target population meets).	Sampling frame with telephone numbers of target population required.	Sampling frame with addresses of target population required (or interviews are conducted in public spaces where target population meets).	Sampling frame with telephone numbers or email addresses of target population required.

	ENUMERATOR-ADMINISTERED		SELF-ADMINISTERED	
	1 Face-to-face	2 Telephone interview	3 Paper-based	4 Internet-based
Checklist	• What are the COVID-19 restrictions in the target country? Would they allow this survey mode? Are they likely to change? Would this mode put enumerators/ respondents at risk? • How scattered are respondents? This has implications on the travel costs and duration of fieldwork. • How safe is travelling to places where respondents are located? • Do you have respondents' addresses? If not, how could you get these or reach them? • Can the interview be conducted in a safe space (confidentiality)?	• Do you have a list of phone numbers? • If not, how would you get this information? • What share of the target population uses a phone? • What share of respondents uses WhatsApp (or Viber, Telegram)?	• What share of respondents is illiterate? • Can incentives be considered? If yes, what amount and what type of incentives would be more appropriate? Would incentives costs be within your budget? • Do respondents have a (fixed) address? Do you have their address? • Is the post office reliable?	• Do you have a list of phone numbers/email addresses? • If not, how would you get this information? • What is the share of respondents that uses Internet? • What is the share of respondents that has a phone? • Does each household member have his/her own phone or is it a shared phone? • What is the share of respondents that is illiterate? • Can incentives be considered? If yes, what amount and what type of incentives would be more appropriate? Would incentives costs be within your budget?

7. REFERENCES*

Evidence in Government and Politics (EGAP)

n.d. Methods Guides. Available at https://egap.org/methods-guides.

Harder, N.

2018 Replication Data for: Multidimensional Measure of Immigrant Integration (V1 ed.). doi: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/MF2Q7U.

Harder, N., L. Figueroa, R.M. Gillum, D. Hangartner, D.D. Laitin and J. Hainmueller.

2018 Multidimensional measure of immigrant integration. *Proceedings of the National Academy of Sciences*, 115(45): 11483–11488. doi: www.pnas.org/doi/full/10.1073/pnas.1808793115 Supplementary information available at www.pnas.org/action/downloadSupplement?doi=10.1073%2Fpnas.1808793115&file=pnas.1808793115.sapp.pdf.

International Organization for Migration (IOM)

2020 2019 Return and Reintegration Key Highlights; Available at: https://publications.iom.int/books/2019-return-and-reintegration-key-highlights.

2021 Survey Bank on Migrant Integration and Social Cohesion. Available at https://gmdac.shinyapps.io/integration/.

2022 Recent Migration Trends in the Americas. IOM, Buenos Aires and San José. Available at https://rosanjose.iom.int/sites/g/files/tmzbdl1446/files/documents/oim_recent-migration-trends-in-the-americas-eng-07062022.pdf.

Organisation for Economic Co-operation and Development (OECD)

n.d. What are equivalence scales? Available at www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf.

Plataforma de Coordinación de Interagencial para Refugiados y Migrantes de Venezuela (R4V)

2022 R4V América Latina y el Caribe, Refugiados y Migrantes Venezolanos en la Región - Sept 2022. Available at www.r4v.info/es/document/r4v-america-latina-y-el-caribe-refugiados-y-migrantes-venezolanos-en-la-region-sept-2022.

United Kingdom Home Office

2019 Home Office Indicators of Integration Framework. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1074688/home-office-indicators-of-integration-framework-2019-horr109.pdf.

* All hyperlinks were active at the time of writing this report.

IOM
UN MIGRATION

International Organization for Migration
17 route des Morillons, P.O. Box 17, 1211 Geneva 19, Switzerland
Tel.: +41 22 717 9111 Fax: +41 22 798 6150
Email: hq@iom.int Website: www.iom.int