A Methodology for Laboratory Experiments in Developing Countries: Examples from the Busara Center*

Johannes Haushofer[†], Marie Collins[‡], Giovanna de Giusti[§], Joseph Muiruri Njoroge[¶], Amos Odero[‖], Cynthia Onyago[**], James Vancel[††], Chaning Jang[‡‡], Maneesh Varghese Kuruvilla[§§], Conor Hughes[¶¶]

February 24, 2014

Abstract

While Randomized Controlled Trials and lab-in-the-field experiments have become more common in development economics in recent years, the rigor and control of laboratory-based experiments has so far been difficult to access in developing countries. Here we describe the use of various cognitive tests in developing countries, illustrated using data from the Busara Center for Behavioral Economics, a state-of-the-art laboratory for behavioral and experimental economics in Nairobi, Kenya. In addition to we provide information on the development of the lab itself, including data on physical and technical setup and infrastructure, protocols for study administration, respondent and data flow, and subject recruitment, payment, and subject pool composition, and outline how researchers can use the lab.

JEL Codes: C93, D03, D87, O12

Keywords: development, laboratory experiments, experimental economics, behavioral economics, psychology

*Our gratitude for making Busara possible goes to the staff og Innovations for Poverty Action, in particular KAren Levy, Colin Christensen, Daniele Ressler, Delia Walsh, NIall Keleher, Justin Oliver, Nicole Grant, Joann Phelan, and Annie Duflo; the staff at the Unviersity of Zurich, in particular Stephan Schmid, Michael Hohl, and Adrian Etter. We thank Ernst Fehr, David Laibson, Iris Bohnet, Mark Edington, Amal Devani, and Jasper Grosskurth for valuable feedback and advice; Anne Kioi, Japeth Juma, Febrone Atieno, Anne Wynne, Carol Aluchio, Nonnie Kamau, Lillian Achibeala, and Peter Gachanja for research support; and Anne Healy for comments on the manuscript.

[†]Abdul Latif Jameel Poverty Action Lab, MIT joha@mit.edu **(Corresponding Author)**

[‡]Google mcollins@gmail.com

[§]Maseno University giovanna.degiusti@gmail.com

[¶]Innovations for Poverty Action, Kenya jmuiruri@poverty-action.org

[‖]Financial Sector Deepening aodero@poverty-action.org

[**]Innovations for Poverty Action, Kenya conyago@poverty-action.org

[††]Innovations for Poverty Action, Kenya jvancel@poverty-action.org

[‡‡]University of Hawai'i at Manoa chaningjang@gmail.com

[§§]University of St Andrews mk542@st-andrews.ac.uk

[¶¶]Abdul Latif Jameel Poverty Action Lab, MIT conorbhughes@gmail.com

1 Introduction

Since their beginnings in the 1960s, laboratory experiments have become an increasingly popular tool for studying economic behavior in controlled settings. In his early double-auction experiments, Vernon Smith (e.g.Smith 1962; Smith 1982) showed that core features of microeconomic theory can be tested and proven accurate in laboratory experiments, and thus provided an early proof-of-concept for the usefulness of laboratory experiments in economics. Since these early successes, experimental economics has become a large and productive sub-discipline whose reach is still expanding; for instance, recent research has begun to model macroeconomic phenomena in laboratory settings (e.g. Lei and Noussair 2002). The rise of laboratory experiments in economics also prepared the ground for the rise of behavioral economics, notably through the work of Amos Tversky and Daniel Kahneman (1979). The goal of this research program was to inject psychological realism into models of economic behavior, and it relied heavily on laboratory experiments to do so. In addition, and partly owing to inspiration from behavioral economics, recent years have seen a surge of interest among psychologists, cognitive scientists, and neuroscientists in the cognitive and neural processes underlying reward, valuation, and decision-making (e.g.Schultz 2006; Glimcher 2003; Glimcher 2009); these studies, too, are mostly conducted in lab settings. Together, these developments within experimental, behavioral, and neuroeconomics have ushered in a new era of psychologically inspired research on economic preferences and decision-making in the laboratory, and today these fields are among the most prolific and rapidly evolving subdisciplines of economics.

Surprisingly, however, laboratory studies in behavioral science have so far largely been restricted to developed countries, frequently using university students as subjects. In psychology, for instance, a 2008 survey by Arnett (2008) found that 96% of subjects in studies published in top journals were from "WEIRD" (Western, educated, industrialized, rich, and democratic) backgrounds. This fact represents a significant challenge to the external validity of the findings obtained in these contexts. Indeed, those experiments that did study non-Western subjects often found considerable variability across cultures (Henrich et al. 2001; Cardenas and Carpenter 2008). These results, in turn, led to calls for more research on the behavior and preferences of people from non-Western contexts (Henrich, Heine, and Norenzayan 2010).

Despite these developments, to our knowledge the rigor and control of laboratory environments have not been readily accessible in developing countries to date. We report here on the creation and features of the Busara Center for Behavioral Economics in Nairobi, Kenya. The core of Busara is state-of-the-art laboratory with 20 networked PCs, combined with a subject pool of over 5000 residents of the Nairobi slums, and an additional subject pool from the University of Nairobi. Use of touchscreen computers and special paradigms enables even illiterate respondents to participate. Busara can be used by researchers around the world; its goal is to provide economists, psychologists and cognitive scientists, political scientists, anthropologists, and other behavioral scientists with an infrastructure which facilitates the administration of high-quality laboratory experiments in a developing country. The goal of this paper is twofold: first, it aims to provide enough detail to enable others to create similar labs in other countries; in the long run, we hope that Busara will become a node in a network of social science laboratories in developing countries. Second, the paper aims to serve as a reference and manual for researchers interested in running studies at Busara. The rest of this paper is organized as follows: Section 2 reviews the merits of laboratory-based experimental studies in relation to field experiments, lab-in-the-field experiments, and online studies. Section 3 reports the results of various cognitive tests conducted using Busara's touchscreen computer system on both a general population sample and subjects from Unviersity of Nairobi.

2 Advantages of laboratories

Before we describe the example cognitive tests, setup, and features of Busara, we briefly list the advantages a laboratory setting has over field or lab-in-the-field experiments. In field experiments, data is typically collected when a field officer conducts one-on-one interviews with respondents, either with a paper questionnaire or with a laptop computer. Lab-in-the-field experiments (cf.Cardenas and Carpenter 2008) attempt to establish laboratory-like conditions in field settings – e.g., participants might be gathered in the classroom of a local school to play a trust game using paper-and-pencil communication. In our view, dedicated laboratories have several advantages over such settings, which we briefly describe below.

First, lab settings allow researchers to maintain the anonymity of participants and their responses to a much greater degree than field or lab-in-the-field experiments: in the latter form of experiments, participants often know the other participants and their choices in an experiment, either because choices have to be made publicly (especially when participants are illiterate), or because the setting is such that even private decisions become public immediately due to the difficulty of restricting communication between participants during the experiment. In contrast, lab settings allow participants to be seated in private cubicles and give responses anonymously. A related advantage is that lab settings allow for sensitive information to be communicated confidentially, with neither the researcher nor other participants becoming aware of a subject's response to any given item; in field settings this is often not possible, mainly because responses are recorded by field officers.

Second, field or lab-in-the-field experiments are frequently conducted in small communities, where participation is public knowledge and many respondents in a particular session may know each other. While such familiarity may be useful or necessary for some experiments, it is more desirable to induce it through laboratory techniques in a sample of unacquainted respondents (e.g. provision of selected pieces of information about other participants, or a minimal group paradigm) than to use the endogenously formed familiarity patterns already present in a sample of acquaintances and family members. Economics labs make this approach possible, not only because they make it easier to manipulate perceived familiarity among participants through experimental techniques, but also because they often have large subject pools of potential participants, drawn from geographically disparate areas. Selecting a random sample from such pools, especially when it happens privately such as through text messages, enables researchers to ensure that a minimum number of participants know each other before the study.

Third, computerized lab experiments allow researchers to conduct interactive games with much greater ease and efficiency than field or lab-in-the-field settings. Interactive economic exchange games, such as Trust, Dictator, Ultimatum, Public Goods Games, or Double Auctions, are backbone paradigms of experimental and behavioral economics, but they are cumbersome to implement as paper-and-pencil experiments in field or lab-in-the-field settings. However, their administration is easy and fast in lab settings with networked computers.

Fourth, and relatedly, a computer lab allows for richer stimulus sets than are possible in paper-and-pencil experiments. For instance, researchers can easily employ pictorial, audio, and video stimuli, as might be used in priming studies. In paper-and-pencil experiments, this is either not possible or more difficult.

A fifth and related advantage concerns respondent comprehension. This is a key concern for economic experiments, especially in developing countries with low education levels. Computerized administration makes it easier for researchers to facilitate and test understanding with animated instruction videos, practice rounds, and test questions.

Sixth, while paper-and-pencil experiments usually require at least basic literacy, computerized lab experiments in developing countries allow for the participation of both computer-illiterate and entirely illiterate respondents. In particular, we describe below how we use touchscreens to enable participants without familiarity with computers to participate, and how specially developed paradigms can make lab studies accessible to entirely illiterate participants.

Seventh, computerized lab studies obviate the need for data entry. For paper-and-pencil studies, data entry can be a significant cost in terms of time and money, and an important source of errors; in computerized studies this factor is eliminated, increasing data quality and decreasing cost.

Eighth, computerized lab experiments are efficient; data can be collected from many participants at once, with a small number of research assistants (RAs). This is generally not possible in survey experiments, which are often face-to-face.

Finally, Busara has important advantages over internet data collection through tools like Amazon's Mechanical Turk (MTurk). In recent years, such internet studies have rapidly gained popularity, especially in the psychology community. Platforms like MTurk allow researchers to put experimental paradigms online, and thus rapidly gather data from hundreds of subjects whose collection would otherwise take a much longer time in the lab. However, a critical limitation of studies on MTurk is that they require a) access to a computer, b) basic reading and, often, writing skills, c) a bank account or other formal financial mechanism for receiving payment. The large majority of the world's population fails to meet any one of these criteria, let alone their combination. Thus, MTurk studies still exclude most of the world's population, especially the poorest. In contrast, to participate in studies at Busara, participants do not need a computer, reading/writing skills, or a bank account;

the only requirements are that participants be 18 years or older and have access to a cell phone and the MPesa mobile money system. Recent data collected in the Nairobi slums by Tavneet Suri shows that, despite their poverty, more than 90% of people there have both a cell phone and MPesa access. Thus, Busara is set up to include even the poorest participants, and thus provides a unique opportunity to study respondents at the very "bottom of the pyramid", who are not accessible through internet tools such as MTurk.

3 Cognitive tests

Here we provide examples of simple yet informative laboratory-based cognition tasks that can be administered to subjects from a wide variety of backgrounds using computer interfaces. We illustrate using data from a series of sessions conducted at Busara.

3.1 Screen tasks

One of the most basic tests implemented was a series of "screen" tasks. The central feature of Busara's computer setup that allows us to work with our special subject population is the fact that all lab computers are equipped with touchscreens. During initial testing we put great emphasis on ensuring that participants could actually use these touchscreens. In particular, in July 2012 we tested 71 randomly selected subject pool members on a series of simple tasks to be performed on the touchscreen. Specifically, participants were asked to perform the following tasks (cf. Table 1); each correct response was rewarded with KES 5 (~USD 0.07).

1. Tasks 1-4: Two black boxes, each ~"3x3" in size, were displayed on a gray screen in a particular configuration. The boxes were either arranged next to each other, or stacked on top of each other. Participants were asked to choose either the left or the right box in the former case, or the top or the bottom box in the latter case, by pressing it with their finger. The instruction was given both on the screen in English ("Please put your finger on the {left; right; top; bottom} black area.") and verbally, by a research assistant reading out the same sentence translated to Swahili. Table 1 reports the mean proportion of correct responses and its standard deviation for this task. Participants performed it almost perfectly, with only a few errors.

2. Tasks 5-9: A horizontal row of 5 boxes (each "3x3") was displayed on the screen, each with a different color (red, green, black, blue, yellow). In each of 5 consecutive trials, participants were asked to press a certain color. The order of the boxes was randomized with respect to the order of the questions. Again the instruction was given both in English writing on the screen, and orally in Swahili. It can be seen from Table 1 that participants had little trouble with this task.

3. Tasks 10-14: A horizontal row of 5 boxes (each "3x3") was displayed on the screen, each showing a different object (cow, hammer, tree, chair, car). In each of 5 consecutive trials, participants were asked to press a certain object. The order of the boxes was randomized with respect to the order of the questions. Again the instruction was given both in English writing on the screen, and orally in Swahili. Again Table 1 shows that participants were able to easily complete this task.

The nearly flawless responses to these simple tasks suggest that the touchscreen interface is easily understood, even by subjects who may be computer-illiterate or possess poor reading and writing skills. Having demonstrated this proof of concept, we now turn to a series of more challenging cognitive tasks, and present results from both a general population sample and a sample of University of Nairobi students.

3.2 Arithmetic tasks

Table 2 presents the results of a basic arithmetic task. Respondents were asked to give a series of responses on the touch screen to assess their math skills, which are essentil in many paradigms in exerimental and behavioral economics. Each correct response was rewarded with KES 5 (~USD 0.07). Each problem was displayed in writing on the screen; e.g. a one-digit addition problem would be displayed as "6 + 2 = ?". Below the question, a 4x4 number pad was displayed on the screen, containing the numbers 0-9, a red field labeled "Clear' and a green field labeled "OK". In one-digit addition, two one-digit numbers had to be added, and the result was also a one-digit number; in two-digit addition, two two-digit numbers had to be added, and the result was also a two-digit number. The one- and two-digit subtraction, multiplication and division tasks were constructed analogously.

Table 2 shows the results of this series of questions for both the general population and University of Nairobi students. The student group not only consistently answered these questions correctly more often than the general population, but also answered more quickly. The differences between the two groups, in terms of the proportion of correct answers and response time for each question, are all statistically significant at the 1% level. Interestingly, despite these differences, respondents from both groups seemed to have the hardest time answering the same questions. For example, on the one-digit group of questions, both groups performed most poorly on the one-digit division question (79% correct in the general population group, 98% correct in the student group). Furthermore, both groups missed the two-digit multiplication question more than any other question, which also took them the longest time to answer (46% correct with 61.35 seconds average reaction time in the general population group, 85% correct with 46.58 seconds average reaction time in the student group).

3.3 Raven's progressive matrices

Raven's progressive matrices are a well-known and commonly used cognitive test. It was developed to be an easily administered method of directly measuring a subject's general intelligence (Raven and Raven 2003). It is nonverbal, and requires no reading, writing, or mathematical training, making it feasible to use in a wide variety of contexts, and the results comparable between subjects of different backgrounds.

In each question, subjects are presented with a partial group of diagrams, where each individual diagram shares a feature of some kind with the others. The subjects are then presented with multiple choices of other diagrams, and asked to choose which one best fits the pattern. The patterns increase in difficulty as the test progresses, but they are all based around asking the subject to discern a pattern between objects, and then extrapolate from them to choose the next object that best fits.

Table 3 shows the results of a three matrix question set that was administered to subjects. Similar to the arithmetic results, students did significantly better than the general population, on average getting 77% of the questions right, as opposed to 49% in the general population group. The differences for both proportion of correctly answered questions and reaction time is again statistically significant at the 1% level for every question. The student group also tends to answer more quickly than the student group, with the exception of the third question, where the students took significantly longer than the general population group to respond. Since the general population group overall wasn't very succesful on the third question (31% got it correct), the lower response time to this question might be a consequence of respondents giving up and simply making a guess about the correct answer.

3.4 Cognitive reflection test

Many behavioral economists and psychologists have recently contrasted what seem to be two distinct modes of cognition: "fast" and "slow" processes, labeled "System 1" and "System 2" processes by Stanovich and West (2000). System 1 processes are made quickly and heuristically, requiring little mental effort. For example, recognizing that a voice belongs to a close family member or friend is largely a System 1 process, in that it occurs instantly and intuitively. System 2 processes, on the other hand, are involved when a problem cannot be resolved utilizing System 1 processes alone, and involve a higher degree of mental effort and concentration. While the problem of recognizing the voice of a close friend can be solved quickly and easily, the same cannot be said when addressing a problem like $\int_0^{\pi/2} \int_0^a \frac{ar}{\sqrt{a^2-r^2}} \delta r \delta \theta$, which would require considerably more thought and System 2 processes to solve correctly.

The three item Cognitive Reflection Test ("CRT") contains simple questions designed to suggest an immediate and intuitive answer, which is incorrect. The correct answer is not difficult to calculate, but in order to find it, subjects must be able to override their initial System 1 reaction and bring System 2 processes into play. The questions on the original three item CRT are (Frederick 2005):

1. A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

3. In a lake there is a patch of lily pads. Every day the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake?

The content of these questions was modified slightly to make sure the objects and concepts involved would be recognizable to a Kenyan subject, but the structure of the questions remained the same. In each case, there is an incorrect yet immediately appealing answer (10 cents, 100 minutes, and 24 days, respectively) which subjects must be able to put aside in order to find figure out the correct answers (5 cents, 5 minutes, and 47 days). Despite its simplicity, the performance on the CRT has been shown to be a very strong predictor of an individual's decision making processes, especially with respect to risk and time preferences. It also correlates with other commonly used measures of intelligence, such as the SAT, ACT, and the Wonderlic Personnel Test (Frederick 2005).

Table 4 shows the results of the CRT as administered to our general population and student groups. Similar to previous tests, students were significantly more likely to find the correct answer compared to the general population. While the relative difference is large, both groups scored quite low: over the three questions, students on average were correct 26% of the time, and general population respondents only 4% of the time. As opposed to the earlier tests, the student group had a slower response time on each question than the general population sample, lending further support to the idea that the student group was better able to overcome fast, System 1 decision making processes during the testing. However, most of the differences in reaction time are not statistically significant.

4 Physical infrastructure

The Busara Center for Behavioral Economics is located in the Kilimani neighborhood of Nairobi, Kenya, within the premises of Innovations for Poverty Action Kenya (IPA-K). Kilmiani is a relatively safe and upscale neighborhood and is home to many NGOs. Nevertheless, it is also located one mile from Kibera, Nairobi's largest slum. The compound has a 24-hour security service, with two guards at the gate at all times who have access to radios and panic buttons for calling backup in the case of security threats. Busara offers a state-of-the-art infrastructure for experiments in behavioral economics and other experimental social sciences. The facilities available include: a networked computer cluster; a waiting room; four individual survey cubicles for private interviews; laboratory space for saliva and blood sampling; and an office with several desks for staff, visiting researchers, and students.

1. The networked computer cluster consists of 20 workplaces for participants that are equipped with 20 computers, and one additional control computer at the front of the room. All computers are Hewlett Packard TouchSmart 310 All-in-One PCs, equipped with keyboard, mouse, camera, microphone, and headphones (described in more detail below). A central feature of the computer setup is that all computers have touchscreen monitors, and therefore can be used by illiterate subjects or subjects who are unfamiliar with operating a mouse or keyboard. The computers are arranged in four rows of 5; each individual desk is 70 cm wide and 60 cm deep, and is separated from the neighboring desks through dividing panels that provide complete privacy and anonymity while performing tasks. The panels are 60 cm high on three sides of the desk, and extend 30 cm beyond the edge of the desk in the horizontal direction, to ensure that participants cannot observe each other from the side. To turn the laboratory into a classroom, the dividing panels can be fully removed, without tools, by simply sliding them out of guide grooves and stacking them on the side of the room. The cubicles are numbered busara01-busara20, and each cubicle is fitted with a sign indicating its number. When participants arrive for a study, they are randomly assigned a particular cubicle and given a placecard that matches that of the cubicle. The placecards are marked not only with the cubicle names, but also with 4 different colors (one for each row: Red, Green, Blue, Yellow), and 5 different symbols (circle, eye-shape, triangle, square, and star for the first to fifth PC in each row). Together, the colors and symbols of the placecards uniquely identify the cubicles, allowing even illiterate participants to find their assigned seat.

2. A large office provides 6 desks for staff as well as visiting investigators and students; in addition, it houses a conference table which can be used for meetings, and an additional Hewlett Packard TouchSmart All-in-One PC, with the same configuration as the computers in the main lab, that can be used for developing and testing programs and paradigms.

3. A 20-person waiting room accomodates respondents while they wait, and for briefings before and after studies.

4. An additional room provides 4 individual survey cubicles in which participants can be interviewed one-on-one, with complete privacy. The walls of the individual cubicles extend from floor to ceiling, creating ompletely closed and separate rooms. Each room has both natural and artificial light and is air-conditioned.

5. Finally, Busara houses a small laboratory equipped for obtaining saliva and blood samples. A -20 deg C freezer is available on-site for specimen storage; continuous and uninterrupted power supply is ensured through a backup system for power outages up to 8 hours. In addition, a -80 deg C freezer owned by Busara is located at Kenyatta National Hospital, about 1 mile from Busara; this freezer is attached to Nairobi's hospital power grid and is therefore not affected by power outages. A phlebotomist with a degree in laboratory sciences from the University of Nairobi is available to perform blood draws using the Vacutainer system. Supplies for saliva

samples and blood draws are available locally, and a stock is kept in the Center. Analysis capabilities for saliva and blood samples (e.g. cortisol, CRP, cytokines) are available at Lancet Labs, a state-of-the-art commercial laboratory and long-time Busara partner, located half a mile from the Center.

5 Technical infrastructure

5.1 Hardware

5.1.1 Lab computers

The 22 computers at Busara are Hewlett Packard TouchSmart 310-1126 All-in-One high performance PCs. Each computer has a 1TB harddrive, 6 GB RAM, an Intel Core 2 Duo processor at 2.90 GHz, a webcam, a CD/DVD and Blu-Ray player and burner, Bluetooth capability, several USB ports, a media card reader, a wireless network card and ethernet ports, high-quality speakers, a headset with microphone, and a 20-inch widescreen touch-sensitive display. The PCs are all-in-one machines, i.e. all hardware is integrated into the screen, obviating the need for placing a tower case by respondents' feet. In addition, a 1.5 TB network attached storage system (NAS) is available; it is mounted on all computers and used for data storage and common access, as well as backup of all experimental and subject pool data. In particular, the NAS allows all experiment PCs to access the same files, e.g. images to be used in experiments, or batch files to start programs. Thus, if a batch file involved in starting up the lab needs to be changed, it is sufficient that this change be made once.

The entire lab is fully networked using both wired and wireless connections, to create redundancy in case of the failure of one system. Network connectivity is provided by a 48-port switch located in a server room. A total of 36 active wired connections are available, of which 24 are allocated to the lab computers; 4 are available in the individual survey cubicles; and 8 are available in the researchers' office. In addition, wireless network connectivity is provided using DHCP throughout the entire lab, allowing for a large number of simultaneous connections. All PCs are assigned fixed IP addresses in a dedicated IP address range on both the wired and wireless connection. The building is connected to the internet through a high-speed fiberoptic line; the arrival of fiberoptic internet in Kenya in 2009 has dramatically increased attainable connection speeds and bandwidths: download speeds at Busara are typically 2 Mbps, and upload speeds 0.5 Mbps. These speeds allow, for instance, for the unproblematic use of YouTube videos as stimuli in studies. Electricity is supplied by Nairobi City Council and backed up through an uninterrupted power supply system, which can provide power to the lab for 8 hours during power outages.

The TouchSmart PCs are distributed as follows: 20 PCs serve as client computers in the main computer lab (computer names: busara01-busara20); one PC serves as the control computer at the front of the lab (busaracontrol); and another PC is located in the researchers' office and serves as a backup machine and testing computer for developing programs and paradigms (busaraoffice). The cubicles are labeled with printed and laminated A5-sized signs with the number of the computer, a color, and a symbol, which uniquely identify the computer as described above.

5.1.2 Enrollment and Identification Netbooks, Fingerprint Scanner, and Recruitment Modems

In addition to the lab infrastructure described above, we use three Hamster IV fingerprint scanners from 360 Biometrics, in combination with three Asus EEE Netbooks, to record participants' fingerprints when they first sign up for the subject pool, and to identify them when they come to the lab for sessions for which they were invited by SMS. Automatized sending of bulk SMS invites is performed with a laptop equipped with a Huawei USB modem containing a Safaricom SIM card.

5.2 Software

5.2.1 Lab computers

All lab computers at Busara run Windows 7 Home Edition and Norton Antivirus. All notifications, automatic updates, and screensavers are disabled to facilitate optimal performance of the experimental software and avoid potential distraction or interruptions of study sessions with participants. Any necessary operating system updates are performed manually once per week. The NAS is mounted on all computers. The NAS is also set up for VPN access, allowing researchers to access it remotely with a password.

The main software used for interactive economic exchange games is zTree [fischbacher1999ztree]. This software, also developed at the University of Zürich, is specially designed for the creation and administration of computerized economic experiments. The main instance of the program is run by the experimenter on the busaracontrol machine, and client programs ("zLeaves") are opened automatically upon startup of the client PCs. The connection between the main zTree instance on busaracontrol and the client computers is established via the IP address of the server; thus, a subset or all of the client computers could in principle be located in a different lab (e.g. in Europe or the US).

The computers are started and shut down by using LabControl, a software developed at the University of Zürich, which uses Wake-on-LAN to remotely boot machines. Upon startup, all machines execute a batch file that opens a zLeaf and thus the connection to the host computer. The successful startup of both the computers and the zLeaves can be monitored on busaracontrol. Thus, it is possible to make the entire lab ready for running an experiment with one click.

It is possible to run more than one experiment during a study session by running more than one instance of zTree on busaracontrol, and connecting subsets of zLeaves to different zTree instances using zTree's built-in channel feature. The different instances of zTree on busaracontrol are distinguished by using different channels, and the zLeaves and the zTree that belong to one experiment must use the same channel.

In the course of each experiment, the session data is stored in an Excel file, which can easily be used for data analysis. zTree is configured such that output files are writing to a password-protected directory on the NAS, and thus are accessible through VPN.

5.2.2 Subject Database, Enrollment, Identification, Study Sign-up, and Respondent Payment

On recruitment days in the slums and at the University of Nairobi, participants' demographic information is registered using the Blaise computer-assisted interviewing software, which runs on two Asus EEE Netbooks ("Enrollment Netbooks"). Blaise is interfaced with an MS Access Database on these netbooks; at the end of a recruitment day, the databases from the two Enrollment Netbooks are consolidated and stored on a desktop PC in the Busara offices, as well as backed up on the NAS. In addition, each newly registered participant provides a fingerprint during registration; this fingerprint is recorded with a Hamster IV scanner connected to the Enrollment Netbooks using the PersonID software by 360 Biometrics. Each respondent's fingerprint is linked to their demographic data in the MS Access database through a unique identifier; this identifier is assigned automatically to both the fingerprint and the demographic data, thus precluding possible errors from manual entry.

When subjects report for a study session at the lab, their fingerprints are scanned at the gate of the Busara compound using an additional Hamster IV scanner connected to another Asus EEE Netbook ("Identify Netbook"). The unique identifier for the subject is automatically populated into a field in a Blaise program called "Identify", which then retrieves that subject's name and age from the MS Access database. This information is then used to confirm that the subject appears on the list of invitees for the particular session. The Blaise program then records the attendance of the subject pool member at the particular session; this information is later added to the MS Access subject pool database, and can subsequently be used to a) track session attendance by respondent, study, and sign-up mechanism (phone vs. SMS), b) selectively exclude participants from participation in particular studies depending on their participation in earlier studies. E.g., researchers conducting a study on trust might want to exclude participants who have previously played a Trust Game at Busara; the tracking of attendance and participation through fingerprints ensures that the subject pool contains up-to-date information on the experimental history of all respondents in the subject pool.

The list of invitees for a particular session is generated as follows. First, a Stata script draws a random list of respondents from the subject pool database, according to the demographic criteria specified for the particular study (e.g. age, education), and prior study participation (see above);

second, the bulk SMS software FrontlineSMS, running on a laptop equipped with a Huawei USB modem and a Safaricom SIM card, automatically sends personalized SMS messages to all respondents on this invitee list. Respondents reply to these SMS messages with the word "YES" to register for a particular session, or "NO" to decline the invitation. FrontlineSMS automatically decodes these responses and generates a "YES"-list for each individual lab session, which is then used at the gate for respondent identification.

Finally, after a session is complete, subjects are paid through an MPesa transfer to their cell phones. This is achieved as follows: during the experiment, zTree (or another survey software) keeps track of subjects' earnings in the tasks performed by subjects. At the end of this experiment, a 200 KES (~USD 3) show-up fee is added to this amount, as well as a KES 50 (~USD 0.60) bonus if the participant arrived on time. A Stata script then generates a payment list including names, phone numbers, and amounts to be transferred, for all participants in the session. This file is sent to the IPA-K finance team, who initiate the MPesa transfer to respondents' cell phones. The time between completion of an experimental session and the arrival of the payment in participants' MPesa accounts is 2 hours on average.

6 Subject pool recruitment

Busara maintains a subject pool with currently over 5000 members, who are recruited from the Nairobi slums and the University of Nairobi. Permission for recruitment was obtained from the District Commissioners (DCs) of Langata and Embakasi districts of Nairobi; in addition, Busara's activities are covered by Institutional Review Board (IRB) approvals at the Massachusetts Institute of Technology (MIT), Harvard University, Innovations for Poverty Action-Kenya (IPA-K), and the Kenya Medical Research Institute (KEMRI). In addition to subject recruitment, these IRB approvals also cover many standard behavioral games. Researchers running studies at Busara need to obtain IRB approval from their home university.

Before carrying out recruitment for the respondent database in a particular area, Busara representatives met local area leaders and informed them about the lab. The DCs next facilitated meetings with local government and community leaders in settlements found within these districts; in particular, we focused our recruitment efforts on Kibera slum in Langata District, and Viwandani slum in Embakasi District.

The Kibera slum is the largest informal settlement in Nairobi and is situated 5 km from the city center and 2 km from Busara's office on Ngong Road. Its population was estimated at 170,000 in the 2009 National Census. The slum is divided into 9 smaller villages: Kianda, Soweto East, Gatwekera, Kisumu Ndogo, Lindi, Laini Saba, Siranga, Makina, and Mashimoni.

Viwandani is a smaller slum located 7 km from the city center, in Nairobi's industrial area. Established in 1973 on land that was left by the City Council as reserve area on the bank of Ngong River, Viwandani is divided into 5 villages: Paradise, Jamaica, Lunga Lunga, Donholm, and Kingston.

Pairs of recruitment officers moved from village to village in each of the slums over the course of several weeks between April-July 2012. In each area, two community leaders were recruited to mobilize the community. On recruitment days, potential participants first received a short presentation about the lab, its mission, and their opportunities for study participation. Interested participants read and signed a consent form (illiterate participants were read the consent form and gave their fingerprints to indicate consent), and were signed up for the database by providing their names, contact phone numbers, and some demographic information. In particular, for each participant, we recorded name, age, ethnicity, marital status, number of children, and education level. In addition, a fingerprint was scanned for each respondent, both as a unique identifier that obviates the need to require respondents to have national IDs, and to alleviate security concerns when participants are called to the lab to participate in studies.

The eligibility requirements consisted in being above the age of 18, and having access to a cell phone and mobile money system MPesa, which was used for payment. Although the penetration of both cell phones and MPesa registration (which is free) is extremely high in the Nairobi slums, respondents who were not signed up for MPesa were allowed to provide the number of a friend or family member. (This occurred in a minority of cases.) Together, these criteria ensure that participation was not limited by socio-economic status, and in particular ownership of a phone, Safaricom line, a national ID, or an MPesa account. In addition to the slum sample, Busara has recently established contacts with local universities to expand recruitment to university students, with the aim of diversifying the subject pool available provided for research. Formal permissions for recruitment at the University of Nairobi were obtained from the Deputy Vice Chancellor for Student Affairs, and the Deputy Vice Chancellor for Research and Extension Programs, and the Dean of Students provided logistical support.

7 Subject pool characteristics

As is typical for subject pools in experimental and behavioral economics, participants are recruited with a convenience sampling methodology. Busara's field officers set up desks in different areas of the slums and at the University of Nairobi, in places that are visible and easily accessible, while community mobilizers from the particular slum area brought potential respondents to register. The subject pool thus represents a community sample of slum residents. For each registered respondent, the database contains information on gender, age, education, ethnicity, marital status, and number of children. Recruitment of subjects for studies can be stratified according to these criteria. In addition, phone number and fingerprints were recorded for communication and sign-up before a study, identification at the gate, and payment after the study.

Between April-July 2012, we recruited 1973 potential participants in the Nairobi slums with further recruitment carried out during the remainder of the calendar year, and 200 potential participants at the University of Nairobi. This paper focuses on the slum sample. Table 5 illustrates in detail the area of residence of the slum residents registered in Busara's subject pool. Broadly, the large majority of subjects (around two thirds) live in Kibera, and a smaller proportion (one third) were recruited in other slums of Nairobi. Kibera residents are mainly from the Kianda, Makina and Gatwekera villages, while the majority of residents in other slums were recruited in Lunga Lunga, Sinai and Donholm.

Table 6 summarizes the gender composition of the Busara subject pool, and compares it to the composition of Nairobi and Kenyan populations. Females are slightly over-represented in the Busara subject pool (around 54 percent of the sample), possibly because men may be more likely to have jobs that prevent them from participating in studies during working hours.

The age of registered subjects ranges from 17 to 93 years, with a mean age of 31.34 (SD: 10.37) for all subjects (N=4196), 30.00 (SD: 9.78) for men (N=1908), and 32.50 (SD: 10.69) for women (N=2288). Table 7 summarizes the age composition of the subject pool; the top panel shows the frequency and percentage of Busara subject pool members in various age groups, while the bottom panel shows the data for Nairobi. Busara's age composition matches that of Nairobi almost exactly; this is even true for older respondents, who might be expected to be less likely to sign up with Busara due to the computerized nature of the studies. However, this appears not to have been a deterrent. Women aged 21-40 are again slightly over-represented in the Busara subject pool, suggesting that, as mentioned above, women of this age group may be more flexible with their time than men. Table 8 shows the ethnic composition of Busara subject pool, and compares it with the population of Kenya as a whole. A large proportion of registered subjects belong to the Kikuyu, Luhya and Luo communities, the largest Kenyan tribes. Since a considerable number of Nairobi slum residents are of Luo origin, members of this community tend to be over-represented in Busara subject pool comparing to the Kenyan population.

Table 9 displays education attained by registered slum residents, in comparison with Nairobi and Kenya figures. People registered in Busara's subject pool are somewhat more likely to have some level of education, and to have attained secondary education in comparison with the Nairobi and Kenyan population, while also being somewhat less likely to have attained primary education only. They are relatively less likely to have university education in comparison to the Nairobi/Kenyan population; this may be due to the fact that tertiary education is difficult to attain for slum residents. Women in our subject pool are relatively more likely than men not to be educated, or to have primary education only; in contrast, they are underrepresented at all other levels of education. This is broadly consistent with the patterns found in the data for Nairobi and Kenya.

An analysis of the marital status of registered slum residents (Table 10) shows that a higher proportion of registered subjects (both women and men) are single in comparison to the Kenyan population; also, the Busara subject pool includes a relatively lower proportion of married/cohabiting subjects than the Kenyan average.

Table 11 shows the data for the number of children of registered subjects. The majority of registered women have one to four children, reflecting the statistics of the Kenyan population. By contrast, the majority of registered men are relatively more likely not to have children than Kenyan men in general. Together with the data on marital status, this suggests that single men with no children are slightly over-represented in the Busara subject pool.

8 Special protocols for lab studies at Busara

A central goal of Busara is to enable researchers to conduct computerized behavioral and experimental economics studies with participants from the Nairobi slums. This combination of a high-tech infrastructure with a very special population of respondents has necessitated the development of special methods and protocols, which we summarize briefly below.

8.1 Recruitment through SMS and phone calls

At the beginning of a study, a researcher generates a list of potential participants based on the Busara subject pool database, stratified by demographic variables if desired. Once this list has been generated, two recruitment methods are available to researchers.

In the first, and standard, method, the bulk SMS software FrontlineSMS automatically sends personalized SMS invitations to subsets of participants to invite them to a particular session, indicating date and time of the session, and reminding potential participants of Busara's location. To sign up for participation in the study, participants simply reply "YES" to the SMS; to decline, they do not reply, or reply "NO". Replies are automatically processed and sorted by FrontlineSMS, and a list of confirmed attendees is generated based on the "YES"-responses. Note that participants' inability to read the text messages received from Busara is rarely an issue because it is common practice for illiterate people to have text messages read out to them by literate friends or family members.

The list of confirmed attendees is used on the day of the study when arriving subjects are fingerprinted at the gate to confirm that they are in fact invited for the day's session. The advantages of this recruitment method are that it is a) easy to administer because of its automaticity; convenient for respondents because they usually carry their cell phones with them and can sign up with minimal effort and cost (note, however, that a "Yes" response is more costly in this recruitment method than in the phone call method described below; sending an SMS costs subjects between KES 1-2, i.e. (~USD 0.02-0.03); and c) cheap, because it can be completed easily and quickly by a single RA, and the cost of sending bulk SMS using Safaricom and FrontlineSMS is ~USD 0.15 for an unlimited number of text messages per day.

In the second method, RAs call potential respondents and invite them by phone to participate in sessions. This method is much more labor-intensive and costly than the bulk SMS method, since in our experience it takes a single RA an entire day to recruit a full session of 20 participants. However, the advantage of this method is that respondents can be pre-screened and selected for baseline criteria that are not contained in the Busara database.

8.2 Fingerprinting for identification

All subject pool members are fingerprinted during recruitment, and the fingerprint information is stored in the subject database. When participants arrive at the gate for a study, they are identified by fingerprint: participants put their thumb on the fingerprint scanner, and the fingerprinting software recognizes them and automatically populates the participant's name and age on the screen of the Identify Netbook. The research assistant welcoming participants at the gate then checks this information against the list of invitees for the particular session to ensure that the participant is invited for that session, and to record their attendance. The attendance information is later added to the subjects database to allow tracking of participation. Fingerprinting has three distinct advantages: first, in a high-crime city like Nairobi, it serves as a safety precaution; opening up an environment with expensive computer equipment to any participant from the Nairobi slums above the age of 18 potentially creates vulnerabilities to raids and theft, and anecdotal evidence suggests that the fact that Busara keeps fingerprint records of all participants serves as a powerful deterrent against such activities. Second, fingerprinting is superior to all viable alternatives for identification purposes before studies; in particular, although every Kenyan is in principle required to have a National ID, many of our respondents do not own one, and even if they did, requiring them to bring it both for signing up to the subject pool, and to identify themselves at the beginning of each study, would leave to a large number of potential subjects unidentified and unable to participate. Fingerprinting solves this problem completely; in addition, it is extremely quick, requiring only a few seconds per participant. Finally, fingerprinting at the gate when participants arrive for a study allows us to keep track of which members of our subject pool have participated in which studies; this enables us to track attendance by respondent, study, and sign-up mechanism (phone vs. SMS), and can later be used to selectively exclude participants from participation in particular studies depending on their participation in earlier studies.

8.3 Subject instruction

Another important element in running studies at Busara is to be careful instruction of respondents. Our subject pool members are highly unfamiliar with a computer lab environment in which they are asked to complete decision-making tasks and participate in economic exchange games, and confusion and misunderstanding is frequently the consequence. Below we summarize the main strategies we employ to facilitate optimal understanding of tasks and games.

1. **Briefing before and after the study.** Each lab session begins with a briefing of participants in the waiting room. They are welcomed to Busara and told that they are about to participate in a study about economic behavior and decision-making using touchscreen computers. Invariably participants are worried that they cannot participate due to their unfamiliarity with computers; this initial briefing procedure has proven useful in alleviating these worries and assuring participants that they will have few problems in using the computers, and that

help is available for those who require it. At this stage, subjects are also told to switch their phones off, and asked not to talk to each other during the study, but instead raise their hand if they have a question to attract the attention of an RA who will then come to assist them.

2. **Oral explanation by research assistants.** Each experimental session is run by 2-3 RAs; one of them operates the control computer which displays the tasks and paradigms, while the other two instruct the respondents and answer their questions. Specifically, a second RA stands by the side of the room and reads through the protocol for the experimental session in Swahili, which contains all instructions for the tasks to be performed, including information which is displayed on the screen in English. If respondents have questions, they can raise their hand, and the third RA comes to their cubicle to assist them. Thus, before respondents make any response, they receive detailed information about the choice they are about to make from three sources: first, the information displayed on the screen in English; second, the oral explanations given by the RA in Swahili; third, personalized assistance at their workstation if they require it. Together, these measures have proven effective in facilitating respondent understanding of tasks.

3. **On-screen instruction in writing and through animation and audio.** It has already been described above that on-screen written instructions in English have proven effective in facilitating understanding when combined with oral explanations by an RA in Swahili. In addition to this technique, we have also used custom-built video animations and audio recordings to explain more complex tasks. For instance, Edward Miguel, Simon Galle, Kelly Zhang and collaborators administered a modified Public Goods Game to participants from our subject pool at Busara in the summer of 2012; beacuse this game was somewhat complex, they employed a custom-made animation video (produced with Microsoft PowerPoint), combined with audio explanations recorded by a Busara RA in Swahili, to explain the game to participants both visually and auditorily. This method of explanation was highly effective in terms of respondent understanding, as well as efficient in terms of the RA's time.

4. **Test questions.** After explanations for a task have been given to all respondents, we frequently probe understanding by presenting questions about the paradigm on the screen which respondents have to answer. Two approaches have been used in this context: in one, participants are required to enter the correct answer to proceed to the actual game; this method is desirable if it is essential that all participants understand the task exactly, but it may require extra explanations by the "circulating RA" when respondents cannot advance to the task because they do not know the answer to a probing question. In the other method, participants are free to proceed to the game as soon as they have given a response to the probing question. In this case, it is not guaranteed that they have understood the game entirely, and their data may have to be discounted or excluded later based on this fact; however, this method speeds up study administration considerably because less explaining is required of the RAs.

5. **Practice rounds.** Finally, as is common practice in behavioral economics labs around the

world, nothing clarifies paradigms to subjects quite as well as a few practice rounds. It is important to demarcate the practice rounds very clearly, both orally and in writing on the screen, and make subject aware when the "actual" task starts. In our experience, playing 1-5 practice rounds, depending on the game, is highly effective in promoting understanding.

8.4 Special screen displays

Because many of our respondents are unfamiliar with computers, and in particular, mice and keyboards, studies at Busara depend critically on successful use of the touchscreen by subjects. To this end, we have developed several response paradigms which enable subjects to intuitively and easily enter responses on the touchscreen; we describe them briefly below. Upon request, zTree files with examples for each of these response formats are available from the authors.

1. **Yes/No questions.** For questions which can be answered with a simple binary response, subjects are presented with the question at the top of the screen in English; in addition, it is read out in Swahili by an RA. Below the question on the screen, a red field marked "No" and a green field marked "Yes" are shown; respondents are instructed to press the field whose color corresponds to their answer. Their understanding of colors (see above) suggests that little confusion is likely in making these responses.

2. **Multiple choice questions.** For questions allowing for multiple different answers, subject are presented the question as described above; the answer possibilities appear as a list on the screen, written in English, and are read out by the RA in Swahili. Next to each option is a colored box; the RA instructs subjects to press the colored box corresponding to the answer of their choice.

3. **Graded responses.** To elicit graded responses from participants – e.g. subjective ratings – we use a 1-100 Likert scale. Again the question is presented as described above; below the written question on the screen, a horizontal bar appears, which shows a color gradient from green to blue. When respondents place their finger anywhere on the gradient, an arrow mark appears at the top of the horizontal bar in the exact location where the respondent placed their finger; in addition, the number corresponding to that particular location (1-100) appears above the horizontal bar. Respondents can change the location of the arrow mark as many times as they wish until the desired location has been selected; then they confirm by pressing a red "OK" button displayed below the horizontal bar.

4. **Numerical responses.** As described above, numeric responses can be entered by subjects through a number pad, structured like that of a cell phone, that is displayed on the screen. Subjects enter the number they desire, and then confirm by pressing a green "OK" button. Corrections can be made by pressing a red "Clear" button. Due to the widespread use of cell phones in the Nairobi slums, our respondents are intimately familiar with cell-phone like

number displays; in addition, basic numeracy is excellent (see above), suggesting that entering numerical information poses few problems to our subjects.

8.5 Payment through MPesa

Finally, after a lab session is completed, subjects receive payment for their participation and, in many cases, their responses in economic games and other incentivized tasks. This payment is transferred to their MPesa accounts immediately after the end of the experimental session through a Safaricom corporate account administered by the Finance Team at IPA-K. To complete the transfer, we first download the identification data collected on respondents as they enter the session. This data records the computer at which each respondent sat, and allows us to pay them based on performance and responses, as zTree questions can be assigned payment values and the program will tally up earned payout for each subject individually throughout the session. The identification data also records the phone number the respondent would like payment sent to. For all studies, regardless of whether respondents earned additional money in the games, we pay a base rate of KES 200 (~USD 3) to respondents from Kibera, and KES 400 (~USD 6) to respondents from Viwandani. Each respondent can earn an additional KES 50 (~USD 0.60) shillings for arriving "on time", i.e. 30 minutes before the intended start of the study. If we overbook a session and respondents who have arrived on time have to be turned away, we pay them the full show up fee and bonus. If a respondent is not on time and less than 90 minutes late, they are reimbursed for the cost of transport at a rate of KES 90 (~USD 1.20) for Kibera respondents and KES 200 (~USD 3) for Viwandani respondents. Respondents who are more than 90 minutes late are paid nothing.

It takes on average 2 hours between the end of a session and the receipt of payment in subjects' MPesa accounts. Once in a respondent's MPesa account, funds can be retrieved at one of many agents found in the Nairobi slums and all around Kenya, or sent directly to other MPesa account holders using MPesa. There are both transfer and withdrawal fees for using MPesa; the fees vary depending on the size of the transfer, and Busara covers only the transfer fees, while withdrawal fees have to be borne by respondents. They are, however, much smaller than the minimum show-up fee.

MPesa is uniquely suited to the needs of Busara for three reasons. First, virtually all residents of the Nairobi slums use it; recent data on cell phone usage in Kenya by Tavneet Suri suggests that upwards of 90% of Nairobi slum residents have access to MPesa. In addition, to receive payment for studies completed at Busara, subjects need in fact not have MPesa access themselves; it is sufficient if they provide the cell phone number of a friend or relative to whose account the money can be transferred. Thus, using this method of payment excludes virtually nobody from participating in Busara studies. Second, MPesa is much safer than paying in cash: the latter method would require keeping relatively large amounts on-site at Busara, which would additionally increase the risk of raids or theft. Processing payments through MPesa entirely eliminates this risk. In addition, IPA-K's corporate MPesa account is operated with an additional security feature: after payment data is entered into Safaricom's online MPesa interface by the IPA-K finance team in Nairobi immediately after a session, it first has to be reviewed and approved by a second member of IPA-K's finance team, working in another location several hundred miles outside of Nairobi. Thus, a raid of the Nairobi office alone would not be sufficient to force MPesa transfers to be made. Finally, the use of MPesa for payment is highly convenient from the point of view of particular paradigms in experimental and behavioral economics, in particular those which require delayed payments. For instance, intertemporal choice tasks routinely require payments to be made after a delay; this is not only easy to implement with MPesa, but the fact that both immediate and delayed transfers are made through MPesa additionally controls for a familiar problem in intertemporal choice tasks, namely holding transaction costs constant. With MPesa, transaction costs are exactly identical, making this payment mechanism superior to many other methods standardly used in this literature.

9 A Busara Session from Beginning to End

Busara is open to researchers from institutions around the world; we hope that it will provide an infrastructure for economists, psychologists, political scientists, sociologists, anthropologists, and other social scientists to better understand the behavior and preferences of people living in poverty. In the following we summarize a complete experimental session at Busara to give researchers a sense of what they can expect when running studies here. This protocol is flexible, however, and can be adapted to diverse experimental needs.

9.1 Overall setup and throughput

A typical study session at Busara Center takes 3 hours on average, and we typically run 2 sessions per day: a morning session (9am-12pm) and an afternoon session (1pm-4pm). Running sessions earlier or later is problematic due to the significant travel time required in Nairobi for both RAs and respondents; in particular, travel after dark is dangerous, and therefore a timely end to the day's sessions is imperative.

Each session can accommodate 20 participants; we typically obtain between 26-28 "Yes" responses for a given session to ensure that the session gets filled with 20 participants. With 2 sessions a day, this throughput corresponds to a theoretical maximum of 200 participants per week and 800 participants per month. These numbers can decrease somewhat due to unforeseen circumstances that arise during the lifecycle of a study; however, Busara can operate very close to the theoretical capacity maximum: in the summer of 2012, Busara ran 620 subjects between mid-July and mid-August.

9.2 Gate identification procedure

A study session is preceded by participant identification at the gate, which commences 1 hour before the scheduled start time of the study. Respondents are asked to arrive at least 15 minutes before start time; the KES 50 (~USD 0.60) incentive for arriving on time has greatly increased respondent punctuality.

Upon arrival at the compound, respondents are identified by fingerprints as described above; specifically, the Identify program identifies respondents' fingerprints and then reads respondent information from an external database that is pre-loaded before each study with the names and Survey IDs of only those subject pool members who were invited to the day's session. During the identification procedure at the gate, personal information such as marital status, number of children, phone number, or education level of the invitees is double-checked and updated if necessary. The Identify software also randomly assigns them a place number, and each respondent receives a laminated card with their place number; as described above, these cards are additionally color-coded and marked with symbols, and together the colors and symbols on each card uniquely identify the computer at which the respondent will be sitting. The rationale for randomizing seating is to avoid friends sitting close to each other and potentially colluding during studies. In addition, after identification, each respondent receives a security badge allowing them to enter the compound. They are then escorted by a staff member to the Busara Center waiting room.

If more than the maximum number of participants specified for a particular session arrive on time, those subjects who arrive last are turned away; however, they are paid the full show-up fee and a bonus. If subjects arrive less than 90 minutes late and have to be turned away, they are paid a transport allowance. If a respondent is late by more than 90 minutes, they receive no payment.

9.3 Waiting room protocol

Once all invitees are registered and seated at the waiting room, a welcome speech is given by a Busara staff member. Participants are reminded about the compensation they will receive after the session for their transport and time; they are also told whether the particular session in which they are about to participate will involve tasks in which they can earn additional income. Phones are required to be turned off before participants enter the testing room. In addition, respondents are not allowed to talk to each other once they have entered the computer lab; if questions arise, they are requested to raise their hands, upon which an RA will come to assist them. We also encourage participants to use the toilets before the session begins. After this welcome briefing, participants are shown into the computer room and asked to sit down next to the computer with the number indicated on their place cards.

Many of our respondents have young children to take care of; we advise these participants to leave their children at home, or bring along someone who will watch over the children during the session. In cases where neither is possible (an average of 1-2 participants per session of 20), they have the choice of sitting at the workstation with their babies as long as the latter are not disruptive, or of Busara staff watching over the children in the waiting room.

9.4 Computer room protocol

Once in the computer room, consent forms are handed out containing more information about the study; participants have a choice of English or Swahili forms. The Busara RA running the session goes over the consent form orally in Swahili; after any questions have been answered, participants are asked to sign the form. Participants who are unable to read or write provide fingerprints.

After consent forms have been signed and collected, the RA gives an introduction on how to use the touchscreen computers to make responses in the experimental tasks. The RA then goes through the study protocol, explaining each task, asking respondents to answer test questions and complete practice rounds. In some cases paradigms are explained through on-screen animations and audio explanations delivered through headphones. Participants who do not understand the tasks or need help using the workstations are assisted by Busara staff. Once the instructions, test questions, and practice rounds are completed, participants perform the actual experimental tasks.

9.5 Leaving protocol

Once all experimental tasks are completed, subjects are informed of their final payout through a display on their screens. The RA thanks them for participation and reminds them that they will be paid through MPesa later the same day. Participants who want to update their mobile payment phone number are asked to do so before they leave. Those who have given third party numbers (relatives, friends, spouses) are reminded to approach the people to whose numbers they asked their payment to be transferred to collect the payment. Participants are then asked to leave their workstations; Busara staff collect the place cards and security tags as the respondents are escorted to the gate. In some studies, a few participants are randomly selected to remain behind for an exit debrief interview lasting 5-10 minutes.

9.6 Stratification and subject-specific treatments

One of the advantages of recording a number of demographic variables is that recruitment lists are generated in different ways depending on the study. We can target a specific pool of people (i.e. those who have completed primary school, are between 20-30 years old, are married, etc.) and can stratify by any of the demographic variables we collect during recruitment if a study needs a particular composition of respondents. Stratified sampling of course does not lead to a perfectly balanced sample, since to fill a session of 20 respondents we normally invite about 35 people, and the number of people from each stratum who actually attend will vary somewhat. Averaging over a number of sessions, however, we have been able to achieve close to perfect balance for stratified studies.

If study treatments vary depending on a particular demographic characteristic, a list matching which treatment they are to receive to the client number of the computer they are participating on must be made at the start of each session. When participants arrive at the session, their thumbprint is scanned and they are assigned to a random computer number. Once all participants have arrived, the identification data can be downloaded, converted into Stata format, merged by fingerprint ID with the recruitment data, and then exported again in zTree format, assigning each respondent to the correct treatment. For instance, if treatment varies by gender, the program can produce a .csv file that can be copied and pasted directly into a zTree file to generate subject-specific treatments that vary by demographic variable.

10 Conclusion

Laboratory research in behavioral and experimental economics in developing countries is still in its infancy, and the effort described here is only a first step towards enabling social scientists to conduct rigorous laboratory studies with respondents from the world's poorest communities. It is our hope that social scientists from around the world will make intensive use of the infrastructure and opportunities that Busara offers, and that the insights generated by this work will bring us closer towards better understanding the behavior and preferences of people living in developing countries. Ultimately, we hope that this knowledge will translate into interventions that can be used to improve their welfare. Researchers interested in running studies at Busara are encouraged to contact the first author (joha@mit.edu).

References

Arnett, Jeffrey J. 2008. "The neglected 95%: Why American psychology needs to become less American." *American Psychologist* 63 (7): 602–614.

Cardenas, Juan Camilo, and Jeffrey Carpenter. 2008. "Behavioural Development Economics: Lessons from Field Labs in the Developing World." *The Journal of Development Studies* 44 (3): 311–338.

Frederick, Shane. 2005. "Cognitive reflection and decision making." *Journal of Economic Perspectives* 19 (4): 25–42.

Glimcher, P. W. 2009. *Neuroeconomics: Decision making and the brain.* Academic Press.

Glimcher, Paul W. 2003. "Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics." March.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Herbert Gintis, Richard McElreath, and Ernst Fehr. 2001. "In Search of Homo Economicus: Experiments in 15 Small-Scale Societies." Artefactual field experiments 0059, The Field Experiments Website.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "Most people are not WEIRD." *Nature* 466 (7302): 29–29 (July).

Kahneman, D., and A. Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica: Journal of the Econometric Society*, p. 263–291.

Lei, Vivian, and Charles N. Noussair. 2002. "An Experimental Test of an Optimal Growth Model." *American Economic Review* 92 (3): 549–570.

Raven, John, and Jean Raven. 2003. Chapter Raven progressive matrices of *Handbook of Nonverbal Assessment*, edited by R. Steve McCallum, 1, 223–237. Springer US.

Schultz, Wolfram. 2006. "Behavioral Theories and the Neurophysiology of Reward." *Annual Review of Psychology* 57 (1): 87–115.

Smith, Vernon. 1982. "Microeconomic Systems as an Experimental Science." *The American Economic Review* 72 (5): 923–955.

Smith, Vernon L. 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy*, vol. 70.

Stanovich, Keith E., and Richard F. West. 2000. "Individual differences in reasoning: Implications for the rationality debate." *Behavioral and Brain Sciences* 22 (5): 645–726.

Table 1: Screen tasks

	(1) Obs.	(2) Mean	(3) Std. Dev.
Choose left	71	0.94	0.23
Choose right	71	0.97	0.17
Choose up	71	0.99	0.12
Choose down	71	0.96	0.20
Choose red	71	0.99	0.12
Choose green	71	0.96	0.20
Choose black	71	0.99	0.12
Choose blue	71	1.00	0.00
Choose yellow	71	1.00	0.00
Choose cow	71	1.00	0.00
Choose hammer	71	0.94	0.23
Choose tree	71	0.99	0.12
Choose chair	71	1.00	0.00
Choose car	71	0.97	0.17

Notes: Results of simple screen tasks run on subsample of respondents, asking them to identify directions, colors, and pictures using the touch screen interface.

Table 2: Arithmetic tasks

	General Population			Students			Difference	
	(1) Obs.	(2) Percent Correct	(3) Mean RT	(4) Obs.	(5) Percent Correct	(6) Mean RT	(7) Percent Correct	(8) Mean RT
One-digit addition	154	0.92	18.94	104	1.00	10.67	−0.08	8.27
		(0.28)	(11.13)		(0.00)	(5.87)	(0.00)***	(0.00)***
One-digit subtraction	187	0.92	18.25	104	1.00	9.06	−0.08	9.18
		(0.27)	(16.03)		(0.00)	(6.02)	(0.00)***	(0.00)***
One-digit multiplication	154	0.92	12.40	104	1.00	6.96	−0.08	5.44
		(0.28)	(9.34)		(0.00)	(3.52)	(0.00)***	(0.00)***
One-digit division	154	0.79	16.35	104	0.98	7.10	−0.20	9.24
		(0.41)	(13.87)		(0.14)	(4.19)	(0.00)***	(0.00)***
Two-digit addition	154	0.90	26.30	104	0.98	12.79	−0.08	13.52
		(0.31)	(29.19)		(0.14)	(6.19)	(0.00)***	(0.00)***
Two-digit subtraction	154	0.78	27.37	104	1.00	13.88	−0.22	13.49
		(0.42)	(26.24)		(0.00)	(7.75)	(0.00)***	(0.00)***
Two-digit multiplication	154	0.46	61.35	104	0.85	46.58	−0.39	14.77
		(0.50)	(49.24)		(0.36)	(22.08)	(0.00)***	(0.00)***
Two-digit division	154	0.61	51.55	104	0.96	25.94	−0.35	25.61
		(0.49)	(48.10)		(0.19)	(18.66)	(0.00)***	(0.00)***
Mean		0.57	22.81		0.76	13.45		

Notes: Results of simple arithmentic tasks as presented to general population and student test groups. Columns 1 and 4 report the number of observations for each group, columns 2 and 4 report the respective proportion of correct answers for each question as well as the standard deviations, and columns 3 and 5 report the mean and standard deviation of the response times. Columns 6 and 7 report the mean differences between these variables for the general population and student groups, in addition to an unpaired two-sample t-test for difference in means with unequal variances. The bottom row presents the means across all questions in the set. * denotes significance at the 10 pct. level, ** at the 5 pct. level, and *** at the 1 pct. level.

Table 3: Raven's progressive matrices

	General Population			Students			Difference	
	(1) Obs.	(2) Percent Correct	(3) Mean RT	(4) Obs.	(5) Percent Correct	(6) Mean RT	(7) Percent Correct	(8) Mean RT
Raven's matrices 1	154	0.67	92.67	104	0.88	43.38	−0.22	49.29
		(0.47)	(65.90)		(0.32)	(21.56)	(0.00)***	(0.00)***
Raven's matrices 2	154	0.49	49.19	104	0.88	39.65	−0.38	9.54
		(0.50)	(29.22)		(0.33)	(18.58)	(0.00)***	(0.00)***
Raven's matrices 3	154	0.31	51.99	104	0.56	75.67	−0.25	−23.69
		(0.46)	(28.56)		(0.50)	(28.59)	(0.00)***	(0.00)***
Mean		0.49	64.61		0.77	52.90		

Notes: Results of Raven's progressive matrices tasks as presented to general population and student test groups. Columns 1 and 4 report the number of observations for each group, columns 2 and 4 report the respective proportion of correct answers for each question as well as the standard deviations, and columns 3 and 5 report the mean and standard deviation of the response times. Columns 6 and 7 report the mean differences between these variables for the general population and student groups, in addition to an unpaired two-sample t-test for difference in means with unequal variances. The bottom row presents the means across all questions in the set. * denotes significance at the 10 pct. level, ** at the 5 pct. level, and *** at the 1 pct. level.

Table 4: Cognitive reflection test

	General Population			Students			Difference	
	(1) Obs.	(2) Percent Correct	(3) Mean RT	(4) Obs.	(5) Percent Correct	(6) Mean RT	(7) Percent Correct	(8) Mean RT
CRT 1	154	0.04	60.18	104	0.30	74.46	−0.26	−14.27
		(0.19)	(34.39)		(0.46)	(52.05)	(0.00)***	(0.02)**
CRT 2	154	0.06	88.10	104	0.19	88.94	−0.13	−0.84
		(0.25)	(43.08)		(0.40)	(54.22)	(0.00)***	(0.90)
CRT 3	154	0.01	94.85	104	0.28	102.93	−0.27	−8.08
		(0.11)	(44.05)		(0.45)	(46.15)	(0.00)***	(0.16)
Mean		0.04	81.04		0.26	88.78		

Notes: Results of the Cognitive Reflection Test as presented to general population and student test groups. Columns 1 and 4 report the number of observations for each group, columns 2 and 4 report the respective proportion of correct answers for each question as well as the standard deviations, and columns 3 and 5 report the mean and standard deviation of the response times. Columns 6 and 7 report the mean differences between these variables for the general population and student groups, in addition to an unpaired two-sample t-test for difference in means with unequal variances. The bottom row presents the means across all questions in the set. * denotes significance at the 10 pct. level, ** at the 5 pct. level, and *** at the 1 pct. level.

Table 5: Location distribution of subject pool

Kibera	**Freq.**	**%**	**Viwandani Greater Area**	**Freq.**	**%**
Kianda	707	16.85	Lunga Lunga	569	13.56
Makina	654	15.59	Sinai	261	6.22
Gatwekera	239	5.70	Donholm	208	4.96
Lindi	168	4.00	Kingston	168	4.00
Soweto	142	3.38	Highridge	157	3.74
Mashimoni	135	3.22	Jamaica	134	3.19
Raila	130	3.10	Riverside	85	2.03
Kisumu Ndogo	111	2.65	Viwandani	61	1.45
Laini Saba	86	2.05	Westlands District	18	0.43
Siranga	61	1.45	Kawangware	11	0.26
Undugu	1	0.02	Other	90	2.14
Total Kibera	2,434	58.01	Total other slums	1762	41.98
Total				4196	

Notes: Distribution of the Busara subject pool across slums and villages within slums. Kibera locations are shown in the left columns, other slums on the right.

Table 6: Gender composition of subject pool

	Busara		**Nairobi**		**Kenya**	
Gender	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**
Male	1,908	45.47	1,605,230	51.15	19,192,458	49.71
Female	2,288	54.53	1,533,139	48.85	19,417,639	50.29
Total	4,196	100	3,138,369	100	38,610,097	100

Notes: Busara data are shown in the leftmost columns, data for Nairobi and Kenya in the middle and on the right. nairobi and Kenya data were obtained from Nairobi and Kenya statistics from *https://opendata.go.ke/Population/Census-Volume-1-Question-1-Population-Households-a/wd27-eki2*.

Table 7: Age composition of subject pool

	Whole sample		**Men**		**Women**	
Age	**Freq.**	$\% \mid > 21$	**Freq.**	**$\% \mid > 21$**	**Freq.**	**$\% \mid > 21$**
			Busara			
18-20	341		167		174	
21-30	2,085	54.10	1,063	61.09	1,022	48.34
31-40	1,019	26.44	425	24.43	594	28.10
41-50	470	12.20	152	8.74	318	15.04
51-60	208	5.40	75	4.31	133	6.29
61-70	67	1.74	23	1.32	44	2.08
71-80	4	0.10	1	0.06	3	0.14
80+	1	0.03	1	0.16	0	0.00
Total	4,195	100	1,907	100	2,288	99.99
			Nairobi			
21-30	940,149	49.13	494,464	55.10	445,685	43.85
31-40	553,761	28.94	236,679	26.38	317,082	31.20
41-50	253,604	13.25	98,594	10.99	155,010	15.25
51-60	105,861	5.53	39,875	4.44	65,986	6.49
61-70	38,188	2.00	15,769	1.76	22,419	2.21
71-80	13,645	0.71	6,759	0.75	6,886	0.68
80+	8,523	0.45	5,175	0.58	3,348	0.33
Total	1,913,731	100	897,315	100	1,016,416	100

Notes: Busara data are shown in the top panel, data for Nairobi in the bottom. Separate data for men and women are shown across columns. Nairobi data were obtained from *https://opendata.go.ke/Population/Nairobi-Pop-Pyramid-Age-Groups-2009/u6e3-zmvi*.

Table 8: Ethnic composition of the subject pool

	Busara		**Kenya**	
Ethnicity	**Freq.**	**%**	**Freq.**	**%**
Kikuyu	1,057	25.04	6,622,576	17.15
Luhya	822	19.47	5,338,666	13.83
Luo	809	19.16	4,044,440	10.48
Kamba	640	15.16	3,893,157	10.08
Kisii	346	8.20	2,205,669	5.71
Nubi	227	5.38	15,463	0.04
Kalenjin	66	1.56	4,967,328	12.87
Borana	63	1.49	161,399	0.42
Meru	61	1.44	1,658,108	4.29
Taita	23	0.54	273,519	0.71
Embu	18	0.43	324,092	0.84
Other	90	2.13	9,105,680	23.58
Total	4,222	100	38,610,097	100

Notes: Busara data are shown in the left columns, data for Kenya on the right. Kenya data were obtained from the 2009 Population and Housing Census conducted by the Kenya national Bureau of Statistics *http://www.knbs.or.ke/censusethnic.php*.

Table 9: Education levels of subject pool

	Busara			**Nairobi**		**Kenya**	
Education Level	**Whole Sample**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
			Percentages				
None	2.97	1.93	0.68	5.53	5.57	6.7	16.2
Pre-school	0.17	0.19	0.21	5.59	5.42		
Primary	47.81	38.32	26.94	37.34	36.55	47.1	50.2
Secondary	39.99	45.04	51.10	32.61	31.97	31.5	22.9
College	5.33	7.24	9.54	10.48	10.89	14.7	10.7
University	3.72	7.27	11.53	7.87	9.05		
Other				0.28	0.28		
Total	99.99	99.99	100	100	100	100	100

Notes: Busara data are shown on the left columns, data for Nairobi and Kenya on the right, separately for men and women. Empty cells in the Nairobi and Kenya data indicate that the sources used did not provide data for the particular education level. Nairobi data were obtained from *https://opendata.go.ke/Education/Population-3-years-and-above-by-Sex-and-Highest-Le/x4e7-whsh*; Kenya data were obtained from the 2008-2009 DHS.

Table 10: Marital status of subject pool

	Busara						Kenya			
	Whole sample		**Women**		**Men**		**Women**		**Men**	
Marital status	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**
Single	2,005	47.79	938	41.00	1,067	56.00	853	13.6	598	25.6
Married or cohabitating	1,881	44.84	1,093	48.00	788	41.32	4,593	73.22	1,617	69.22
Divorced, separated, widowed	309	7.37	257	11.23	52	2.73	827	13.18	121	5.17
Total	4,195	100	2,288	100	1,907	100	6,273	100	2,336	100

Notes: Busara data are shown in the left columns, data for Kenya on the right, separately for men and women. Kenya data were obtained from the 2008-2009 DHS.

Table 11: Composition of subject pool by number of children

	Busara						**Kenya**			
	Whole sample		**Women**		**Men**		**Women**		**Men**	
Number of children	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**	**Freq.**	**%**
None	1,427	34.02	432	18.88	995	52.18	722	11.51	699	29.92
1 or 2	1,567	37.35	988	43.18	579	30.36	2,131	33.97	688	29.45
3 or 4	778	18.55	563	24.61	215	11.27	1,858	29.62	549	23.5
5 or more	423	10.08	305	13.33	118	6.19	1,562	24.9	400	17.12
Total	4,195	100	2,288	100	1,907	100	6,273	100	2,336	100

Notes: Busara data are shown on the left, data for Kenya on the right, separately for men and women. Kenya data were obtained from the 2008-2009 DHS.