The Sandbox Model: A Novel Approach to Iterating while Implementing an Emergency Education Program in Lebanon during the COVID-19 Pandemic

Author(s): Michèle Boujikian, Alice Carter, and Katy Jordan

Source: Journal on Education in Emergencies, Vol. 8, No. 3 (December 2022), pp. 215-228

Published by: Inter-agency Network for Education in Emergencies

Stable URL: http://hdl.handle.net/2451/63993

DOI: https://doi.org/10.33682/rj45-k7z7

REFERENCES:

This is an open-source publication. Distribution is free of charge. All credit must be given to authors as follows:


The Journal on Education in Emergencies (JEiE) publishes groundbreaking and outstanding scholarly and practitioner work on education in emergencies (EiE), defined broadly as quality learning opportunities for all ages in situations of crisis, including early childhood development, primary, secondary, non-formal, technical, vocation, higher and adult education.

Copyright © 2022, Inter-agency Network for Education in Emergencies.

The Journal on Education in Emergencies, published by the Inter-agency Network for Education in Emergencies (INEE), is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, except where otherwise noted.
THE SANDBOX MODEL: A NOVEL APPROACH TO ITERATING WHILE IMPLEMENTING AN EMERGENCY EDUCATION PROGRAM IN LEBANON DURING THE COVID-19 PANDEMIC

Michèle Boujikian, Alice Carter, and Katy Jordan

ABSTRACT

Jusoor’s Refugee Education Program helps Syrian refugee children living in Lebanon integrate back into formal schooling. When schools closed due to the COVID-19 pandemic, the refugee program adapted to distance learning by developing Azima, a novel program that used WhatsApp to enable children to keep learning. Azima had to respond immediately to the emergency context while maintaining high education standards, and it also needed to find an effective way to test and refine its content quickly. To do this, the Azima program adopted an innovative experimental approach called a sandbox. A sandbox model operates in rapid iterative cycles and uses multiple methods to quickly test a program’s assumptions about how it will meet its goals. In this field note, we use Azima as a case study to report on our experience of applying the sandbox model. We reflect on the benefits and limitations of this novel approach in supporting the use of education technology in a crisis situation.

INTRODUCTION

The outbreak of the COVID-19 pandemic in March 2020 had consequences that reached far beyond the health sector. Education was hit particularly hard, with school closures being one of the most widely implemented policy responses (Hale et al. 2021). According to the World Bank (2021), “COVID-19 has created the worst crisis to education and learning in a century,” disrupting school-based education.
for 1.6 billion children worldwide. Education technology (ed tech) solutions were often at the center of rapid responses to this crisis, as many governments turned to remote learning options. Devices replaced classrooms, starting a new education modality that children, teachers, and parents had to adjust to quickly.

Due to the scale and speed of the emergency, program evaluation was often deprioritized. The emergency context also exposed some existing challenges in ed tech, such as inappropriate evaluation techniques and a lack of accessible and understandable research (King et al. 2016; Cukurova, Luckin, and Clark-Wilson 2019). There clearly was a need for evaluation and impact studies that would generate evidence rapidly enough for practitioners to use when making decisions in situations where they were asked to act immediately (Tauson and Stannard 2018).

In this field note, we introduce the sandbox model, which embeds research and evidence-informed decisionmaking into program implementation. It uses rapid iterative cycles to evaluate an intervention and triangulate understanding of what is working and where barriers persist throughout the implementation. We present the sandbox model through a case study of Azima, a program of Jusoor, which used WhatsApp to support refugee education in Lebanon during the COVID-19 emergency. In the next section, we discuss the rationale behind the need to use different evaluation and research approaches in ed tech and introduce the sandbox methodology. We then demonstrate how the components of the sandbox model were concretely applied in the context of Azima. Finally, we reflect on our experience of carrying out a sandbox approach and the usefulness of this approach for making rapid decisions, and for the education in emergencies field in particular.

THE NEED TO TAKE A DIFFERENT APPROACH

Education practitioners aim to find the most cost-effective interventions to improve learning outcomes. Historically, this has meant relying on costly and time-consuming evaluation approaches. One option is randomized controlled trials, which are regarded as the gold standard for generating causal evidence. However, decisionmakers—particularly those in low-resource, highly uncertain settings—are increasingly seeking other forms of research appropriate to these contexts (Crawfurd et al. 2021).

---

1 Jusoor is an international nongovernmental organization whose mission is stated as “Investing in Syria’s Children & Youth for a Better Tomorrow.”
Furthermore, there often is a disconnect between the academic world that produces evidence and the practitioners who are expected to act on it. Evidence generated by evaluation studies that have been conducted by researchers can feel inaccessible to practitioners, which results in its limited use for decisionmaking (Hennessy et al. 2021). Evaluations also commonly occur at the end of a project and thus assess the impact of the activities concluded, rather than intending to inform future iterations of the intervention.

With this in mind, we argue for the need to consider methodologies that are better suited to supporting ed tech implementers. Best practice in the digital realm tends to encourage practitioners themselves to take the lead in generating evidence, to work in multidisciplinary teams, and to conduct formative assessments that will inform future iterations effectively and in a timely manner. These methods and approaches can be borrowed and adapted for the education sector.

**THE SANDBOX METHODOLOGY**

The term “sandbox” has been used for decades in software engineering to describe a space that enables developers to test new code before rolling it out. The concept also has been used in other sectors, including by the EdTech Hub, which has developed the sandbox methodology to focus on its application in ed tech initiatives (Rahman et al. 2021).

The sandbox method creates a space in which interventions operating in uncertain conditions can be tested and iterated. By using a combination of research and design methods, the approach enables its implementers to put their findings into practice and then to re-evaluate them quickly and iteratively. To enable learning, adapting, and scaling to take place throughout implementation, the sandbox combines elements from backcasting (Robinson 1990), lean start-up (Chang 2019), user-centered design, agile methodology (Kaiser 2019), and behavioral innovation (Simpson 2019).

The key steps in the sandbox method are as follows:

1. **Articulate the desired impact of the intervention and surface critical assumptions**
The primary focus of the sandbox model is to address a problem and have an impact, rather than to assess a product or solution. To do this, a team of stakeholders must first articulate a hypothesis for the intervention; this is a backcasting methodology that orients the team’s effort around a desired future, from which it then works backwards toward the current state, rather than making marginal gains forward. The team determines the hypothesis or “big idea” and then considers what must be true for the idea to work. This can be done using a design hypothesis format—“If we...then...”—or by creating a theory of change for the intervention. The next step in the sandbox approach is to collect evidence to validate (or invalidate) the assumptions underpinning the hypothesis. An assumption is something that needs to be true in order for the intervention to work (Chang 2019).

Education technology exists within complex and often fragmented systems. To account for this, the EdTech Hub proposes a model that goes beyond testing the technical components of an intervention and considers the broader system of factors that need to work together to have an impact at scale. By using the 6Ps Framework—people, product, pedagogy, policy, place, and provision (Plaut et al. 2020)—practitioners are encouraged to identify assumptions across all aspects of the education system that might affect the intervention (see Figure 1).

Figure 1: The 6Ps: A Framework for Considering the Education System
2. Design lean experiments to test the assumptions

The assumptions are prioritized from most to least critical; the most critical assumptions are those that would undermine the entire project if proved invalid. Experiments are then designed to test the most critical assumptions. An experiment is the smallest batch of work that can be done to get informative feedback to iterate an intervention (Murray and Ma 2015). This approach is drawn from the lean method, an approach to new product development that prioritizes feedback loops and accelerates the pace of learning about what works (Chang 2019). Practitioners themselves design and conduct the experiments and draw from all relevant data-collection methods: qualitative or quantitative, primary or secondary, descriptive or experimental. Experiments vary in length. At the start of a sandbox, when uncertainty about the merits of an idea is still high, experiments typically are short and inexpensive; the financial and time investment increases in keeping with growing confidence in the validity of the assumptions and the efficacy of the intervention. Finally, experiments are designed to generate measurable and clear insights into people’s real behavior rather than their opinions or claims.

3. Reflect, learn, and iterate at regular intervals

The sandbox model centers on the importance of iteration and adaptation throughout the implementation process. The approach includes formal moments during the program that allow practitioners to focus on learning and make space to change plans. Experiments are packaged into “sprints,” a term used in agile methodologies to describe “a short, time-boxed period when a scrum team works to complete a set amount of work” (Rehkopf 2022). Between each sprint, the team reflects on what has been learned and uses the information to review and redesign the next sprint.

THE AZIMA SANDBOX IN PRACTICE

Jusoor has supported out-of-school Syrian refugee children living in Lebanon since 2013 through its Refugee Education Program, with the aim of getting them back into formal schooling. Lebanon is the country with the world’s highest number of refugees per capita (UNHCR 2022). Since its onset in 2011-2012, the refugee crisis has been synonymous with an education crisis. United Nations data for the last pre-COVID academic year (2018-2019) put the number of school-

---

3 A scrum team designates practices that emphasize daily communication and a flexible reassessment of plans.

4 To date, Jusoor has enrolled a total of 4,631 Syrian refugee children in its schools.
age Syrian refugees at 666,491; of these, only 42 percent were attending school (Norwegian Refugee Council 2020, 4).

Despite attempts to provide access to education for Syrian refugees, the Lebanese public school system has struggled to absorb the substantial number of school-age children. Through the provision of nonformal education, nongovernmental organizations, including Jusoor, have played an important role in ensuring that refugee children do not become part of a lost generation.

Jusoor’s Refugee Education Program is typically facilitated through three education centers. However, following the outbreak of the COVID-19 pandemic in March 2020, Jusoor switched to online learning by creating a new program that used WhatsApp. This new program was called Azima, which means “determination” in Arabic, referring to the children's determination to continue learning.

The Azima program was developed quickly and it relied on teachers to send video lessons and materials to their students via WhatsApp. Students were sent homework to complete and return for correction and feedback. Teachers recorded attendance through completed homework submissions. The goal was to preserve children’s current learning opportunities and their future chance to receive quality education. Overall, the initial WhatsApp model appeared to work, but some students’ attendance was poor, which suggests that they had difficulty engaging with the program.

Jusoor and the EdTech Hub formed a partnership to test and refine Azima using the sandbox method while responding to the urgency of the crisis. The partners began by articulating the following hypothesis: “If we provide lessons and assignments via WhatsApp to out-of-school refugee children at the primary level and engage their caregivers, then children will be able to continue learning and will have a greater chance of accessing formal education in the future.”

The assumptions were formulated and the experiments were designed through a series of joint workshops between Jusoor and the EdTech Hub. The organizations’ staff used the 6Ps Audit Tool (EdTech Hub 2022) to reflect on each element of the education system and to assess the level of certainty in the different parts of the intervention. They articulated assumptions by taking into account answers to questions related to the 6Ps, the intended goal of the program, and the team’s observations and results from Azima. Each component of the 6Ps Framework was associated with critical assumptions, which workshop participants ranked from the most to the least critical. The most critical were prioritized for experiments that were designed to test each of them. Developing the experiments relied on
a combination of the field and research knowledge of Jusoor’s staff and the innovation expertise of the EdTech Hub team.

The full set of six experiments and how they map onto critical assumptions within the 6Ps Framework is shown in Table 1.

Table 1: Overview of the Critical Assumptions, Experiments, and Research Methods

<table>
<thead>
<tr>
<th>6Ps Framework</th>
<th>Assumptions</th>
<th>Experiments and Research Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedagogy</strong></td>
<td>Teachers are technically and pedagogically equipped to deliver remote learning, including curating and creating content themselves that will have maximum impact</td>
<td>Understanding teachers’ practices in action through <strong>classroom observation and evaluation</strong></td>
</tr>
<tr>
<td><strong>Product</strong></td>
<td>WhatsApp is the best platform for online learning for our beneficiaries in the context of school closures, or for out-of-school children</td>
<td>Testing the suitability of WhatsApp, explored through a <strong>tool criteria evaluation</strong></td>
</tr>
<tr>
<td><strong>People</strong></td>
<td>Parents are interested in supporting the education of their children remotely if given the resources and support</td>
<td>Testing how to alleviate barriers to engagement for learners, which involved conducting a <strong>survey of nonengaged students and their caregivers</strong> and testing two interventions: a short <strong>information campaign</strong> and a <strong>one-off cash grant</strong></td>
</tr>
<tr>
<td><strong>Place</strong></td>
<td></td>
<td>Modeling the costs of the program by applying <strong>cost-effectiveness modeling</strong> tools</td>
</tr>
<tr>
<td><strong>Provision</strong></td>
<td>WhatsApp, in combination with additional support, can be used cost-effectively to ensure learning outcomes across all groups</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The sixth P, policy, was considered out of scope, given the timing of this sandbox.

The Azima sandbox extended from September 2020 to March 2021 and included two sprints (Figure 2).
Sprint 1 focused on exploring the assumptions across the 6Ps Framework to understand which had the highest level of uncertainty and where it would be most useful to invest additional resources in further investigation.

**Sprint 1 Experiments**

**Experiment 1: Provision—Cost Modeling**

**Method:** We collected data on the program’s direct costs and expenditures, as well as projected costs for alternative models, in order to weigh the costs of different interventions (e.g., take-home learning supplies, cash, internet, devices, internet and devices, and information campaign interventions).

**Result:** The highest costs were data plans and teachers’ salaries, both essential to the model’s effectiveness, thus only limited savings could be achieved without further exploration of the different needs and elements of the Azima program. Additional work would be needed to understand how to reduce costs in order to attract donors while maintaining student engagement and learning.

**Experiment 2: Pedagogy—Virtual Classroom Observations**

**Method:** The WhatsApp classroom group was monitored over the course of one week. Each of Jusoor’s 29 teachers was observed by two evaluators, one each from the sandbox team and the Jusoor staff. The evaluators used a customized classroom observation tool based on the World Bank’s (2019) Teach Primary tool. The scores and evaluators’ notes were compared to the results of the previous academic year (when distance learning started) to assess the progress achieved.

---

5 “Internet” refers to the provision of an internet-enabled data plan only. “Devices” refers to the provision of device hardware. “Internet and devices” means providing both the data plan and device hardware.
Result: An overall improvement was observed. 14 teachers were rated “Good,” 11 obtained a “Very good” score, and 4 teachers were rated “Excellent,” which suggests that the teachers were applying the training they received. This provided evidence that, with adequate training and consistent follow-up, it is possible to build teachers’ capacity to transition to remote learning.

Experiment 3: Product—Testing the Suitability of WhatsApp

Method: WhatsApp was chosen as the distance learning platform because of its prevalence and familiarity among the refugee community in Lebanon. To assess its broader suitability for the program, a set of desirable criteria was developed based on existing research (Jordan and Mitchell 2020) and applied by three evaluators from the sandbox team. The criteria focused on function—what we wanted the ideal tool to do—and on contextual factors—what context-specific factors might affect the choice of a tool. A total of 23 criteria were listed and evaluated to determine whether or not Azima’s current education provision met them (Khalayleh 2021).

Result: Most of the criteria were rated as being met by Azima’s current provision; only two were unachievable. This indicated that WhatsApp could be used to implement distance learning and that most improvements could be explored without changing the tool. However, it also suggested that there was a role for complementary tools to fill some of the gaps, such as Google Drive for organizing content.

Experiment 4: People and Place—Learning Barriers Survey

Method: To understand the reasons behind some students’ nonengagement, a survey targeted students who were registered in the Azima program but not participating (nonparticipation was defined as a participation rate of 0 percent six weeks into the program) and their caregivers. The survey targeted all three locations where Jusoor operates, and we were able to survey all nonparticipating students (n=196).

Result: Access to devices was identified as the biggest barrier. While 88 percent of refugee families had access to a smartphone, this did not always translate into participation. The (generally) one device many families could afford was needed by the breadwinner to meet their priority of securing work or responding to work requirements.
The Sprint 1 experiments generated evidence that served to validate the program’s assumptions about pedagogy, product, and provision. On the other hand, the survey informed the decision to further explore our assumptions about people and place, and to understand how to alleviate barriers to children’s engagement most effectively. Two experiments were designed in a second sprint to test the assumption that parents are interested in supporting their children’s education and to understand what resources and support would enable them to do so.

**Sprint 2 Experiments**

**Experiment 5: People and Place—Information Campaign**

**Method:** A tailored information campaign was designed to send tips to parents on practical things they could do to help their children learn at home. A series of weekly messages was sent via videos and voice notes over the course of four weeks.

**Result:** The information campaign targeted 916 families (all families of Jusoor’s students), yet only 66 percent (n=602) interacted with its content. Of those, 65 percent found the information helpful and claimed to have adjusted their behavior as a result. However, no noticeable increase in children’s engagement was observed. This may be because the campaign did not tackle the root cause of the problem, or that a longer timeframe was needed to see the full impact of children’s behavioral changes.

**Experiment 6: People and Place—Cash Experiment**

**Method:** In one informal settlement, all 194 families were offered simple, no-strings-attached grants of US$25 to spend however they wished. The following options were explicitly offered to the families prior to the distribution of funds: keep the cash, rent a phone and a data card, rent a phone only, buy a data card only. The impact on engagement was then monitored through a feedback form, attendance tracking, and in-depth interviews.

**Result:** As a result of distributing the cash, engagement with WhatsApp learning increased by 16 percent (to 64%, from an average of 48% before the experiment). Most families (58%) decided to use the cash on a device and/or data. Combining phone rental and data had the greatest impact, as observed through the highest increase in educational engagement. The attendance of students whose families

---

6 The grants were distributed to all families of Jusoor students in the settlement, irrespective of the children’s attendance or performance results.
chose the phone and data option jumped to 78 percent following the experiment, from an average of 50 percent before the experiment.

The combined findings from experiments 4, 5, and 6 supported the assumption that parents prioritize their children’s education when given the support to do so. These findings indicate that solutions such as unconditional cash assistance could be scaled to address the issue of disadvantaged children’s participation in distance learning.

At the end of the sandbox, the team held a final workshop in which they reflected on options for scaling the program and on possible future paths, including an adaptation of Azima to target permanently out-of-school children in the next academic year, when lockdown restrictions were alleviated and Jusoor’s students returned to physical classes.

**CRITICAL REFLECTION**

In this field note, we have presented a case study of an application of the sandbox model—a novel approach to rapidly develop and refine a new education program in an emergency context. This case study illustrates that the sandbox model represents a useful, low-cost, rapid way to conduct research that informs iterations throughout the implementation of a program.

The sandbox approach puts the implementers themselves in the role of researcher, which might present some risks or weaknesses. Practitioners may not have time to collect data with the same rigor as those conducting a more formalized research project, and they may not have the necessary research skills. Having the practitioners evaluate their own projects may create the risk of bias and raise questions about the independence of the research. On the other hand, sandboxes can empower individuals to take a new perspective on their own practices and focus on generating evidence that will be of practical value in their particular context.

Sandboxes result in the production of evidence that is highly contextualized and focused on iterating the particular intervention. As a result, inferences from the data generated may not be generalizable and should be tested carefully before implementing in other contexts. The sandbox approach itself, however, is highly flexible and transferrable, as reflection and iteration are at its core. Building in time to surface critical assumptions and to re-evaluate as new insights emerge during sprints are principles that can be applied in a wide range of settings. By working in sprints, we regularly stopped and reflected together on what we were
learning and used those insights to inform the next steps and experiments. For example, the nature of our experiments changed quite radically between the first and second sprint, as we realized the importance of fully exploring what it would take to alleviate barriers to engagement.

The 6Ps Framework also provides a systematic way of thinking through the range of factors that can intersect in any given ed tech implementation and is a transferable tool for thinking through these issues in different contexts. Within this framework, the choice of particular experiments or data-collection techniques can be adapted to the specific context (Rahman et al. 2021). As such, the sandbox approach has the potential to be applied in a wide range of education contexts, particularly where existing evidence is lacking and urgent action is needed.

ACKNOWLEDGMENTS

We would like to thank the Jusoor team for trying this new approach and allowing us to share the resulting insights in this field note; UNHCR for guidance throughout the sandbox; and the UK Foreign, Commonwealth & Development Office and the Bill & Melinda Gates Foundation for their funding of this work through the EdTech Hub. Most of all, we are grateful for the collaboration of the teachers, parents, and students in the refugee camps in Lebanon, without whom this work would not have been possible.

REFERENCES


