CGAP

ITU

WORKING PAPER

Mystery Shopping for Digital Financial Services: A Toolkit

Michelle Kaffenberger and Danielle Sobol

CGAP
The United Nation's International Telecommunication Union's Focus Group on Digital Financial Inclusion

TABLE OF CONTENTS

ABOUT THIS TOOLKIT

This toolkit was developed with funding from the Bill & Melinda Gates Foundation, in partnership with CGAP and the International Telecommunication Union's (ITU) Focus Group on Digital Financial Services (DFS). It is designed to help DFS regulators, providers, research firms, and other stakeholders use mystery shopping to monitor DFS markets. A mystery shopping exercise conducted in Zambia, in partnership with the Zambia Information & Communications Technology Authority (ZICTA)—the national telecommunications regulator—is cited throughout this toolkit for illustration.

This toolkit lays out a step-by-step guide to conducting a DFS mystery shopping study. The main audiences for this paper are regulators and supervisors and the field research firms they hire to carry out the mystery shopping. Both parties should read this entire toolkit, but each should pay special attention to specific sections that address their particular roles:

- ***Regulators and supervisors should concentrate on sections A, B, C, and G of Part II***. These sections guide regulators and supervisors through the process of identifying research objectives, selecting a research firm, creating a research plan, determining study content, and participating in an iterative analysis process. Regulators with greater capacity or experience are also encouraged to participate in the steps directed at the research firm, or to carry them out in-house.
- ***Research firms should concentrate on sections C, D, E, F, and G of Part II***. These sections will help research firms to prepare logistics, develop research tools, carry out training and fieldwork, and analyze results.

I. INTRODUCTION: WHY MYSTERY SHOPPING?

Digital financial services (DFS) are playing an increasing role in providing low-income individuals with access to financial services. The rapid expansion of DFS makes regulations relevant to consumer protection increasingly important in promoting responsible service provision to maximize benefits and minimize risks to consumers.

Mystery shopping is a valuable tool regulators can use to help them understand consumers' experiences using DFS and interacting with DFS agents. Mystery shopping can help identify problems in the market, inform regulations, and monitor the market for regulatory compliance. It can be used at various points in market development, such as helping understand the customer experience following the rollout of new services or regulations, to investigate specific market issues, or on an ongoing basis to check continual compliance with existing regulations.

Mystery shopping has been used in many environments to identify critical insights for regulators, supervisors, and other stakeholders. CGAP, for example, developed a technical guide on mystery shopping for financial services that provides practical guidance on how to assess behaviors in financial markets using this methodology, based on CGAP's insights and learnings from applications in seven lower-income markets (Mazer, Gine, and Martinez 2015). The case studies identified market conduct issues ranging from limited presentation of mandated financial product information, steering of customers toward less suitable products, fraud, overcharging of customers, and provision of different information to more experienced customers. For example, a multi-country audit study in Ghana, Mexico, and Peru, conducted by CGAP and the World Bank, found that bank staff provided information on costs associated with credit or savings products only after being asked, and that customers were rarely offered the cheapest product, likely because of the financial institutions' sales incentives. This indicated that transparency and disclosure of key product terms and conditions were often insufficient for consumers to make informed decisions (Gine and Mazer 2016).

In another setting, the World Food Program (WFP) in Kenya and CGAP partnered to conduct mystery shopping for humanitarian cash-based assistance programs. WFP Kenya had shifted some of its food distribution programs to cash, delivering the cash digitally into bank accounts, mobile wallets, and mobile vouchers. Such a transition can present challenges for beneficiaries, who often represent the most vulnerable and least literate populations in a country, and requires extra accountability for agents who help beneficiaries redeem or cash out their transfer. The mystery shopping exercises, which included training actual beneficiaries as shoppers, revealed that agents often charged beneficiaries unauthorized fees for withdrawing their funds, merchants increased their prices for goods purchased with a bank card, both agents and merchants frequently entered beneficiaries' PIN for them, and agents and merchants were not always able to adequately address beneficiaries' problems or grievances (WFP forthcoming and Mazer and Baur 2014).

A mystery shopping study conducted in partnership with CGAP, ITU, and the Zambian telecommunications regulator, ZICTA, serves as an example throughout this toolkit (Kaffenberger 2017). The study helped ZICTA measure compliance with existing DFS regulations, such as identification requirements for registering for an account, as well as identify areas that need further action, such as potential disclosure requirements for service fees.

This toolkit provides guidance on the process that a regulator and a research firm can use when conducting a mystery shopping study of the DFS experience. It provides examples of research objectives and scenarios that can be tested, model structures for setting up fieldwork teams and training, and suggestions for conducting analysis. It can serve as a practical guide for regulators, or other stakeholders, who are ready to implement a study to gain deeper understanding of the DFS market.

II. CONDUCTING A MYSTERY SHOPPING STUDY

A. IDENTIFY OBJECTIVES FOR THE STUDY

To begin the mystery shopping study, regulators should first identify the key objectives they wish to accomplish. They may consider existing regulations for which they want to measure compliance; potential problem areas for which they are considering new regulations; or general inquires for monitoring market development. "Tool #1. Sample Study Objectives" suggests example objectives that the regulator might aim to accomplish.

TOOL #1. Example Research Objectives

Objective #1: Test agent compliance with existing regulations, such as (depending on current regulatory requirements):

- Does the agent check ID upon registration?
- Does the agent check ID upon deposit, withdrawal, or any other times when it should be required?
- Does the agent display fee charts, customer care contact information, fraud warnings, and anything else required?
- Does the agent perform an over-the-counter (OTC) transaction through the agent's own account, when asked (if this is not allowed)?
- Does the agent allow transfers over the transfer limit, and if so how does the agent conduct these?
- Is the agent able to explain certain key features of DFS when asked? (This could indicate a need for greater agent training.)
- Is the agent able to help the customer identify proper recourse channels when a problem arises? (This could indicate a need for more agent training.)

Objective #2: Identify issues in the market that may require new regulations, such as the topics listed in Objective #1, if not already required by regulation.

Objective #3: Understand developments in the market, consumer experience, and agent behavior more generally, such as the following:

- Agents' abilities to describe and explain new products and services being offered, such as digital credit or insurance.
- Agents' explanations of personal identification numbers (PINs), their use, and the need to keep them secret/secure.
- The general state of agents' shops. Are they secure? Do they allow private transactions?
- The extent to which agents struggle to maintain sufficient liquidity.

B. SELECT A FIELD RESEARCH PARTNER

After identifying the objectives of the study, regulators who do not have internal research capacity or experience need to select and hire a research firm to conduct the study. Selecting an experienced and competent partner is crucial. "Tool #2. Selecting a Field Research Partner" gives an overview of the main features to look for in a research partner.

Even after hiring a research firm, the regulator should stay closely involved in several aspects of the study. Areas where regulators should be most involved in developing ideas and decision-making include deciding which scenarios to test, and with whom and where to test the them (both described in Section C); reviewing and approving the research tools to ensure that all of the desired information is captured (Section D); and, through an iterative analysis process, ensuring findings fully meet their needs and adequately address the research objectives (Section G).

TOOL #2. Selecting a Field Research Partner

The ideal field research partner will have the following qualities:

- ***The firm will be local***. It is critical that the research firm have contextual knowledge and understanding of the local DFS environment. Make sure that the firm hires locals as mystery shoppers, to ensure they are treated the same as typical customers, and that the managers who will be organizing and supervising the fieldwork have intimate knowledge of the market.
- ***The firm should have mystery shopping experience***. Ideally, regulators should work with a firm that has conducted a similar mystery shopping study focused on customer service and experience objectives. Having a mystery shopper go through full financial transactions with an agent is a complex form of mystery shopping, and thus, if possible, it is better to identify a firm with experience conducting complex mystery shopping studies. If this is not possible, a firm with some mystery shopping experience should be identified.
- ***The firm must have the capacity to organize for and carry out the complicated logistics of the study***. The firm will need to map out the logistics of mystery shopping scenarios, which will likely include (at minimum) fund deposits, withdrawals, and transfers among a large number of mystery shoppers in different locations. The firm will need to have one or two managers dedicated to running these logistics.

C. WORK WITH THE RESEARCH FIRM TO CREATE A RESEARCH PLAN AND DETERMINE STUDY CONTENT

C.1. Determine a timeline for the study

The regulator and research firm should develop and agree on a timeline that is feasible for the firm and that meets the needs of the regulator.

C.2. Decide which scenarios will be tested, based on the study objectives

Scenarios are the stories that the mystery shoppers will act out to test certain regulations or to learn more about how agents respond to certain events. Example scenarios

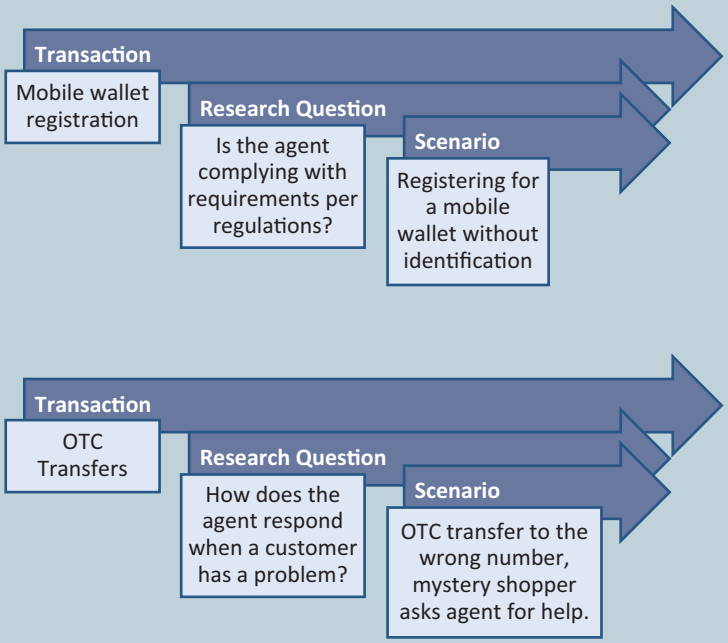
TOOL #3. Drafting Mystery Shopping Scenarios

The following three steps (followed by illustrative examples) will guide the process of drafting mystery shopping scenarios.

1. **List the DFS transactions available to consumers:**
   Mobile wallet registration, bill pay, OTC transfers, mobile money transfers, mobile wallet deposits and withdrawals

2. **Develop research questions based on the study objectives:**
   - Is the agent complying with requirements per regulations?
   - Are agents providing correct and thorough information?
   - How does the agent respond when a customer has a problem?

3. **For each type of transaction, what scenarios can be used to investigate the research questions?**



include registering for an account without proper identification, conducting an OTC transaction, and asking the agent for more information about a particular service. Tool #3 describes the process for developing the scenarios to be tested. Regulators should consider the first list of scenarios a draft—the scenarios may need to be revised based on experiences during the mystery shopper training and piloting stages of the study.

C.3. Determine *which providers and products* will be tested, and *where*

Which providers and products are tested and where the mystery shopping is conducted depend on many factors, including the research objectives and available budget. The Zambia study tested agents serving the top six providers in the market, including two mobile wallet providers, three OTC transfer services, and one bank product that

incorporates a mobile component with a widespread network of banking agents. Testing different geographic areas provides additional insight into how customer/agent interactions and experiences differ based on location. The Zambian study covered five locations: three urban centers, one peri-urban center, and one rural area.

Finally, the regulator, in collaboration with the research firm, will need to determine a sample size. This will depend on the research objectives and available budget. Mystery shopping studies often include between 50 and 100 mystery shopping visits. Larger sample sizes allow testing of more scenarios and transaction types. The Zambia study used a sample size of 300 visits, which allowed testing of 11 scenarios across seven transaction types.

C.4. Establish a methodology for validating the research plan

The research firm should validate the scenarios that will be tested, to ensure that all scenarios are relevant to DFS users and that no common problems experienced by actual users have been missed. In-depth interviews (IDIs) and/or focus group discussions (FGDs) with actual DFS users in the market can be used to accomplish these aims, and should be conducted before the scenarios are finalized.

The information derived from IDIs and/or FGDs will help the regulator and research firm identify issues that should be included in a scenario, such as a type of agent fraud that mystery shoppers should watch for, or a common type of user problem that should be included in a scenario. The results can also help identify user experiences or concerns with particular providers that could be relevant for mystery shopping. The results can also help inform the mystery shopping questionnaire, ensuring that it captures all the information relevant to the user experience.

The decision of whether to use IDIs, FGDs, or both will be based on the context and project resources. For example, in some countries, FGDs are difficult to organize and conduct, while in other countries, FGDs might be the most efficient and effective way to gather information. Usually one approach or the other will be selected, but both may be used if the research firm or regulator believes different insights will be gained from each based on their local context.

The regulator and the research firm should review the results of the IDIs or FGDs together, and decide on changes to the scenarios and questionnaires that may be needed.

D. DEVELOP RESEARCH TOOLS

Three research tools will need to be developed for the study:

1. The IDI or FGD guide, which will be used to validate the scenarios and which involves qualitative and open-ended elements (Section D.1.).

2. The mystery shopping scripts, which will inform the mystery shoppers on how to act out each scenario (Section D.2.).

3. A close-ended questionnaire the mystery shoppers will complete after each mystery shopping visit (Section D.3.).

It is important to keep in mind that the mystery shopping scripts and questionnaire, like the scenario list, may be revised at multiple stages: first as a result of validation

exercises with DFS users; again after mystery shopper training, when the mystery shoppers can give feedback on improvements; and finally as a result of piloting the materials.

D.1. In-depth interview or focus group discussion instrument

The IDI or FGD instrument informs the regulator and research firm about interactions between actual DFS users and DFS agents, validating the scenarios and ensuring the mystery shopping visits accurately reflect real-life interactions.

This instrument should include a short list of open-ended questions. The IDIs or FGDs should target respondents who are similar to those who will be the focus of the mystery shopping. For example, if the mystery shopping study will have a particular focus on the experience of female DFS users, then women should be targeted for the IDIs and FGDs. The IDIs and FGDs should include respondents who have used DFS services recently, for example, in the past two weeks or month, because they are likely to be able to easily recall experiences. Questions should cover relevant topics. The following are examples of questions respondents may be asked:

- What is your typical experience conducting a transaction with an agent?
- What information or documentation (such as forms of identification) do agents usually ask you to present?
- What do agents say or do when you interact with them?
- What kind of transaction confirmations have you received?
- What problems have you experienced while using DFS, particularly when you were interacting with a DFS agent?
- How willing and able have agents been to help you resolve problems?
- What are the customer service issues you've encountered? What have been your experiences when calling a customer care line or visiting a customer care center?

The Zambian study included 12 IDIs.

D.2. Scripts

Scripts explain to mystery shoppers exactly how to play out each scenario when interacting with an agent. After a few days of training and practice, the mystery shoppers internalize the scripts and learn to be flexible and responsive to the actual situation. Scripts will include shopper profiles and a scenario guide.

Shopper profiles. To measure differential treatment among different population segments, shoppers should be assigned a consumer profile that is based on the types of consumers in the local context. The following are just a few examples of profiles:

- A low-income woman with no DFS experience.
- An elderly man with limited DFS experience.
- A young woman with significant DFS experience.

TOOL #4. Creating Shopper Profiles

1. Brainstorm and compile a list. In the local context, what populations might receive especially poor or differentiated treatment when interacting with DFS agents? What population segments are regulators most interested in or concerned about? Here are some examples to think about:

Shopper profiles should be relatively simple so that useful comparisons and conclusions can be drawn. It is ideal to choose two or three criteria on which to focus. For example, the profiles could include low-income men and low-income women, to allow conclusions about how gender affects the way agents treat low-income customers. There could also be profiles for inexperienced, low-income men and inexperienced low-income women. Combining more than three criteria narrows the focus, and makes it difficult to test each profile enough times to draw conclusions. (For example, a profile for a "middle-aged, low-income man of a certain ethnicity with no DFS experience" would be too detailed.)

2. Choose which traits will be actual traits and which will be assumed traits. For some traits, such as gender, the mystery shopper must have the actual trait. Other traits, however, can be assumed or acted out, such as income level or experience with DFS. For example, a profile of low-income men could include male mystery shoppers who are "acting" as though they are low-income, such as by adjusting the way they dress and the way they talk. Ideally shoppers would have as many of the *actual traits* of the profiles as possible.

3. Match profiles to scenarios. Profiles need not be the same for each scenario. For example, you may want to see how agents interact with different experience levels when registering a new customer, but you may be more interested in how agents interact with different poverty profiles when sending cash OTC.

Whenever possible, shoppers should have the actual traits of the profile they are projecting. Some traits, such as age, gender, and ethnicity, are obvious. For other traits, available shoppers who do not fit a profile may have to assume the traits of the profiles needed for the study. For example, someone with no experience using DFS would ideally carry out a scenario with the profile of an inexperienced shopper. However, when this is not possible, a shopper will "play the role" of a consumer who has no experience using DFS. Similarly, a shopper can adjust the way he dresses so as to appear low-income, even if he is not. Scripts not only describe these profiles, they also explain to shoppers, for example, how to behave like someone who is inexperienced. See "Tool #4. Creating Shopper Profiles."

Scenario guide. The script will lay out how the transaction should occur, and how the shopper can guide the conversation to address all the elements necessary to later complete the questionnaire. The guide should not include scripted lines meant to be memorized verbatim, because each shopper and interaction will be a little bit different. Rather, the script acts as a guide to conversation flow and specific points to address. See "Tool #5. Components of a Mystery Shopping Script."

TOOL #5. Components of a Mystery Shopping Script

Mystery shoppers should be given a packet of information outlining all the elements they need to carry out the mystery shopping visits successfully. During training, they will work through each element until they know it well and can act it out from memory. The following are elements that should be included:

- ***Customer profile,*** and tips on how to best "play the role" when traits are assumed. This may include dress, language, a background story, or specific words to use or not to use. The script may also include the amount that each profile should use when transacting if this is used to indicate income level.
- ***Conversation flow***, including how to approach the agent, what to do or ask, and when. This can be numbered if the conversation is meant to follow a specified order, and may include instructions on when to wait for an agent to bring something up and when to actively approach the topic.
- ***Probing questions*** that match the questions in the questionnaire. Probing questions allow shoppers to gather additional information they will need to complete the questionnaire.
- ***Directions on how to react to likely developments.*** For example, shoppers may be instructed to wait while an agent calls a customer care number, or to get very angry if an agent does not allow them to perform an illegal transaction.

D.3. The Questionnaire

Shoppers should fill out the questionnaire immediately following a mystery shopping visit. They should report exactly what happened during the interaction. The questionnaire should include primarily close-ended questions that can be analyzed quantitatively. Data should be collected electronically such as through a mobile device or tablet where possible, to limit data entry errors and allow for efficient quality-control measures.

The questionnaire for the Zambia study began with basic questions about the shopper, scenario, and location. It then moved on to specific questions about each scenario.

- *Section I: Agent Information* gathered basic descriptive information about each agent.
- *Section II: General Indications* covered all the information that was not directly related to a specific scenario, such as wait time, information posted at the agent shop, and cleanliness and privacy of the agent location.
- *Section III* was different for each scenario and concentrated on exactly what that scenario is testing, such as registering without identification documents or making an OTC transaction.

"Tool #6. Helpful Hints for Developing the Questionnaire" provides some ideas and suggestions for creating a questionnaire.

TOOL #6. Helpful Hints for Developing the Questionnaire

1. **Test know-your-customer compliance.** For scenarios where an agent may ask for identification, the questionnaire should include a question about information and documents the agent requested, so that the number of agents that asked for proper identification across all relevant scenarios can be assessed. Including options in the questionnaire that are not necessary for the agent to know, such as marital status or occupation, checks whether agents are gathering unnecessary information or possibly discriminating against certain types of individuals.

2. **Test how well agents are relaying product, consumer safety, and recourse information.** Agent provision of information can be tested through a series of question types. An example that can be applied to many topical areas is illustrated here. The example involves creating a PIN when registering for a mobile wallet:

   - **Addressed or Not addressed:** *Was the topic covered?*
     Question to include: Did the shopper and the agent discuss PIN creation?
   - **Spontaneous or prompted:** *How was the topic brought up?*
     Question to include: Did the agent spontaneously offer information on why the shopper needs to create a PIN, or did the shopper have to probe for information?
   - **Mentioned or explained:** *How thorough was the explanation?*
     Question to include: Did the agent provide an in-depth explanation or a shallow explanation of why a PIN is important?

3. **Test general consumer experience and protection indicators.** Each questionnaire should include general questions on consumer experience and protection issues, such as privacy, agent trustworthiness, agent knowledge, and comprehensiveness, and other points of interest. For example, a three-point scale could be used to measure whether the agent was (1) very knowledgeable, (2) somewhat knowledgeable, or (3) not at all knowledgeable on a particular topic.

4. Design the questionnaire for easy reporting and analysis. Instead of multiple-choice questions, allow the shopper to answer yes or no to as many questions as possible. Here is an example:

INSTEAD OF THIS:	ASK THIS:
What forms of identification did the agent ask for? ☐ National ID Card ☐ Passport	Did the agent ask to see your National ID Card? ☐ Yes ☐ No Did the agent ask to see your Passport? ☐ Yes ☐ No

The "Ask This" option forces shoppers to think about each form of identification individually, and ensures all relevant information is captured.

5. **Include a notes and comments section at the end of each questionnaire.** There will inevitably be things that happen during a mystery shopping visit that are not captured in the questionnaire. An open-ended notes and comments section allows shoppers to add information that otherwise would not be recorded, and gives them the opportunity to record their personal feedback on the interaction.

TOOL #7. Example Research Team Structure

Research Lead. Ideally, the research lead is the same person who created the research tools. This person will lead the mystery shopper training and oversee the project.

Project Managers. Project managers will work under the research lead. They will be responsible for project logistics, including managing the money needed for each scenario (e.g., ensuring shoppers who are depositing money have the cash they need to do so, and ensuring shoppers who are withdrawing money have the e-money in their wallets to do so). They will also be responsible for quality control and monitoring.

Team Leaders. Team leaders will be mystery shoppers in the field who directly manage other shoppers. They will be the first line of quality control and monitoring, and will also be the pilot team. Team leaders may be selected from among all shoppers after a few days of training so that the shoppers who quickly and proficiently master the material and study objectives are selected to be team leaders. There should be about one leader for every 4–5 mystery shoppers.

Mystery Shoppers. The number of mystery shoppers needed will depend on the number of scenarios, budget, and timeline. Enough shoppers should be included so that each may specialize in about three scenarios. Whenever possible, shoppers should have the actual traits of the profile they are portraying.

E. FIELDWORK

Before beginning fieldwork, the regulator or research firm will need to assemble a research team. A sample team structure is outlined in "Tool #7. Sample Research Team Structure."

Fieldwork occurs in four phases:

1. Conducting IDIs or FGDs
2. Training
3. Piloting and "Extended Mystery Shopping"
4. Implementing and doing quality control

E.1. In-depth interviews or focus group discussions

The research firm should conduct IDIs or FGDs and discuss results with the regulator. Some of the researchers who will later take on the role of mystery shoppers can conduct the IDIs or FGDs. Sampling can be done randomly among DFS users who reflect the shopper profiles. Once results are discussed with the regulator, changes can be made to the scenarios and to the mystery shopping instrument before mystery shopper training begins. See Box 1 for an example of this phase.

E.2. Training

Training could last a week, depending on the size of the study and the number of mystery shoppers and scenarios. The training should focus on having the shoppers practice going through each script until it is second nature, and giving shoppers the opportunity to fill out the questionnaire enough times that they internalize the information they must remember when acting out the scenario with a real agent. The following is a sample schedule that can be used as a guide when organizing a mystery shopper training. The research lead will conduct the training and ideally will be same person who created the research instruments, because this person will be best positioned to understand how to effectively use the instruments. All project managers should also participate in the training.

Day 1: The first day of training should be spent teaching shoppers the methodology and reviewing the different profiles and scenarios. The mystery shoppers should go over the scripts and discuss the scenarios to gain familiarity and to help determine the feasibility of each scenario and which providers would make sense to target in each scenario. (See Box 2.)

Days 1, 2, and 3: After an overview of the project, the methodology, and the scenarios, the research lead should go through each scenario and questionnaire individually. Here is an example of a process to follow:

- Begin with the research lead acting out a scenario in front of the shoppers as they follow along with their scripts.

BOX 1. Incorporating in-depth Interview Findings into the Zambia Instrument

In the Zambia study, IDIs revealed that almost all agents from all providers have a record book customers must sign to confirm they have made a transaction. Customers may be required to record information about the transaction (type, amount) as well as personal information (National Registration Card number, phone number). The widespread nature of the practice, and the potential that such record keeping could lead to fraud or misuse of information, led us to add a question to the questionnaire on whether the mystery shopper was required to fill out this type of book, so that we would have a quantitative measure of how common this practice is.

We also learned from the IDIs that agents routinely ask customers who receive OTC transfers how much money they are expecting to receive, and that agents do not always inform OTC recipients if the sender paid the transfer fee or if the recipient is responsible for it—both of which could facilitate agent fraud. Accordingly, we added questions to the questionnaire asking if OTC agents did either of these.

BOX 2. Incorporating Feedback from Training into the Zambia Instrument

We originally included a scenario in which the mystery shopper would pay a bill OTC through an agent. While reviewing the mystery shopping scenarios with the shoppers, we learned that it is not possible to pay a bill through a DFS agent account—an agent may help a customer pay a bill on the customers' device, but not OTC through the agent's device. We thus revised the scenario to instead focus on recourse for incorrect bill payments, and not on testing experiences with OTC bill payments.

- Then, talk through the questionnaire as a group, allowing each shopper to read a question and answer it, according to how the scenario was played out. Invite other shoppers to agree or disagree with the answer given.
- Next, have two shoppers act out the scenario in front of the others, discuss the scenario as a group, and go through the questionnaire again as a group.
- Finally, group the shoppers into pairs and have each shopper take one turn as the customer and one turn as the agent, each time talking through the questionnaire with his or her partner.

This exercise should be repeated for each scenario. The project managers should be included in all steps, because they will play an important role in quality control and, therefore, must understand the instrument well.

Days 3 and 4: As shoppers become familiar with the scripts, they can begin practicing in small groups. For example, the research lead may do the following:

- Form groups of four or five shoppers each, and have the groups rotate so that the research lead and the project managers spend time observing each group.
- Have each group start with the first scenario and practice it until every shopper has had a chance to play the customer role. Then move to the next scenario and repeat the exercise.

As the research lead rotates through the groups, he or she will correct shoppers while simultaneously training the project managers to perform quality control checks. The project managers will learn what to look for when performing these checks. This process also helps the research lead and project managers identify which shoppers are the best fits for each scenario.

Days 4, 5, and 6: After the small group practice, shoppers should be given their scenario assignments so that they can concentrate on developing expertise in a few scenarios. Ideally, shoppers should specialize in about three scenarios, but the actual number will depend on the size of the study and number of shoppers and scenarios. Although shoppers will be specializing in specific scenarios, it is still beneficial for shoppers to be familiar with all the scenarios, as this allows for greater flexibility in the field if a shopper needs to replace another. Shoppers should spend the remainder of the training in groups according to their scenario assignments, practicing their assignments as many times as possible.

E.3. Piloting and "Extended Mystery Shopping"

At some point during the training, shoppers who have picked up the material quickly are likely to stand out. In the Zambia project, for example, this group emerged by the end of the second day. Shoppers who quickly mastered the material can go on to pilot the scenarios with real agents, while the other shoppers continue to train and practice.

Piloting allows the researchers to test the research tools, including the scenarios, scripts, and questionnaire, by doing actual mystery shopping visits in the field and identifying which parts of the scenarios and questionnaires work well and which need to be revised. Each scenario should be piloted two to three times, preferably one in each geographic area (e.g., if the study covers urban and rural areas, a scenario would be piloted once in an urban setting and once in a rural setting). Piloting could take one to two days, and depends on the size of the study and locations of the pilots. Pilots should be followed by an in-depth debriefing session that includes the research lead, project managers, and shoppers conducting the pilot. Piloting will likely lead to some changes to the tools.

After piloting, and subsequent changes to the instrument, the research firm can conduct a final stage of testing: "extended mystery shopping." Although this is not a required step, extended mystery shopping can produce extremely informative results and add anecdotal insights to the quantitative mystery shopping findings. For extended mystery shopping, the scenarios are carried out to their full conclusion, beyond what is required during the standard mystery shopping visits. For example, in a standard scenario that tests recourse, a mystery shopper may ask an agent for help and then record in the questionnaire whether the agent called customer care, directed the shopper to a customer care center, or refused to help. In the case of extended mystery shopping, if an agent directs a shopper to a customer care center, the shopper goes to the customer care center and continues to play out the scenario until the problem is resolved or there is no action left to take. This exercise provides insight about recourse experiences in a way that would be logistically difficult for a large sample of mystery shoppers.

Finally, the group of mystery shoppers who conducted the pilots and extended mystery shopping visits may then serve as team leaders during the implementation phase. Their experience acting out the scenarios in the field will help to inform their role as team lead.

E.4. Logistics

Shoppers should be given a list of the specific scenarios they will carry out, the profile they will play, the type of agents they will visit, and where the visit will take place. Coming up with these specifics will require logistical planning and understanding of which shoppers speak which languages, which can travel, and so forth.

Assignments may also involve details that are specific to the study. For example, the Zambia study investigated compliance with a regulatory requirement that a customer may not have two mobile money wallets with the same provider. To test this, the research firm had to gather information on which mystery shoppers

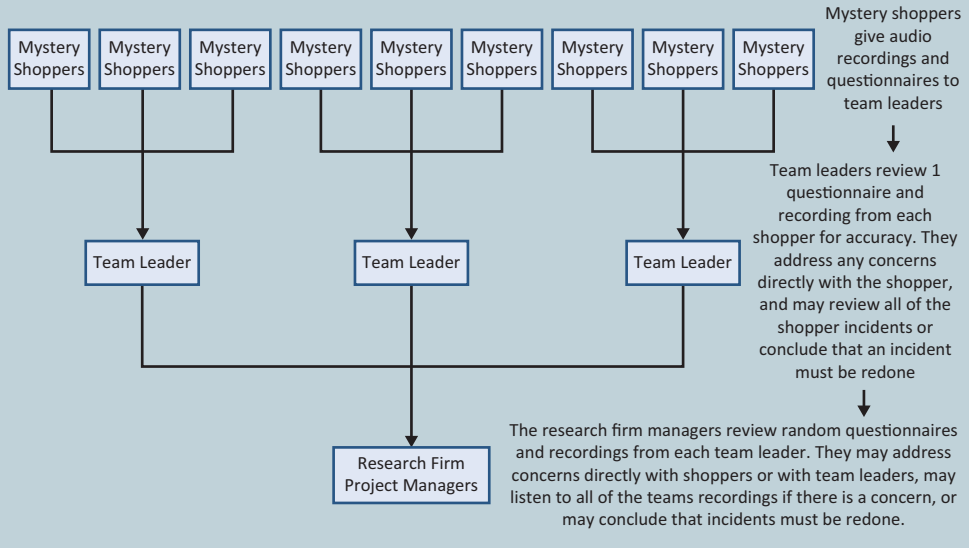
already had a (personal) mobile money wallet and which provider those wallets were with, to distribute scenarios accordingly. The firm also had to coordinate how many SIM cards each shopper needed. Giving shoppers new SIM cards for the study (rather than having them use personal SIM cards) helps to simplify logistics and prevent shoppers from having to disclose personal account information as part of the study.

E.5. Implementation and quality control

Selecting a location can involve preselecting the neighborhoods or towns that the mystery shopper would visit, and then allowing the shopper to randomly select agents within that area. A more structured approach would involve mapping out the neighborhood or town to be visited, and telling shoppers which specific agents to approach. In both cases, shoppers should have specific targets for provider type and location, and should fill out their questionnaire immediately after the mystery shopping visit.

The research firm should implement a robust monitoring and quality control plan to ensure that the scenarios are being carried out correctly, and that questionnaires are filled out properly. In the Zambia study, the firm used a multistage approach. The shoppers recorded the audio of their mystery shopping visit using a dictaphone they kept in their pocket. Each evening, the team leader listened to at least one scenario from each shopper while going through the accompanying questionnaire. If something was not captured correctly in the questionnaire, the team leader discussed the issue with the shopper and went through the shopper's other scenarios. Next, project managers went through a sample of recordings and questionnaires from each team leader's pool, to ensure that the team leaders were performing appropriate quality checks. "Tool #8. Sample Monitoring Plan" diagrams the management structure and quality control arrangement used in the Zambia study.

TOOL #8. Sample Monitoring Plan



In addition to monitoring shopper performance, the research firm should produce regular field reports (e.g., once every other day) for the regulator. These reports should give an overview of progress to date, discuss any challenges that the shoppers have faced in finding agents or carrying out scenarios, and allow shoppers to share insights or interesting stories. In the Zambia study, team leaders compiled this information and sent it to the project management team at the research firm, who synthesized the information into a field report to share with the research lead.

F. POTENTIAL CHALLENGES

A downed network or insufficient agent float may present challenges to successfully carrying out a mystery shopping visit (unless the research questions include how often networks are down or how often agents do not have float). In the case of network problems, the shopper can still fill out the portion of the questionnaire that remains relevant, such as agent demographic information, privacy, and cleanliness. However, this need not count toward the total number of mystery shopping visits the firm will conduct. In the case of insufficient float, agent reaction to this problem may be instructive and interesting to the study. One option is to allow shoppers to follow agent instructions when this happens, such as waiting for float to arrive or visiting another agent, and then have shoppers report the outcome. In some cases, the outcome might be that the scenario cannot be completed. The regulator and the research firm should decide the terms for such situations. One example is to have the research firm count a certain number of these toward the total required mystery shopping visits (in the Zambia study it was agreed that such problematic visits could represent no more than 10 percent of the total).

One unexpected challenge in the Zambia study was that fraud, or suspected fraud, was common in a number of the research locations, making agents naturally suspicious. As shoppers enacted the scenarios, agents were particularly suspicious of mystery shoppers conducting large OTC transfers because they feared these could be fraudulent transfers. One agent called her local mobile network operator representative to her booth because she suspected that a shopper was attempting to commit a scam. In this case, the shopper produced an introduction letter from the research firm that confirmed his transaction was for research purposes. He was forced to "break character" and explain that he was a mystery shopper. It is good practice to give shoppers this type of identification in the event that something goes wrong.

G. ANALYSIS

When analyzing mystery shopping data, it is useful to look at the data through two modalities: as stand-alone scenarios and as a full sample. Within each of modalities, data can be categorized by research question and shopper profile.

Individual Scenarios. Each type of scenario should be analyzed individually, using the research questions to organize results. For example, the research question "Is the agent following rules?" may manifest itself as "Did the agent proceed with registering a new account if the customer did not have necessary identification?" In another scenario it might mean "Did the agent allow a customer to transfer an amount over

the transfer limit?" Simple frequencies—such as the number of shoppers allowed to proceed without identification and the number who were not—are instructive for analyzing agent behaviors. Analysis can focus on how shopper profiles are associated with the outcome—for example, were men allowed to proceed more often than women?

Anecdotes from individual shopper's experiences, provided in the comments section at the end of the questionnaires, can add further depth and context to the findings, thus providing a richer set of project conclusions.

Full Sample. Certain questions are asked across multiple scenarios, such as those regarding cleanliness of agent shops, or agents' provision of information. Analyzing these questions across the full sample can more fully address these research questions and give more weight to analysis of each shopper profile (such as comparing men and women) because the sample size for analysis is larger. For example, the research question "Are agents providing correct and thorough information?" may be answered by looking at all transactions where there is a fee, and analyzing the frequency with which the fee was disclosed to the shopper before the transaction. Analysis can then examine how customer profiles affect these.

Finally, the analysis should be an iterative process. The regulator should be given the chance to look over the first round of analysis and to go back to the research firm with follow-up questions and requests for a deeper dive into the data.

III. CONCLUSIONS AND GOOD PRACTICES

In Zambia, the mystery shopping study showed that most shoppers could register for a DFS account without showing the necessary identification. All who tried were able to bypass OTC transaction limits, and most agents who quoted transaction fees to shoppers told shoppers the wrong amount. According to one shopper, "I was initially informed the transfers attract zero charges, but I later incurred charges after the transfer was complete and the agents were unable to explain this."

Insights like these are critical for regulators and supervisors charged with overseeing DFS markets. Compliance with identification requirements is necessary for the "know-your-customer" (KYC) process. And fee disclosures are a foundation of consumer protection.

Mystery shopping is also an important tool for identifying emerging risks with newer products. In Zambia, among mystery shoppers who sent a bill payment to the wrong account (intentionally, as part of the scenario), none was able to retrieve the funds, and advice from agents was inconsistent. As bill pay services expand, such difficulties will become more common.

Regulators and supervisors should use mystery shopping, along with other tools, to understand market developments and customer experiences and risks in their markets. They can also use mystery shopping to hold specific providers to account, if their agents are consistently noncompliant. The step-by-step guidance in this toolkit enables authorities to carry out such studies and identify needed actions in their markets.

GOOD PRACTICES

- Regulators should develop clear research objectives that will guide each stage of the study.
- Regulators should work with a qualified, creative, and flexible local research firm that has mystery shopping experience. The project managers at the firm should participate in shopper training and understand the instrument thoroughly.
- A longer training period will help ensure that acting out the scenarios and filling out the questionnaires are second nature to the shoppers.
- Using qualitative methodologies, such as IDIs and "extended mystery shopping," helps validate scenarios and augment results, thus providing additional insights.
- Training shoppers in all scenarios, while also allowing specialization in specific scenarios and profiles, provides flexibility in the field so that shoppers can accommodate challenges (such as filling in for a sick shopper) during field work.
- A robust management structure and monitoring and quality control system is critical to ensure high-quality data, and should include regular field reports for the regulator.

IV. REFERENCES

Gine, Xavier, and Rafael Keenan. 2016. "Financial (Dis-)information: Evidence from a Multi-Country Audit Study." Policy Research Working Paper No. WPS 7750. Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/869451468937960883/Financial-Dis-information-evidence-from-a-multi-country-audit-study

Kaffenberger, Michelle. 2017. "How 'Mystery Shopping' Can Help ICT Regulators: 5 Lessons from Zambia." Blog post, 2 March. https://itu4u.wordpress.com/2017/03/02/how-mystery-shopping-can-help-ict-regulators-5-lessons-from-zambia/

Mazer, Rafael, and Silvia Baur. 2014. "Merchant Incentives in the Shift to Cashless Food Aid." Blog post, 24 November. http://www.cgap.org/blog/merchant-incentives-shift-cashless-food-aid

Mazer, Rafe, Xavier Gine, and Cristina Martinez. 2015. "Mystery Shopping for Financial Services: What Do Providers Tell, and Not Tell, Customers about Financial Products?" Technical Guide. Washington, D.C.: CGAP. http://www.cgap.org/sites/default/files/Technical-Guide-Mystery-Shopping-for-Financial-Services-Oct-2015.pdf

World Food Program. Forthcoming. "Behavioral Program Monitoring in Cash Transfer Programmes." Kenya: World Food Program.