The LEGO Foundation

ACER

UERA

Learning Through Play at School
FINAL RESEARCH REPORT
Ukraine

Prepared by the **Australian Council for Educational Research** *(ACER) and the* **Ukrainian Educational Research Association** *(UERA) for the* **LEGO Foundation** *and the* **Ministry of Education and Science Ukraine**

31 March 2024

Rachel Parker, Amy Berry, Kellie Picker, David Jeffries, Prue Anderson, *and* **Oksana Zabolotna**

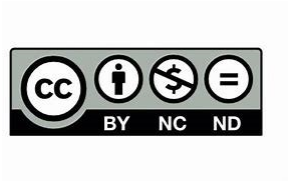
Learning Through Play at School Ukraine: Final Research Report

Rachel Parker, Amy Berry, Kellie Picker, David Jeffries, Prue Anderson, Oksana Zabolotna

The Australian Council for Educational Research Ltd
19 Prospect Hill Road
Camberwell VIC 3124
Phone: +61 3 9277 5555
ABN 19 004 398 145

www.acer.org

ISBN 978-1-74286-738-0



Recommended citation
Parker, R., Berry, A., Picker, K., Jeffries, D., Anderson, P., & Zabolotna, O. (2024). *Learning Through Play at School Ukraine: Final research report.* Australian Council for Educational Research. https://doi.org/10.37517/978-1-74286-738-0

Table of contents

List of Figures

List of tables

Abbreviations

AA	Assessment Administrator
ACER	Australian Council for Educational Research
COVID-19	Coronavirus disease
DIF	Differential item functioning
DMT	Data Management Team
EAP/PV	Expected a-posteriori/plausible value
ISCED	International Standard Classification of Education
LTP	Learning through play
LTP-LSSA	Learning through Play Literacy and Social-emotional Skills Assessment
MNSQ	Weighted infit mean square
MOES	Ministry of Education and Science
NUS	New Ukrainian School
PP	Pedagogy Partner
PV	Plausible values
RQ	Research question
SD	Standard deviation
SEN	Special Education Needs
SES	Socio-economic status
SL	School leader
STF	Student Tracking Form
TCS	Total class sample
TIPP	Teacher Innovative Play Program
TOPP	Training of Pedagogy Partners
TQ	Teacher questionnaire
UERA	Ukrainian Educational Research Association
UNESCO	United Nations Educational, Scientific, and Cultural Organization
ZNO	External Independent Evaluation

Acknowledgements

This report was written by Rachel Parker, Amy Berry, Kellie Picker, David Jeffries, Prue Anderson, and Oksana Zabolotna.

In addition, valuable contributions were made by Luc Le, Van Nguyen, Alla Krasulia, Oksana Roma, Reilly Ross and Patricia Castanheira.

Implementing a research study with substantial fieldwork during a global pandemic and the subsequent invasion presented significant, and at times insurmountable challenges for all staff and participants.

During the preparations for our second round of data collection, Russian armed forces launched a full-scale invasion of Ukraine. This affected every person and participant of the study in myriad ways; family and friends were injured or killed, and homes and businesses were damaged or destroyed. Several staff had to move far from their homes leaving their loved ones, and family members were conscripted to defend the country.

Despite this, our staff, partners, and participants demonstrated unwavering commitment to the study, which speaks volumes about their passion for improving education for the children of Ukraine. We wholeheartedly thank our colleagues, partners, teams, and teachers who contributed to make this study successful, despite the odds.

Most importantly, to the children who participated in this study - we honour you. We were so fortunate to bear witness to your growth and development in your first years of school, your good humour, struggles, and great resilience. We wish you long, happy, healthy, and ***playful*** lives.

Слава Україні!

Executive Summary

Through play we make sense of our world, and in a world of chaos, play is therapeutic. Children in Ukraine live with perpetual uncertainty and disruption. Play enabled them to process their experiences together. One teacher said:

> *For a few minutes during the lesson, the children were constantly sharing their impressions of the noise [of sirens], the way they were running to escape…*

Playful learning environments that promote positive social interactions can provide opportunities to come together, relate to one another and our collective ordeal, and offer some levity, support, and understanding.

Introduction

The Learning Through Play (LTP) at School Research Study Ukraine was a four-year intervention study funded by the LEGO Foundation and implemented by the Australian Council for Educational Research (ACER) and the Ukrainian Educational Research Association (UERA).

The intervention was a two-year professional learning program that blended online, and face-to-face learning called the Teacher Innovative Play Program (TIPP). The TIPP was designed based on documented evidence that reports that teachers need opportunities to experiment and reflect to change practice.[1] The study was guided by three research questions which were revised following the full-scale invasion of Ukraine by Russian armed forces on February 24, 2022. The revised questions were as follows:

1. What are the barriers and enablers that limit and/or support effective implementation of LTP in intervention school classrooms?
2. How do teachers in intervention schools implement LTP and adjust their classroom practices to promote learners' literacy and social and emotional development?
3. How do children's literacy and social emotional skills compare between testing time points including prior to and during the invasion of Ukraine by Russian armed forces?

Methodology

In the original study design, intervention and control groups were selected from five oblasts (regions) in Ukraine: Kyiv, Kharkiv, Zaporizhzhia, Poltava and Dnipro including 20 primary schools, 30 teachers, and 1,465 students[2]. In the revised design, the intervention teacher sample remained the same, but the school and student sample was reduced to seven schools in Kyiv and Poltava and 296 students participating in the final assessment due to the war. Thirty intervention teachers completed 20 interactive online modules where they gained foundational knowledge regarding LTP and worked with experienced Pedagogy Partners (coaches) to enact and reflect on LTP in their classrooms.[3] The study investigated the impact of the TIPP on teachers by comparing the results of data gathered from interviews, classroom videos and planning documentation.

[1] See Clarke & Hollingsworth, 2002.

[2] The sample included 10 intervention and 10 control schools, and 30 intervention teachers. There were initially 2 intervention and 2 control schools in each oblast. All 20 school principals participated in an online questionnaire at endline.

[3] The course was certified by ACER's Accreditation Committee.

The study was initially designed to measure the impact of the intervention on students using the LTP Literacy and Socio-emotional Skills Assessment (LTP-LSSA), designed to align with LEGO Foundation's definitions for play and learning and contemporary research on learning through play The key skills assessed were expressive oral language, listening comprehension, empathy, self-awareness and self-regulation, problem solving and conflict resolution, using playful items scored live using a developmental rubric. The assessment was administered with sample school students at the end of Grade 1 and end of Grade 3[4].

Key Findings

Question 1: What are the barriers and enablers that limit and/or support effective implementation of learning through play in intervention school classrooms?

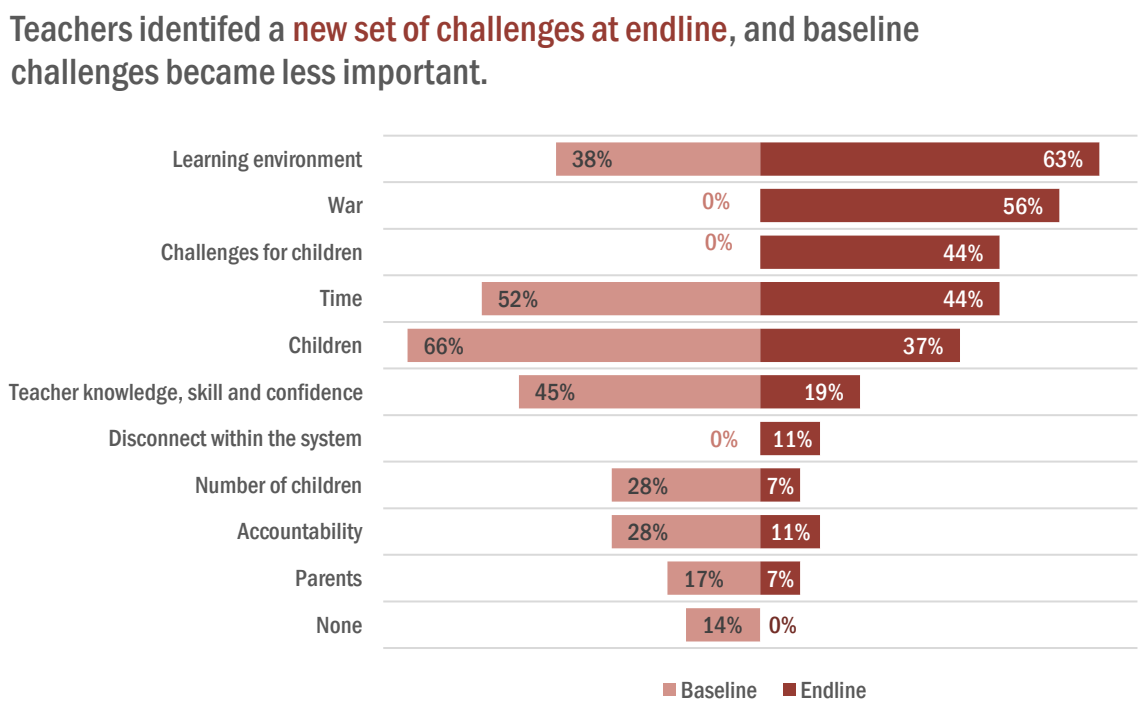
Challenges

As illustrated in Figure 1, challenges changed over time, and new challenges emerged at endline. One important shift was that at baseline most teachers (66%) viewed 'children' as a barrier to learning through play; by not possessing the requisite social-emotional skills including self-regulation they were seen as not ready for LTP. Underlying this is the belief that children must be taught these things *before* they are ready to learn through play, rather than seeing play as a way of supporting children to develop key skills and dispositions for learning.

At endline this had reduced to 37%, and a new barrier emerged - 'challenges ***for*** children' which depicted a fundamental change where teachers described common challenges faced by children when learning through play, with play providing the opportunity for learning and growing. The teachers were now able to take the perspective of the child and appreciate the challenges that children had to overcome when learning through play (e.g., compromising and dealing with setbacks). Instead of seeing deficits in children, they now saw potential and capacity for growth. This new challenge reflects a change in the teacher's thinking as they began to see children as agentic and capable partners in learning. Another new challenge at endline was the disconnect within the system, where teachers questioned the commitment at system-level to extend the active playful approach beyond Grade 4, and the impact of rigid curricula preventing time for deeper LTP. Of particular concern for some teachers was a feeling that children may encounter resistance to their agency and voice as they transition into the higher grades. The invasion of Ukraine presented an insurmountable barrier for some, as did the return to online learning (learning environment).

[4] The original design included three assessment rounds (Grades 1, 2 & 3) which was reduced to two due to the invasion.

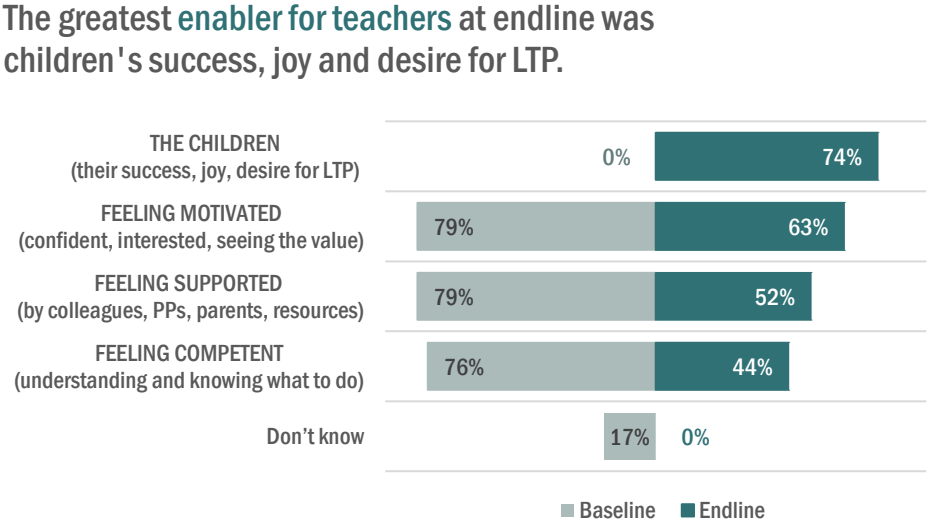
Figure 1: Challenges to implementing LTP at baseline and endline



Enablers

Teachers' perceptions of enablers also changed over time. At baseline, the most common enablers for teachers were feeling motivated, supported, and competent. As shown in Figure 2, at endline a new enabler for teachers emerged; the children themselves - their success, joy, and desire for LTP. This result underlines the value in seeing the impact of change and how it mutually benefits and reinforces teaching and learning. Enablers at baseline such as feeling motivated, supported, and confident were still relevant and important to most teachers at endline, but with more specificity about the source of their motivation and sense of competence coming from the children's success.

Figure 2: Enablers to implementing LTP at baseline and endline



Impact of COVID-19 and the invasion of Ukraine

All 20 school leaders from control and intervention schools, participated in a new online questionnaire designed to unconfound and quantify the impact of COVID-19 and the invasion of Ukraine on participating students and schools. We aimed to gain a better understanding of how students and schools were impacted by COVID-19 and the invasion of Ukraine separately and combined, and whether there were key differences between intervention and control schools.

The results suggested that intervention schools spent more time learning in online mode due to COVID-19 school closures than control schools. Three intervention schools spent more than 31 weeks learning in online mode compared with one control school. Regarding time spent in online mode after the invasion in February 2022, the duration was roughly equal comparing control and intervention groups. More control school leaders reported damage to their schools resulting from the invasion than intervention schools and an equal number of control and intervention schools reported damage impacting internet access. In summary, there was no distinct pattern of responses that suggested one group was much more impacted than the other.

Question 2: How do teachers in intervention schools implement LTP and adjust their classroom practices to promote learners' literacy and social and emotional development?

In responding to this question, the research team recognised it was beneficial to understand not only what **teachers do,** but also how they **think and feel** about LTP, acknowledging the relationship between these dimensions and sustainable change. Researchers have found that the way teachers implement new methods relates to how those methods align with their beliefs (Donnell & Gettinger, 2015). If a teacher does not value play or believe LTP is a way to foster holistic skills, their capacity for change and impact is limited.

Key changes to participating teachers' feelings (attitudes, values) about learning through play after the intervention were:

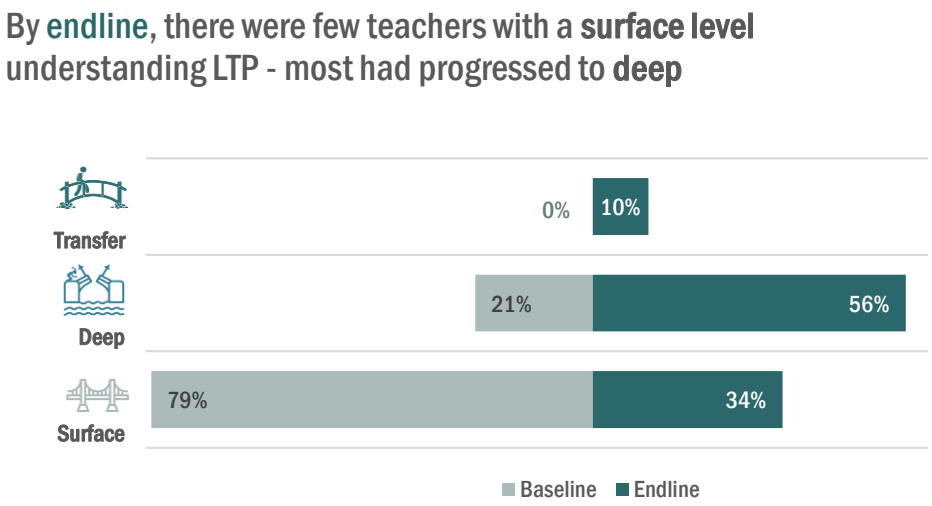
1. 48% of teachers reported increased **motivation and confidence about implementing LTP**
2. 30% of teachers reported a decrease in **fear and anxiety** about implementing LTP

Key changes to teachers' thoughts (beliefs) about learning through play after the intervention were:

1. **About LTP as a pedagogy:** Most teachers (55%) reported seeing LTP differently, as an effective way to foster a range of skills
2. **About children:** Many teachers (33%) saw children as more capable of solving problems and making decisions about their learning than they previously thought
3. **About teaching:** Many teachers (30%) developed a new vision for themselves as partners in and facilitators of learning

To understand the change over time in teachers' depth of understanding of LTP, combining practice, attitudes, and values, we applied the lens of 'surface, deep, and transfer'. This notion views learning as a process of acquiring foundational knowledge, deepening though practice, and at the highest level, transferring new knowledge to new contexts. As shown in Figure 3, at baseline, most teachers reported demonstrating thinking and behaviours at the surface level (79%), and a small proportion at deep (21%). At endline, the percentage of teachers at surface had reduced to 34% and the percentage at deep had increased to 56% with a small number operating at the highest level of transfer.

Figure 3: Depth of teachers' understanding of LTP, baseline and endline



A full description of the levels can be found in section 4.2. In summary they are:

- **Transfer:** I have a whole new way of thinking about teaching and learning
- **Deep:** I see the bigger picture and understand the value of learning through play
- **Surface:** I might see play as a threat to teaching and learning.

Most teachers started the program with a surface level understanding of LTP, but others began at higher levels. Given the levels contained a range of practices and dispositions, it was possible for a teacher to stay at the same level but still grow.

Changes to teachers' practice

Classroom videos revealed many teachers (47%) demonstrated some change to their practice; some teachers demonstrated limited change, and a smaller proportion significant change. Changes to practice observed in videos and documented in reflections included:

- Increase in group and pair work
- Integration of LTP into a range of curricula areas
- Increased skills in reflective practice
- Support for agentic learning

The results showed an increased number of examples of co-constructed approaches to learning through play from baseline at 2 to endline at 14. Co-constructed approaches included children facilitating learning (being the teacher) children making decisions about what they want to learn, children and teachers collaborating to extend learning and teachers taking on student suggestions.

Successful professional learning

Learning through play at school Ukraine contributes to our understanding of how LTP practice evolves by acknowledging the **cognitive, behavioural, and affective** dimensions of change. Professional learning aims to affect teacher practice; however, this aim is rarely met. The findings of this study align with evidence about what works to affect change in practice, that is, providing time and opportunities to apply new learning in practice, including opportunities for reflection, support from a coach, and creating social networks. These features underpinned the success of this study in changing teacher practice.

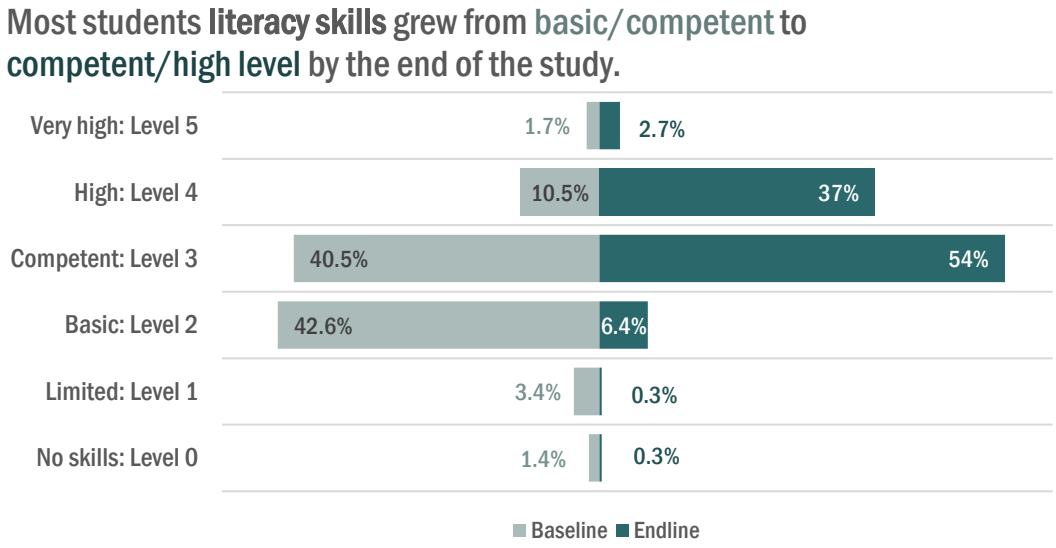
Question 3: When comparing the control and intervention groups, is there a difference between children's literacy, and social and emotional skills including collaboration, at the start and end of the intervention?

Students' expressive and receptive oral language and social-emotional skills were assessed using the Learning through Play Literacy and Social-emotional Skills Assessment, a short playful assessment of these skills designed especially for this study to align with LEGO Foundation's definitions of play and learning and administered 1-1 and live scored using a rubric.

Literacy (oral language) skills at endline

Results at endline revealed substantial growth from baseline. At endline, the distribution was skewed to the higher levels, as expected for an assessment of this nature. The assessment was found to be a well targeted and satisfactory measure of literacy and social-emotional skills for children at the end of Grades 1 and 3, when the assessments were administered. Figure 4 shows the results at baseline and endline for each level.

Figure 4: Literacy skills at baseline and endline, all students



Students with **high or very high literacy** skills could answer all questions about a text read to them, imaginatively generate descriptions of characters, imply ideas beyond the images, and give 3 synonyms for a familiar adjective.

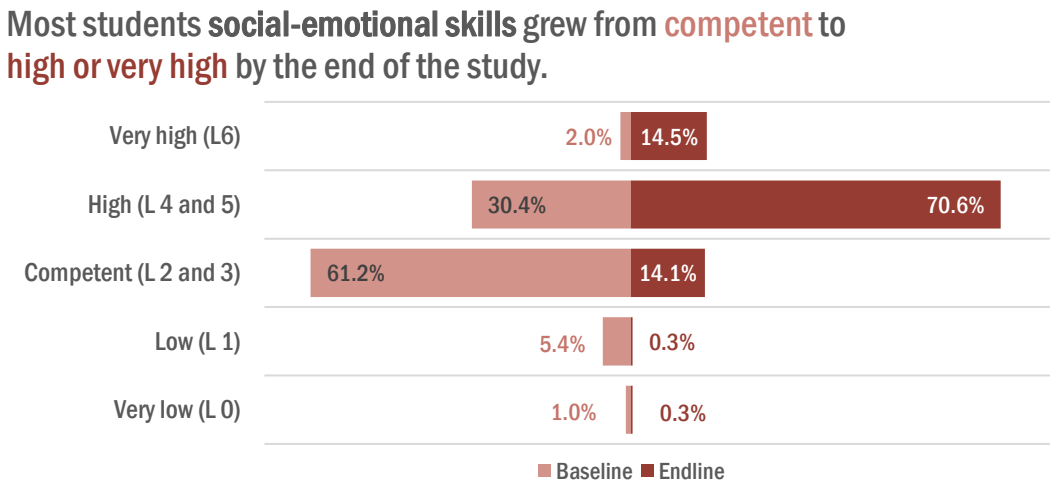
Students with **competent literacy skills** could answer most questions about a text read to them, link features when describing characters and give 1-2 synonyms for a familiar adjective.

Students with **basic literacy skills** could communicate effectively and understand short texts with teacher support.

Socio-emotional skills at endline

As with literacy, students' endline social-emotional skills grew substantially from baseline. Results were skewed toward the upper levels at the end of the study. Figure 5 shows the results at baseline and endline for each level.

Figure 5: Social-emotional skills at baseline and endline, all students



Students with **high or very high social-emotional skills** could propose collaborative solutions to problems, understand ineffective conflict resolution strategies, and show insight and empathy when evaluating a character's behaviour.

Competent social-emotional skills were described as evaluating a character's response to a conflict scenario from one perspective only or proposing a resolution to conflict via an external solution.

Students with **limited social and emotional skills** demonstrated self-awareness, but no self-regulation strategies or actions to demonstrate empathy or solve conflicts.

Difference in performance between intervention and control schools at baseline

Control school students performed higher than intervention schools in both literacy and social-emotional skills, but intervention students grew more between baseline and endline. The gap between groups was starting to close by endline, as illustrated in Figure 6 and Figure 7. Intervention school children were assessed at baseline with an average score of 47.71. At endline their score was 58 which is a difference of 10.29 points growth.

Figure 6: Intervention and control literacy performance and growth

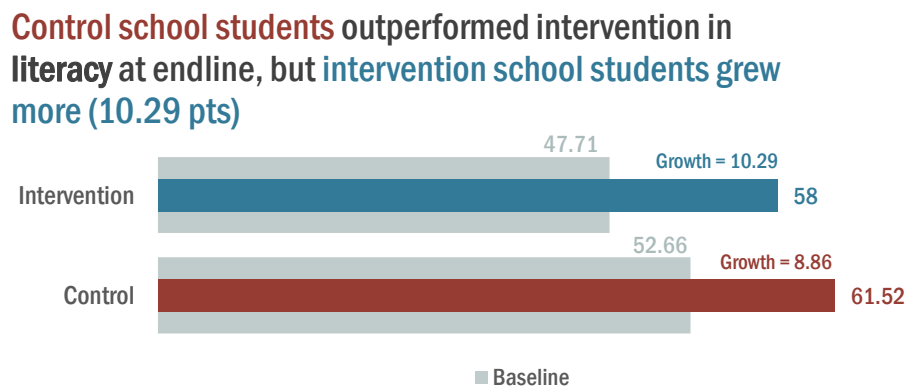
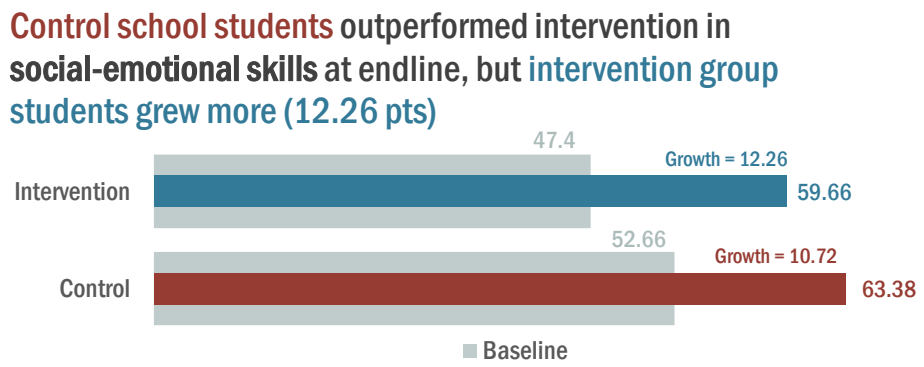


Figure 7: Intervention and control social-emotional skills performance and growth



The difference in performance between the intervention and control groups must be considered along with key differences between the groups. The percentage of low socioeconomic status (SES) students in the intervention group was much higher (44%) than in the control group (25%). This is because SES was not known and therefore not a school selection factor. It was not possible to match control and intervention schools on SES composition. Student-level SES was obtained via the parent questionnaire. In addition, the students in control schools demonstrated higher performance than students in intervention schools in both literacy and social-emotional skills at baseline.

These differences impacted the study regarding the expected growth of the two groups from baseline to end-line. As predicted in the baseline report, the participating students from the intervention schools demonstrated higher levels of growth in literacy and social-emotional skills compared to participating control group students whose teachers were not involved in the TIPP and thus we saw the effect of 'closing the gap'.

Intervention teachers provided detailed descriptions of student growth in literacy and social-emotional skills which helped explain the results from the student assessment. These included descriptions of students' self-management, social awareness, relationship skills, creative thinking, collaboration, communication, and problem solving.

The student assessment results are unique and valuable in that they demonstrate that:

1. Playful learning designed to foster positive learning experiences and relationships might have been a protective factor in supporting children to continue developing holistic skills.
2. Despite the compounded impact of COVID-19 and the full-scale invasion of Ukraine, almost all students demonstrated growth and the magnitude of this growth was measured as very large and large (Intervention EE= 1.3/Control EE= 0.84).
3. Playful assessments that are live scored using developmental rubrics can reliably measure literacy (oral language) and social-emotional skills in Ukraine.

Discussion

This study addresses gaps in the LTP evidence base in the following key areas.

Good practice in LTP: This study has generated practical guidance for teachers to identify where they are positioned on a continuum of depth of understanding of LTP from surface to deep. In doing so, we can better understand what good practice looks like and how it develops over time.

LTP in key transitions: Many teachers made the fundamental shift from viewing children as a problem to LTP as the solution. Initially teachers viewed children as unprepared for play and lacking key skills. At endline, they viewed LTP as the vehicle for developing these skills. This shift can be characterised as moving from a focus on school readiness to making schools ready for children.

LTP requires playful measures: We developed a reliable – and playful – instrument to measure children's literacy and social-emotional skills appropriate to the target group with sufficient room for growth. The assessment results were supported by other valuable sources of data enabling us to triangulate findings and understanding the 'how' and 'why' in addition to 'what'. This underlined the importance of using diverse methods to answer complex questions about LTP.

Evidence relevant to the New Ukrainian School (NUS) reform: The Ministry of Education and Science of Ukraine may find results presented here useful and relevant regarding the shift in focus toward developing competencies over learning content, particularly social-emotional skills. Other areas of interest include the key features of impactful professional learning, extending the NUS to the middle years, and learning and development during extreme events.

1. Introduction

The Learning through Play at School Research Study aimed to test in practice what was theorised from research evidence in the scoping study Learning through play at school (Parker & Thomsen, 2019) to broaden the evidence base to include new countries. Specifically, this study aimed to develop specific guidance and materials on integrating learning through play pedagogies in early primary classrooms in the form of a professional learning program for teachers, and to pilot this program in a specific geographic context. Key findings guide and inform the future implementation of the program in new contexts, as well as providing detail about sustaining changes in Ukraine. This is particularly important against the backdrop of COVID-19 and the full-scale invasion of Ukraine by Russian armed forces, and the ongoing implementation and extension of the New Ukrainian School (NUS) reform.

The study is underpinned by the hypothesis that teachers can be supported to develop the knowledge, skills, and mindsets to enable them to integrate LTP pedagogies in the early primary classroom. It is further hypothesised that these changes to teacher practice will have a positive impact on student learning.

1.1. Research questions

The study is guided by three research questions, revised on 19 December 2022, as follows. Note that question 1 and 2 remain largely unchanged from the original study design.

1. What are the barriers and enablers that limit and/or support effective implementation of learning through play in intervention school classrooms?
   - What barriers (e.g. teacher attitudes and values, training) limit effective implementation of LTP in intervention classrooms?
   - What enablers support effective LTP integration in intervention classrooms?
2. How do teachers in intervention schools implement LTP and adjust their classroom practices to promote learners' literacy and social and emotional development?
   - What practices do teachers demonstrate at the beginning of the study?
   - What practices do teachers demonstrate during the intervention?
3. How do children's literacy and social emotional skills compare between testing time points including prior to and during the invasion of Ukraine by Russian armed forces?
   - When comparing the control and intervention groups, is there a difference between children's literacy, and social and emotional skills including collaboration, at the start of intervention?
   - When comparing children learning in different locations in Ukraine, is there a difference between children's literacy and SE skills?

The study seeks to address the evidence gap on effective implementation of LTP in early primary school classrooms. This includes understanding the necessary preconditions for successful integration, the barriers and enablers to integration, and the adaptations to curricula and teaching required to integrate playful pedagogies in the classroom.

This final report summarises the results of the four-semester professional learning program and data collection that has occurred from the start to end of the study, namely the baseline and endline teacher interviews, student assessment, supporting evidence from pedagogy partner and teacher reflections, and video recordings of classroom practice.

2. Methodology

2.1. Overall study design

The Learning through Play at School Research Study was designed to support a targeted number of schools and teachers in Ukraine to integrate LTP pedagogies in their classrooms. Guided by the research literature, an online and face-to-face blended professional learning program for teachers was developed and was implemented over the course of four semesters. The professional learning program, the Teacher Innovative Play Program (TIPP), included both group and individual learning experiences. The program was monitored and evaluated using quantitative and qualitative methods including interviews, questionnaires, reflections, classroom observations, and a student assessment.

Key features of the methodology were as follows:

- A longitudinal study that follows the same cohort of teachers and students from Grade 1 to 3 enabling researchers to monitor teachers' and students' performance and growth over time.
- Inclusion of intervention and control schools to enable comparison of performance and growth of both groups of students over time.
- An evidence-informed professional learning program.
- All components of the study align with the LEGO Foundation's definitions for play and learning (Zosh et al., 2017) and integrated playful pedagogies (Parker & Thomsen, 2019).

The study was designed to evaluate the impact of the TIPP intervention on teachers and their students. Selection of control and intervention school sites was influenced by the location of the selected Pedagogy Partners, preventing the use of random control measures. The intervention was a four-semester professional learning program comprising online learning modules, workshops, and face-to-face reflection sessions between teachers and their designated Pedagogy Partners (coaches). To measure the changes to children's literacy (oral expressive and receptive skills) and social-emotional skills between two time points, students in control and intervention schools completed a baseline assessment at the end of Grade 1. Students completed another assessment at the end of the intervention at end of Grade 3. Results of the assessment are presented at Chapter 4. Other data collection methods used to answer the research questions included teacher and school leader interviews, an online school leader questionnaire, a teacher questionnaire, teacher planning and reflection documents, Pedagogy Partner interviews and video recordings of classroom practice. A full description of the methodology can be found at Annex 1.

3. The Intervention: The Teacher Innovative Play Program (TIPP)

3.1. Overview

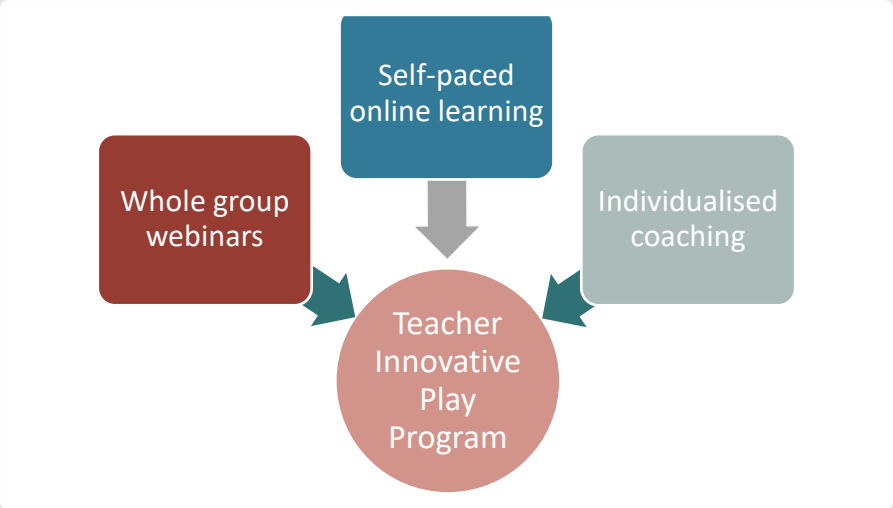
The intervention, called the Teacher Innovative Play Program (TIPP), was a four-semester blended learning program intended to be delivered over two calendar years[5]. It was designed to remain true to the spirit of learning through play, aiming to extend the LEGO Foundation's definition of learning through play and apply it to adult learners (teachers), as well as the children they teach. The design incorporates evidence-based practices that are recommended in the teacher professional learning research literature, remaining adaptable and responsive to context and individual needs, and deliverable within the constraints of the current study. In the design, teachers are supported to establish a strong foundation of knowledge about learning through play, as well as opportunities to connect to the real world of the classroom and deepen their understanding of the concept through

[5] The TIPP final semester was due for implementation from Jan 2022 but delayed one year due to the invasion of Ukraine.

an ongoing, iterative process of professional experimentation and reflection. The practical application of the ideas presented in the online learning modules and support provided by the Pedagogy Partners were designed to help the teacher improve their use of playful pedagogies and empower them to continue the improvement journey beyond the current learning program.

This blended learning program integrated whole group webinars and celebrations of learning, online self-paced learning using the Moodle learning management system, and individual in-person coaching in the Pedagogy Partner sessions (see Figure 8). This approach enabled learners to have some degree of control over the pace they moved through online materials and allowed them to return to modules as often as they like. The coaching program was individualized and embedded in the teacher's classroom practice. Teachers were supported and provided templates to plan a lesson integrating learning through play, implement the lesson, and reflect on the experience with their Pedagogy Partner. The Pedagogy Partners video recorded each of these lessons and shared them with the teacher to support reflection and discussion on the lesson. The Pedagogy Partners provided support to the teachers in both the online setting (e.g., through discussion boards) and in their individual coaching sessions.

Figure 8: The Teacher Innovative Play Program (TIPP)



3.2. TIPP Semester 3 and 4

With the end of the pandemic and education in Ukraine settling back into offline learning, it was thought the final two semester of the TIPP would provide teachers with the opportunity to take a deep dive into the modules and have uninterrupted opportunities to translate the learning to their classroom practice. However, this was not the case, as on 24 February 2022, Russia invaded Ukraine. As expected, this impacted the delivery of the Semester 4 TIPP, which was put on hold. The Semester 4 TIPP modules were initially designed as an action research project, giving teachers agency over their learning and the ability to follow one of three LTP areas of interest – student agency, social-emotional learning, and minds-on learning. Overall, the TIPP was paused for 12-months but before restarting the research team consulted with the LEGO Foundation (Ukraine), UERA and the Pedagogy Partners to gain an understanding of what would best support the teachers current teaching and learning needs. These conversations resulted in the rewriting of the Semester 4 TIPP modules.

An overview of the topics covered in the Semester 3 and Semester 4 of the TIPP is found below.

TIPP Semester 3 modules

The aim of Semester 3 was to respond to the teachers' feedback from the first two semesters of the TIPP by consolidating and extending their understanding of key concepts and supporting them to embed learning through play into their everyday classroom practices. The first module for the semester welcomed and thanked the teachers for their dedication to the project. It summarised ideas previously covered, including the importance of reflection, as well as sharing observations from the videoed lessons and setting LTP goals for the semester. The remaining four modules for Semester 3 focused on pedagogical practices that could be used to support a range of LTP experiences. Each of these modules included playful, engaging teaching strategies and an example lesson plan that could be implemented by the teachers.

- Module 13: Welcome and moving forward
- Module 14: Dialogical practices
- Module 15: Student collaboration and co-operation
- Module 16: Open-ended approaches
- Module 17: Problem-based learning

TIPP Semester 4 modules

The Semester 4 TIPP content was condensed, and explicit connections were made to previous modules. New modules included teaching strategies for use in both the offline and online learning environments. Two modules were developed to help teachers refine their learning and practice in key areas of LTP identified as critically important in the current context – social-emotional learning and student agency. Both modules incorporated minds-on learning experiences.

- Module 18: Introduction to Semester 4
- Module 19: Supporting social-emotional learning
- Module 20: Supporting student agency

Overall TIPP feedback

During the endline interviews, teachers were asked to share their experience participating in the TIPP, which was comprised of the online learning modules and the Pedagogy Partner coaching program. In total 28 Interviews were conducted by UERA. Feedback about the modules and the coaching program are summarised below.

Feedback on the TIPP modules

Feedback about the TIPP modules was generally positive with teachers providing details about how they used the content to support changes in their classroom practice. To begin with, a small number of teachers indicated that they were familiar with the concepts being taught in the TIPP because it aligned with and reinforced what they had learned during NUS training. Other teachers familiar with learning through play suggested that participating in the TIPP connected with their prior knowledge and provided more up-to-date information about its implementation in the classroom. Teachers described how they tried ideas from the modules, others explained how they combined the content from the modules with their knowledge of the children, as well as their own ideas to adapt their LTP experiences, with these adaptations sometimes taking place during the lesson. As an example of the positive impact of engaging with the TIPP content, one teacher explained how she rethought the role of group learning while engaging with the modules. Another teacher was sceptical about what she read in the discussion forums about her colleagues learning through play experiences until she tried the experiences herself.

Favourite TIPP strategies

When asked about their favourite LTP strategies teachers referred to many pedagogies and strategies outlined in the TIPP modules. Strategies to foster social skills such as group, pair work and collaboration were the most popular among teachers. Teachers noted that children enjoyed working together and identified social development as an important life skill. They also used other strategies, such as 'turn and talk' to encourage children to interact and learn from each other.

Considering the impact of the invasion, it is not surprising that the teachers were also interested in developing children's holistic emotional intelligence. One teacher referred specifically to the content from Module 19, as providing strategies to support emotional expression and children developing the ability to respond to emotions in appropriate ways. Another teacher explained that the children often wanted to talk in class about what was happening around them with their family and friends and they used these strategies to encourage emotional expression and to help the children feel emotionally at ease.

Other TIPP strategies favoured by teachers were active, experiential, and problem-based learning, as well as dialogical practices and thinking skills.

Challenges and constructive feedback

As with the first year of the TIPP, there were a few teachers who continued to have technical issues with logging in, internet access, and the loading of TIPP modules. Some teachers expressed initial fears and concerns about the unknowns of the TIPP, such as understanding the content, linking the content to classroom practices and being unfamiliar with the use of reflective practices. A small number of teachers indicated they did not encounter problems with the TIPP and an even smaller number of teachers explained that teaching in an online learning environment reduced the capacity for LTP due to limited lesson time and difficulties in implementing playful pedagogies.

The key constructive feedback from the TIPP teachers was to expand the TIPP modules to increase the number of practical examples showcasing the various learning through play strategies and pedagogies, and their implementation in the classroom. It was suggested this could be done using:

- Downloadable resources, e.g., games
- Excerpts of LTP lessons that could be tried in the classroom
- Observation videos showing how other teachers enact learning through play

It was advocated that the TIPP teacher lesson plans and/or video recorded lessons could be shared on the discussion forums so the teachers could observe the different ways their colleagues implemented learning through play in their classrooms.

Feedback on the Pedagogy Partner coaching program

The TIPP teachers had high praise for the Pedagogy Partners coaching program, particularly the professionalism and support the Pedagogy Partners provided as the teachers navigated the TIPP modules and implemented learning through play in their classrooms. The teachers acknowledged how the Pedagogy Partners were constantly in touch (via video calls, Viber, classroom visits and other means) to provide them with emotional support, motivation, reassurance and scaffolding to help them connect with the content and apply that within their lessons. The teachers learned from the practical advice offered by the Pedagogy Partners, which included useful materials and resources, as well as appropriate teaching topics to support learning through play. The teachers appreciated their encouragement to experiment and take risks, and the time they took to help them reflect. The feedback provided by the Pedagogy Partners inspired and supported the teachers to consolidate and extend their implementation of learning through play. The TIPP teachers did not

face any challenges working with the Pedagogy Partners and highlighted that they were kind, positive, sincere, bright and a pleasure to work with.

4. Summary of findings at endline

Based on the data collected at the end of the study, the following chapter is organised by the three research questions that guide and underpin this study. They are:

1. What are the barriers and enablers that limit and/or support effective implementation of learning through play in intervention school classrooms?
   1.1. What barriers (e.g. teacher attitudes and values, training) limit effective implementation of LTP in intervention classrooms?
   1.2. What enablers support effective LTP integration in intervention classrooms?

2. How do teachers in intervention schools implement LTP and adjust their classroom practices to promote learners' literacy and social and emotional development?
   2.1. What practices do teachers demonstrate at the beginning of the study?
   2.2. What practices do teachers demonstrate during the intervention?

3. How do children's literacy and social emotional skills compare between testing time points including prior to and during the invasion of Ukraine by Russian armed forces?
   3.1. When comparing the control and intervention groups, is there a difference between children's literacy, and social and emotional skills including collaboration, at the start of intervention?
   3.2. When comparing children learning in different locations in Ukraine, is there a difference between children's literacy and SE skills?

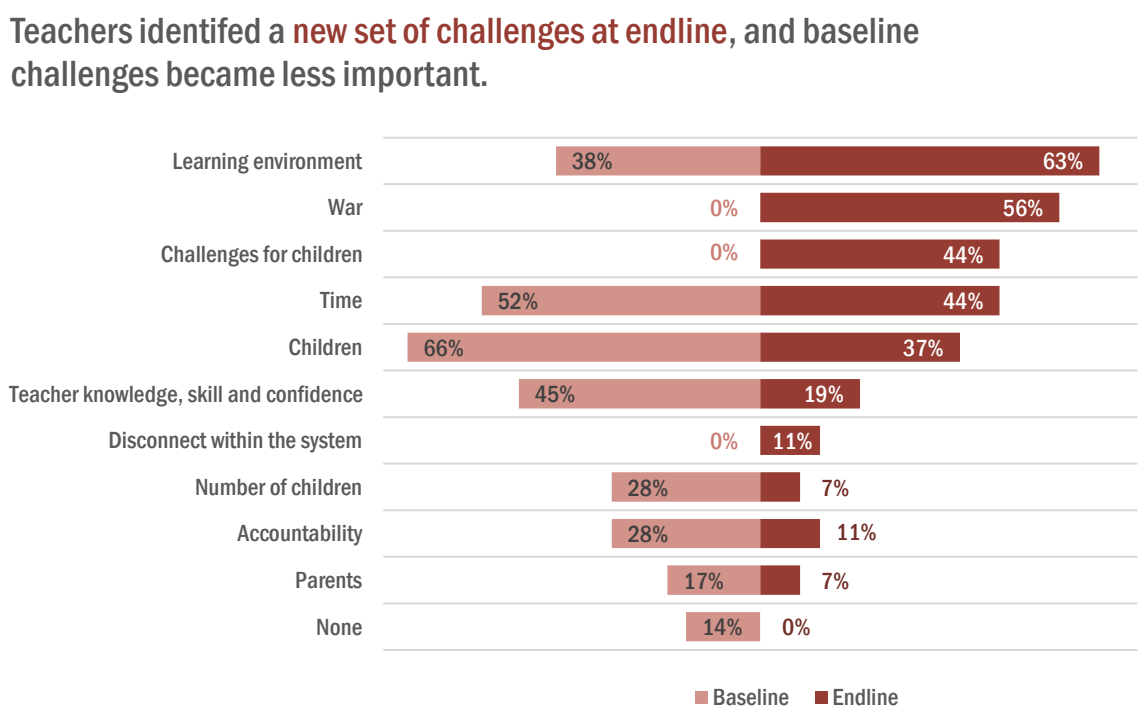
4.1. Research Question 1: Barriers and enablers

In this chapter, we answer the question *What are the barriers and enablers that limit and/or support effective implementation of learning through play in intervention school classrooms?* Understanding the factors that influence teachers' motivation and capacity to implement learning through play is crucial to decision-making about how best to support and sustain meaningful changes to practice. To investigate the barriers and enablers to implementing LTP, we analysed the teacher interviews at baseline and endline.

Barriers to implementing LTP

At both baseline and endline, teachers were asked to describe the challenges they faced when implementing learning through play. This allowed us to investigate not only the barriers to implementing learning through play, but also how those barriers might change over time. Figure 9 shows the results of that investigation.

Figure 9: Teacher identified challenges to implementing LTP



Despite reporting generally positive attitudes towards learning through play, implementing it at school remained challenging for many teachers. The challenges identified at baseline reflect similar findings reported in the literature. The challenges of time (having the time to prepare for LTP and the time to implement it) and accountability (feeling the pressure to perform and achieve expected outcomes) are frequently reported as concerns for teachers, particularly when they have not yet made the connection between play and learning (Bubikova-Moan et al., 2019).

Similarly, teacher concerns about their own readiness to implement LTP is also a familiar barrier reported in the literature (Parker & Thomsen, 2019). This includes a lack of confidence in their understanding of LTP and skills in implementing play-based pedagogies (Bubikova-Moan et al., 2019), and concerns about implementing LTP with large numbers of students (Gray & Ryan, 2016).

The biggest perceived barrier at baseline was the children, something that has been reported in the literature but to a lesser degree than seen here. The concern for these teachers stemmed from a belief that children did not have the necessary skills and dispositions (e.g., emotional regulation) to successfully learn through play at school. Underpinning this is the belief that children must be taught these things *before* they are ready to learn through play, rather than seeing play as a way of supporting children to develop key skills and dispositions for learning.

Looking at the perceived challenges at endline, a different story emerged. As teachers started to change their thinking about children and about LTP, they reflected on the challenges that children faced when learning through play at school (e.g., developing social skills, emotional regulation, confidence, etc.) and could now see how play had helped children to develop these skills. This was an important shift – from seeing children as a challenge that prevented LTP to seeing LTP as the answer to support children to take on challenges and develop important skills. As their knowledge, skills, experience, and confidence grew, teachers moved their focus from how ready the child is for the demands and requirements of school to how the child experiences LTP and learns in this

environment. In short, as they empathised with learners, teachers began to show concern ***for*** children, rather than ***about*** children.

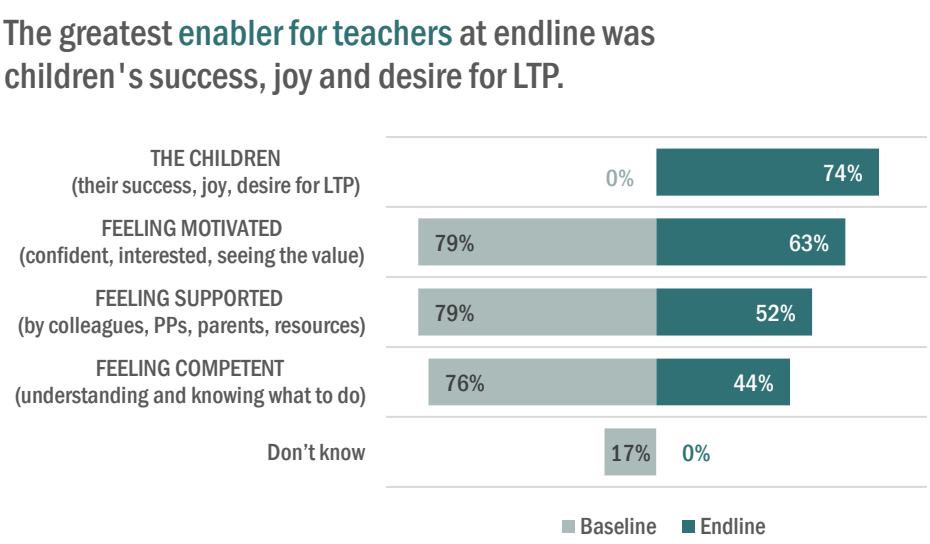
Teachers also began to express concerns about the emerging disconnect within the system. They questioned the commitment to LTP beyond Grade 4 and expressed frustrations that the curriculum did not allow them the time needed to support deeper learning through play. These barriers align with those described in the literature regarding rigid content heavy curricula curtailing time for play (Parker, Thomsen & Berry, 2022).

Unsurprisingly, the war and the subsequent shift to remote learning presented a significant new challenge for intervention teachers. The combined challenge of having to apply newly developed skills to a new context (remote learning) and doing so during a period of extreme stress was understandably difficult for teachers. While some teachers were able to rise to this challenge and continued to develop their LTP practices, others felt they had to pause their journey as they could not find a way through.

Enablers that support teachers to implement LTP

Along with investigating the barriers to implementing learning through play, the study was designed to gather data on what teachers believed supported them to use playful pedagogies in their classrooms. As was the case with barriers, we analysed the baseline and endline teacher interviews to understand how those enablers might change over time. Figure 10 shows the results of that investigation.

Figure 10: Teacher identified enablers that support LTP implementation



As with the barriers, the teachers' perceptions of enablers changed with time. At the start of the TIPP, key enablers were the teacher's feelings of motivation and competence in relation to implementing LTP, as well as feeling supported to implement LTP. While these things remained important enablers at endline, the greatest enabler became the children themselves. Watching the children grow and develop, not only in relation to academic subjects, but also in confidence and competence as learners was motivating for many of the teachers interviewed. One teacher explained what supported her to implement LTP as follows:

> *This is the success of my children. When you see their development and see how they perform tasks, have the opportunity to compare how it was at the beginning and how it is now, it really motivates you to implement learning through play in the lesson.*

Many of the enablers and barriers identified by the teachers in this study echo those previously reported, suggesting a level of generalisability in the evidence from high-income countries that dominate the research literature and lower and middle-income countries such as Ukraine.
Two areas of divergence are worth noting, both relating to the teachers' views of children. First, concerns about children's readiness for learning through play at baseline appeared to be a more significant barrier for these teachers than what has previously been reported in the literature (see Bubikova-Moan et al., 2019). This may relate to the existing cultural traditions and beliefs about teaching and learning in Ukraine. At the core of this perceived challenge is the notion of *school readiness* rather than the children's *readiness for play*. The concern with school readiness and issues with "schoolification" have been identified as particularly relevant in countries where traditional, teacher-centred pedagogies are dominant (OECD, 2006). The second area of divergence is the emergence of children as a key enabler at endline, identifying a new and important idea about what supports teachers to implement LTP at school. While teachers may be willing to try out new approaches, if they do not see any evidence of a positive effect, they may decide to give up and go back to more familiar approaches. The findings reported here show that the teachers were receiving valuable positive feedback from their children that motivated them to keep going. If teachers are to sustain their motivation to implement learning through play over time (as an integral part of the NUS reform) they will need to experience signs of success and see the positive benefits for children and themselves. While the importance of context is well accepted, it is worth challenging our thinking about context. The findings from this study highlight the importance of considering time as a context and how people might change their perception of barriers and enablers over time as their thinking changes. This is especially important when thinking about changing teacher practice within a long-term, system-level reform such as the NUS.

Impact of COVID-19 and the invasion of Ukraine

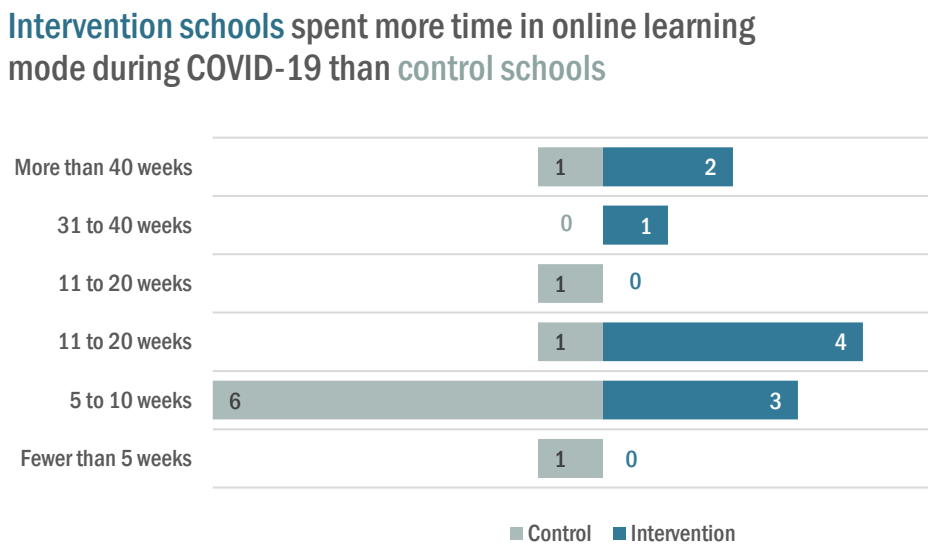
From the beginning of the study, the operating environment was complex, precarious, and ever changing, due to the pandemic which began six months after the study commenced, and the full-scale invasion of Ukraine two years later. To try to understand, unconfound and quantify the impact of COVID-19 and the invasion of Ukraine on learning, we added a new research method at endline – an online school leader interview. Questions included teacher departure due to COVID-19 or the invasion, support required during COVID-19, challenges during COVID-19, weeks learning in online mode, impact of the invasion on school, community, infrastructure and learning, and support required during the war. Data from all school leaders from the full original sample (two intervention and two control schools from each of the five regions, 20 in total) were received.

The results showed that most school leaders had more than four years in leadership at the school. Almost half the schools had at least one Grade 1-3 teacher leave during the study, with teacher retention slightly worse in control schools. Most schools had either support materials, training or both provided to their Grade 3 teachers to support student wellbeing, with intervention schools more resourced. Remaining results are organised by themes as follows.

Impact of COVID-19

Most school leaders (11) reported operating online during COVID for 5-20 weeks, with 7 intervention schools reporting spending between 11-40 weeks in online learning mode prior to Feb 24, 2022. Figure 11 shows the number of intervention and control schools learning online during the pandemic.

Figure 11: Sampled schools weeks of online learning during COVID-19



The most challenging factors for online learning due to COVID-19 according to school principals were student access to internet (18), student access to technology (18), and student/family familiarity with technology (19). According to principals, most teachers (18) required either minimal or some support for online learning due to COVID-19.

Impact of the full-scale invasion of Ukraine

Damage to school infrastructure

Four control school leaders reported damage to school infrastructure, compared with three intervention school leaders (n=20). An equal number of intervention (2) and control school leaders (2) reported damage to surrounding areas caused by the war (n=18). Further, an equal number of control (4) and intervention school leaders (4) reported damage to homes, with the highest proportion in Kharkiv (3) and Zaporizhzhia (3).

Extent of damage disrupting learning

Most schools reported none or minimal damage to heating (17), transportation (18), and water (19), with damage slightly worse in control schools. More schools experienced disruptions to internet with 8 schools somewhat or very affected (n=20).

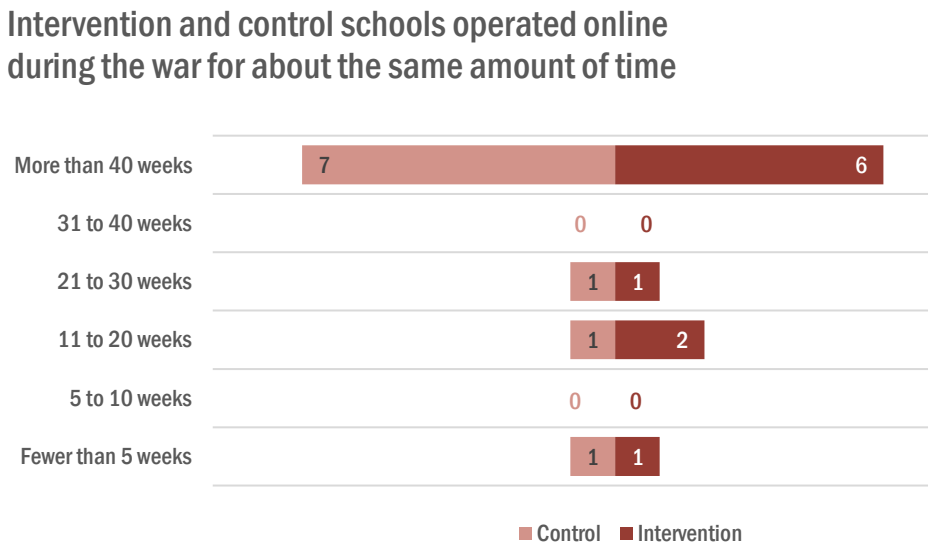
Challenging factors for online learning due to the war

Most schools reported being minimally or somewhat challenged regarding the lack of student access to the internet (17) or technology (17). The lack of access impacted intervention schools slightly more than control schools. Results were more varied with respect to teacher access to technology, with control schools slightly more affected by lack of access.

Effect of war on learning delivery mode

Most sampled schools (15) ceased operating entirely due to the war for fewer than 5 weeks, with one control school in Poltava closing for more than 40 weeks. Some schools shifted to online learning due to the war for up to 30 weeks, with the majority (13) in online learning mode for more than 40 weeks. According to school leaders, most teachers required either minimal or some support for online learning due to the war. Figure 12 shows the amount of weeks control and intervention schools learned online due to the war.

Figure 12: Sampled schools weeks of online learning during the war



4.2. Research Question 2: Implementing learning through play

In this chapter we answer the research question: *How do teachers in intervention schools implement LTP and adjust their classroom practices to promote learners' literacy and social and emotional development?* The question contained two sub-questions:

- *What practices do teachers demonstrate at the beginning of the study?*
- *What practices do teachers demonstrate during the intervention?*

This section focuses on changes in participating teachers and brings together the perspectives of the teachers, the Pedagogy Partners, and the researchers. Understanding changes to how teachers implement learning through play requires looking beyond what they are doing when enacting playful pedagogies. Implementing learning through play, as seen above in response to RQ1, is influenced by how teachers **think** about the concept, what they **know** about it, and how they **feel** about it. When it comes to educational reform, the extent to which teachers feel it is appropriate, reasonable, and likely to be effective is a contributing factor in their decisions about implementing proposed changes (Donnell & Gettinger, 2015). To answer research question 2, and understand how teachers may have changed over the course of the professional learning program, we analysed a range of data from the following sources:

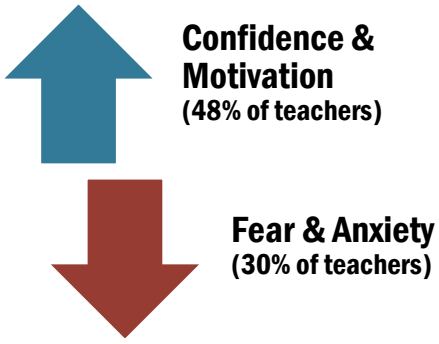
- baseline/endline teacher interviews
- endline teacher reflections (facilitated)
- endline Pedagogy Partner reflections (facilitated)
- classroom videos

While the lesson videos offered insights into what learning through play looked like in intervention classrooms, we had to suspend the collection of lesson videos at the start of 2022 due to the invasion. Accordingly, no videos were collected at the end of the study when teachers had completed the intervention. Using a combination of classroom videos and baseline/endline interviews enabled us to gain an understanding of changes to teachers across the full period of the TIPP. Below, we will explore changes to the teachers' feelings about learning through play, changes in their understanding of learning through play, and changes to their practice.

Changes to teacher feelings about learning through play

When asked to reflect on changes to their feelings about implementing learning through play at the end of the study, just under half of the teachers interviewed described increased feelings of motivation and confidence in their ability to implement learning through play. One-third of teachers described overcoming their initial anxiety and fear about failing or making mistakes. This also correlates with the decrease in 'Teacher knowledge, skill, and confidence" as a perceived barrier to implementing learning through play seen in Section 4.1.

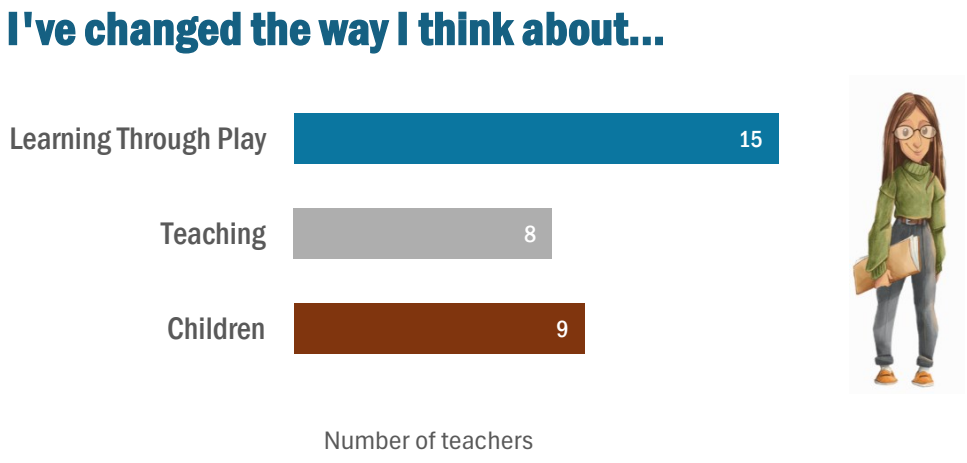
Being filmed by the Pedagogy Partner added additional stress for some teachers at the beginning of the program. When the Pedagogy Partners were asked to reflect on changes in the teachers they worked with, all five identified increased confidence, a greater willingness to take risks and less anxiety about being judged if things did not go to plan. Several Pedagogy Partners noted that as teachers overcame their fears, they also started to show signs of being more relaxed and enjoying themselves during learning through play experiences even when being filmed.



Changes to teacher thinking

Sustainable and meaningful changes to practice do not happen without changes to thinking. There is evidence that the teachers who participated in the TIPP experienced noticeable changes to their thinking. When asked about changes to their thinking, the teachers identified three areas of change as shown in Figure 13 below.

Figure 13: Self-reported changes to teacher thinking



The most noticeable changes to thinking involved:

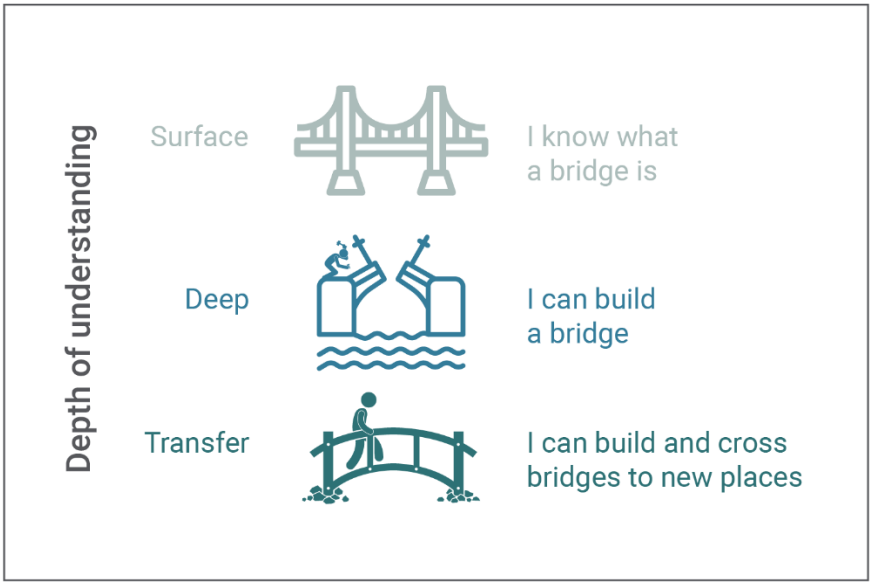
- Seeing the value of learning through play and the connections between play and holistic development
- Seeing children as more capable than they previously thought
- A new vision of themselves as facilitators or partners in learning

As teachers understanding of learning through play deepened, this often led to an expanded vision of the possibilities and potential of play as a pedagogy.

Surface, Deep, and Transfer Learning

The idea of learning as a process of moving from accumulating foundational knowledge to a deeper level of understanding and, finally, the ability to transfer that understanding to new contexts is well established in the field of adult learning (Merriam & Leahy, 2005; Sharff et al., 2017). At the surface level people accumulate facts about something as they develop foundational knowledge about a concept. At a deeper level, the person can see how these things are connected to each other and to other ideas or concepts. Finally, at the transfer level, people can apply what they have learnt to new and unfamiliar contexts. These levels are represented as a continuum of depth of understanding in Figure 14. A compelling argument has been made that one of the biggest challenges facing adult education and training is the lack of success in achieving the goal of learning transfer (Roumell, 2014; Sharff et al., 2017).

Figure 14: Depth of understanding continuum



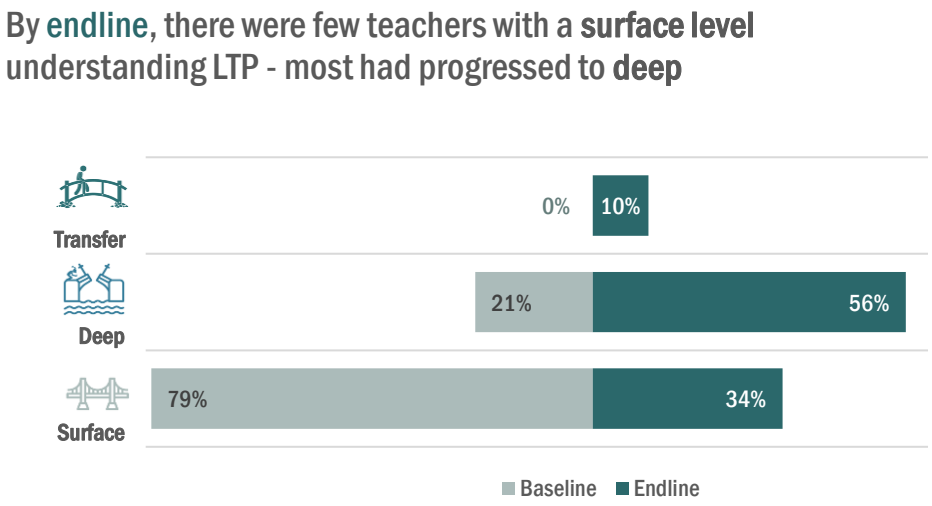
To investigate changes to teacher understanding of learning through play, we applied the lens of surface/deep/transfer to analyse their descriptions of learning through play. We then workshopped this with a group of LEGO Foundation Master Trainers from Ukraine. Table 1 shows the differences between understanding learning through play at a surface level, deep level, and transfer level, with examples from the TIPP teacher interviews presented first, followed by additional examples from the Master Trainers.

Table 1: Teacher descriptions of learning through play

SURFACE LEVEL *I might see play as a threat to teaching and learning.*	*DEEP LEVEL* *I see the bigger picture and understand the value of learning through play.*	*TRANSFER LEVEL* *I have a whole new way of thinking about teaching and learning.*
The children find it fun.	The children find it so meaningful when it is their ideas and their creation.	The children and I are partners in learning.
It's good for motivating children and making learning more interesting for them.	Learning through play helps children to become more confident and not afraid of mistakes.	I've discovered that I am learning through LTP too.
Learning through play is when I add things like games, videos, or hands-on activities into my lesson.	Learning through play is not just an activity, it is a new way of thinking.	Together we find new ways of learning through play.
It can lead to conflict or problems that I have to solve.	The children support each other and find solutions together.	The children and I solve problems together.
It can be hard to find time for play because I have a lot to get through.	The children are learning how to organise themselves and work together in teams.	Children can be teachers too. I love learning from the children!
Additional examples from the Master Trainers		
Working in teams is noisy, children are difficult to control and difficult to calm down.	I understand the value of open-ended tasks and questions and offer them to children	I cannot imagine teaching without learning through play.
Learning through play can be implemented from time to time but the parents are interested in academic skills.	The children are striving for knowledge and are much more motivated to learn at school.	I'm flexible and change my plans to follow the ideas and interests of the children.

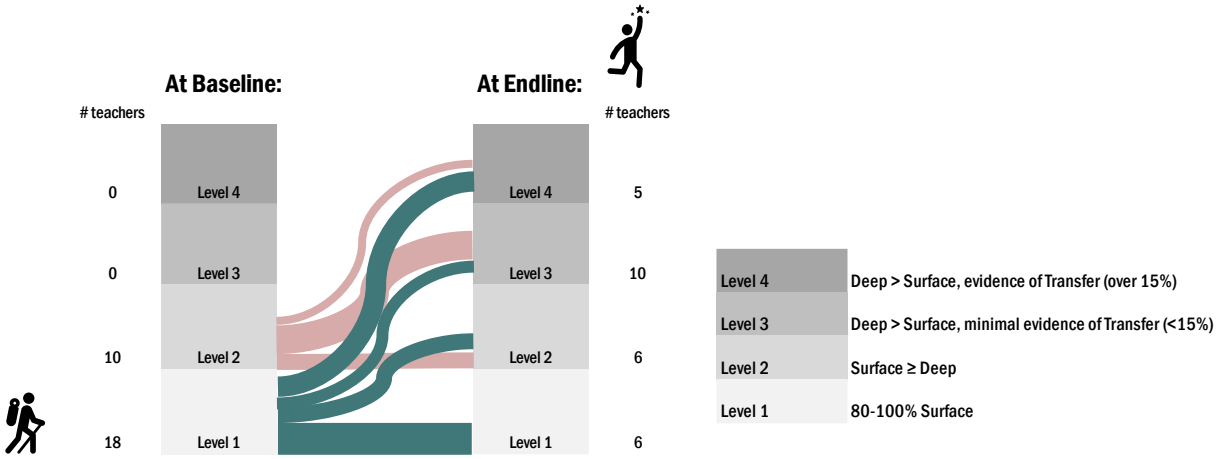
We investigated changes to teacher understanding over time by analysing teacher descriptions of learning through play at both baseline and endline. Looking at the group as a whole, we can see that they were primarily at a surface level of understanding of learning through play at the start of the program, with some evidence of deeper understanding present and no evidence of transfer. At the endline, we see that the group has deepened their understanding and there is evidence of transfer emerging (see Figure 15).

Figure 15: Changes to understanding across all teachers from baseline to endline



Of course, not all teachers had the same journey of learning. Teachers came in at different starting points in terms of their understanding and they ended at different points as well. In Figure 16, we can see the different pathways that teachers took in developing their understanding from baseline to endline. The size of the coloured line reflects the number of teachers that followed that pathway.

Figure 16: Changes to understanding by teacher from baseline to endline



At baseline, most teachers entered the program with a surface level understanding of learning through play, but others demonstrated some evidence of a deeper level of understanding. Looking at the teachers in Level 2 at baseline, they were spread across the regions and schools, with no discernible pattern that connected them.

At endline, we saw that teacher understanding spread across the four levels and ranged from surface level to deep and transfer. It is worth noting that even though a teacher might not have moved from one level to the next, this does not mean there was no learning taking place. For example, one teacher began the program with a limited understanding that learning through play was something children enjoyed but made the teacher's job more difficult, and that it involved using a variety of hands-on materials. At endline, the same teacher was able to list many different activities she used when implementing learning through play and provided more detail on why

children enjoyed it but had not yet made the connections to suggest her understanding had moved into the deep level. She was progressing towards Level 2 but had not yet reached it.

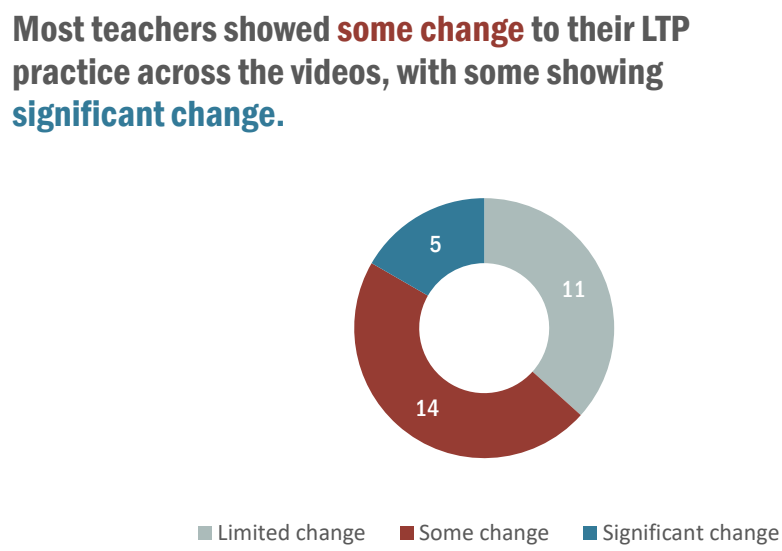
Changes to teacher practice

To investigate possible changes to the way teachers used and facilitated learning through play in their lessons, we analysed endline teacher interviews/reflections, endline Pedagogy Partner reflections, and classroom lesson videos. Over the course of the professional learning program, there were noticeable changes to the way teachers used playful pedagogies with their classes. To begin, we will look at the observed changes in the lesson videos.

Observed changes to teacher practice

Most teachers demonstrated some change to their practice in the lesson videos over the course of the TIPP, as seen in Figure 17. Changes to practice included increased opportunities for children to work together, to make choices about their learning and take responsibility during learning. For example, choices about roles they would take in groups or what resources they would use. There was also evidence that play was becoming more integrated into the learning focus of the lesson, rather than a disconnected activity or break from learning.

Figure 17: Observed changes in videos of LTP experiences



Small group/pair work

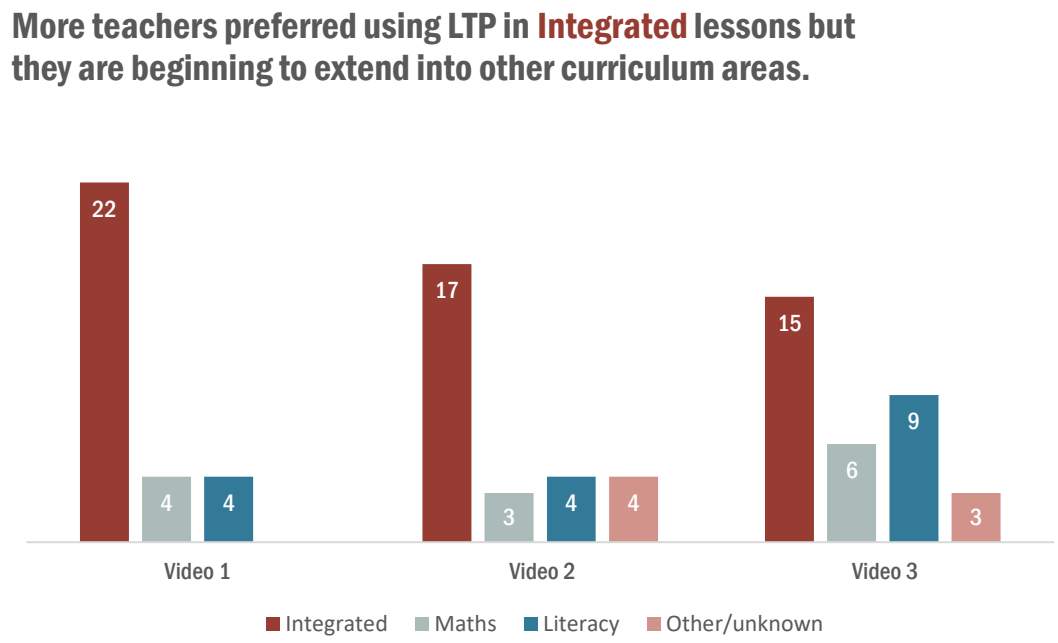
One of the first observable changes in the lesson videos involved an increase in small group/pair activities, suggesting this might be a good first step in transitioning from a traditional, teacher-directed approach towards a more child-centred approach as described in the NUS. The Pedagogy Partners also identified this as an area of focus for many teachers:

> *It was teamwork that could be difficult for teachers. Why? Because usually children sit at desks in pairs, each working in their own notebook. And teamwork is the challenge for now. Teachers called it their developmental area. And it was difficult for teachers to understand and accept that children should work in a team, not individually in their notebooks, but do something together. They realised that children need such skills, and we can give them the opportunity to learn through play.*

LTP across the curriculum

Looking at the lesson videos, we see that teachers chose to implement LTP within the integrated curriculum area of the NUS (*I explore the world*), suggesting this may be the easiest entry point for teachers as they begin to develop their skills in implementing LTP. Over time, teachers did begin to expand their use of LTP into other areas of the curriculum (see Figure 18).

Figure 18: Curriculum area chosen for recorded lessons



Teacher reflective practice

Additionally, three of the five Pedagogy Partners observed a noticeable and important change relating to teacher reflective practice. One described it as the most useful thing the teachers learned because it enabled them to see and understand the changes that were happening to them. They noted that reflection, in the deep and intentional way it was used in the TIPP, was a new practice for these teachers and one that motivated them and "helped them move forward". Being able to reflect with their Pedagogy Partner helped teachers to develop their skills, with the final reflection providing evidence that many had been able to move to a deeper level of understanding because they were able to think about their thinking and practice, see things from different perspectives, make connections and notice changes in themselves and in the children. Without these skills, it is difficult to imagine the teachers would be able to move from a surface level of understanding to a deeper level. The importance of shared reflection (Roumell, 2014) and metacognition (Sharff et al., 2017) to transfer of learning has been identified in the literature.

Support for agentic learning

As teachers developed a deeper understanding of learning through play and became more confident and reflective about their implementation of playful pedagogies, there were other observable changes to their practice. These included becoming more flexible and responsive to input from the children during lessons, becoming more willing to take risks and less anxious about things not going to plan, and becoming more supportive of children's agency and voice within lessons. Teachers described the shift to seeing themselves as facilitators or partners in learning. These changes were identified by the teachers themselves, as well as their Pedagogy Partners. To further investigate how this change in thinking might have been reflected in their implementation of learning through play,

we turned to the teacher interviews. Teachers were asked at both baseline and endline to describe examples of how they implement learning through play, and these were analysed to determine if it was a teacher-driven practice (i.e. where the teacher made all the decisions) or a co-constructed practice (i.e. where they collaborated with children to make decisions). Some teachers gave examples of both within their interviews. Examples of the co-constructed and teacher-driven approaches include:

Co-constructed approach to LTP

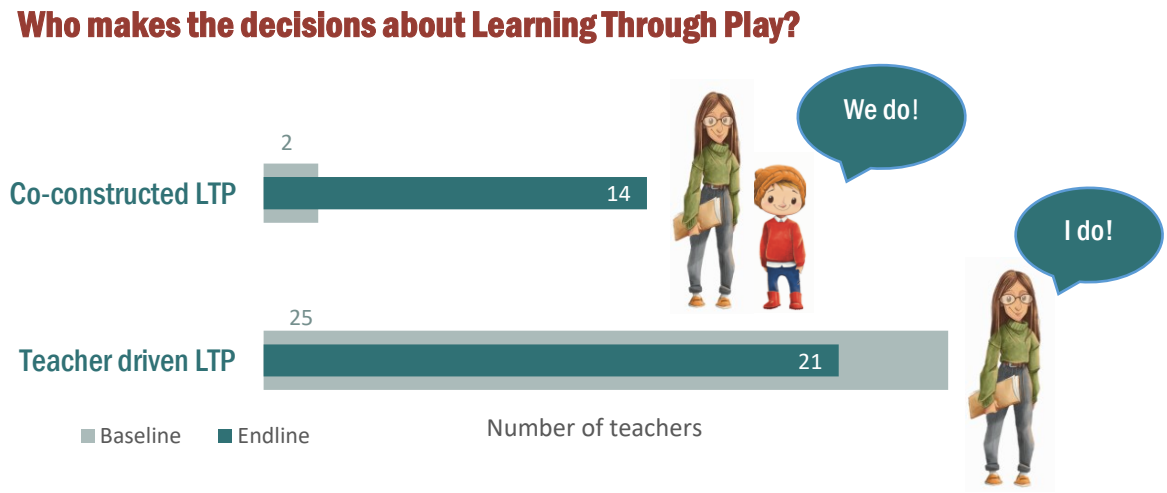
- The teacher following suggestions made by children
- Children becoming the teachers
- Teachers introducing a broad focus for learning and then children deciding on what they wanted to investigate within that focus
- Children and teachers collaborating to extend learning that was particularly interesting for the group

Teacher-driven approach to LTP

- Integrating games or movement activities into a lesson
- Planning for hands-on activities
- Providing opportunities for children to work in groups
- Facilitating discussions so children can share their ideas

The results showed an increase in the number of teachers who described examples of a co-constructed approach to learning through play (see Figure 19).

Figure 19: Becoming a partner in learning through play



Successful professional learning

A key contribution of this study is to expand our thinking about how teachers develop their skills in implementing learning through play at school. Changes to practice do not happen in isolation and are intertwined with the teacher's thoughts and feelings about the change. Looking only at what is happening in the classroom limits our capacity to understand the impact of professional learning programs on teacher learning. The goal of professional learning programs is that what is taught is learned and that the learner is able to transfer that learning beyond the program and into their daily life. Unfortunately, this goal is not often met, with some arguing that this may be due to a lack of intentional designing for transfer. Key features of designs that enable transfer of learning are:

- Providing one-on-one coaching (Merriam & Leahy, 2005)
- Providing time and opportunities to apply learning within the work context (Merriam & Leahy, 2005; Roumell, 2018)
- Embedding opportunities for shared reflection (Roumell, 2018)

- Creating social support networks (Merriam & Leahy, 2005; Roumell, 2018)

The findings of this study suggest the TIPP supported many teachers to move beyond a surface level understanding of LTP, with some teachers demonstrating signs of transfer of learning. This success may be due in part to certain elements within the TIPP design. Of prime importance, according to the teacher reflections and feedback, were the Pedagogy Partner sessions, providing ongoing, individual coaching that focused on supporting teachers to apply the ideas from the online modules in their classrooms. Further, the sessions involved repeated opportunities to engage in shared reflection and goal setting. Beyond these activities, the Pedagogy Partners provided support and encouragement to help teachers engage in the learning program. Finally, the online modules provided practical ideas and examples to support teachers to connect and apply ideas to their own context. The program also encouraged and supported teachers to connect with and support each other within online discussion forums. These intentional design decisions may have contributed to the success of teachers moving from surface to deep and transfer.

Spotlight: Becoming Partners in Learning

One of the biggest changes we saw in our TIPP teachers was a shift in their thinking about children. There was a growing appreciation for the capacity of children to actively contribute to the decision-making about learning through play. Teachers also began to value children as important partners for their own learning, as we can see in the example below.

Teacher:

The most important motivator is the children. Their dedication, shining eyes, support, and enjoyment of the process. When we were studying symmetry, the children were divided into groups, some of them painted on the window, they were involved in the creative process, others made symmetry with LEGO bricks. And it was this interaction, when they supported each other and helped each other, performed collective work, but everyone had their share of responsibility, and everyone tried to make their own personal contribution. Even if it was some kind of support to clean up or present the work, or mechanical actions like to glue something, to help... the children were involved and it was these shining eyes, the children who wanted to learn…all this was very interesting and stimulated me to develop and supported me.

Pedagogy Partner:

The biggest challenge was to let go of the children, to give them freedom and to believe in their capabilities…The situation has changed, but it took time to believe in children. I can already see that children are at the centre of the process. Children who are happy and interested in learning, arguing their opinions, and planning their activities. In learning through play, we mean learning with joy, the concept of joy includes not only laughter or incredible pleasure, but also a situation of success, enjoyment of the process of interaction, frustration in the face of challenges and the success they feel when they cope - we put all this into learning. And one of the teachers was very afraid to show this joy herself, to smile, as we say, to relax and trust the children. And when we finally offered to let the children go, we found out what they needed, what they wanted and how they wanted it. The theme was "Symmetry", it was the New Year, and the children were united in teams, making snowflakes, hanging them symmetrically, creating New Year's decorations from

LEGO bricks, and you know, it was incredible that the teacher allowed the children to paint the windows. And this teacher allowed them to paint the windows in our Ukrainian school! They took the paints, they painted, they stood on the windowsill, this creative chaos, it was so new to her, she saw the children's eyes, she saw how important it was for the children! And when I was at the last lesson, I said: "[Name], you have a smile on your face, you are happy together with the children!".

4.3. Research Question 3: Literacy and social-emotional skills comparison between baseline and endline

In this chapter we answer the research question: *How do children's literacy and social emotional skills compare between testing time points including prior to and during the invasion of Ukraine by Russian armed forces?* This question included two sub questions as follows:

- *When comparing the control and intervention groups, is there a difference between children's literacy, and social and emotional skills including collaboration, at the start of intervention?*
- *When comparing children learning in different locations in Ukraine, is there a difference between children's literacy and SE skills?*

The LTP LSSA was a playful measure of children's literacy (receptive and expressive oral language) and social-emotional skills including empathy, conflict resolution, problem solving and self-regulation. It was designed to be playful and engaging and included items measuring children's experience of undertaking the assessment. It was administered in late 2020–early 2021 when sampled students were starting Grade 1 and again in April 2023 when the same students were completing Grade 3. The number of students participating in the assessment at baseline 2020-2021 was 1460, from 20 schools in Kyiv, Poltava, Dnipro, Zaporizhzhia and Kharkiv. In 2023, the total number of participants at endline was 296 from seven schools: four in Kyiv and three in Poltava. This much reduced sample was due to the number of families leaving the community or country due to the war. This section includes the results of students in Kyiv and Poltava at the two time points, comparing the literacy and social-emotional achievement of intervention and control students.

Comparing literacy and social-emotional skills between two time points

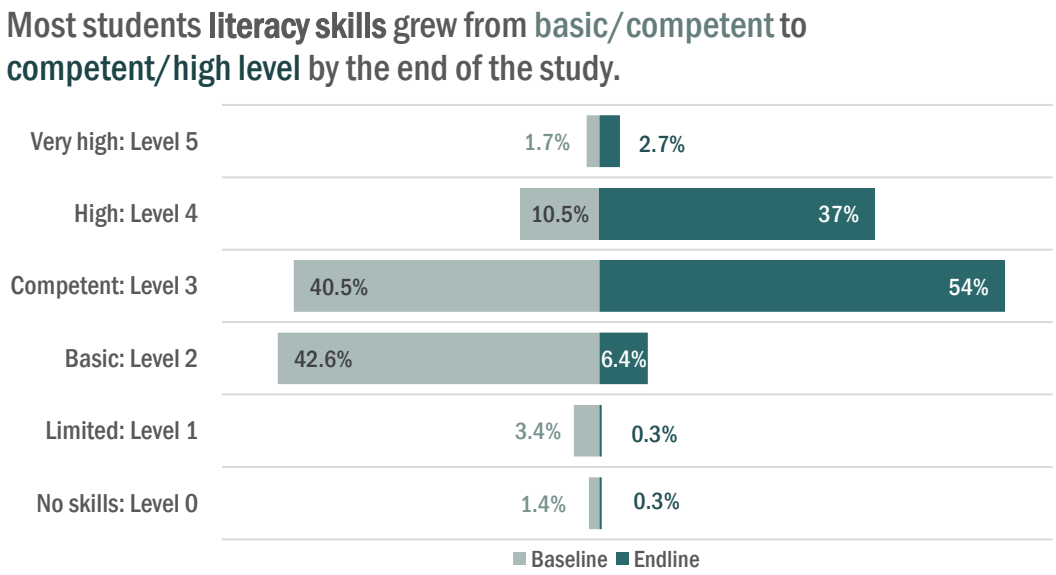
Literacy (oral language) – baseline and endline

Literacy results at endline showed in general **all students grew,** and the most substantial changes were the **increase in the number of students demonstrating high level literacy skills (71% increase)**, and the **reduction of students demonstrating basic literacy skills (85% increase).**

From baseline to endline, the number of students demonstrating basic skills, for example, listing a few disconnected features of a character, and making literal inferences from spoken text, had reduced from near half of the sample (42.6%) to 6.4%. Figure 20 shows the number of students demonstrating high level literacy skills increased from 10.5% to 37%, indicating substantial improvement in students' skills with most students demonstrating competent or high-level oral language skills at the end of the study. Examples of high-level literacy skills were 'Imaginatively generating and extending descriptions of characters and their actions' and giving three different

synonyms for a familiar adjective. Only a small percentage (2.7%) demonstrated 'very high' literacy skills at endline. A full description of the skills at each level can be found at Annex 3.

Figure 20: Literacy skills at baseline and endline, all students



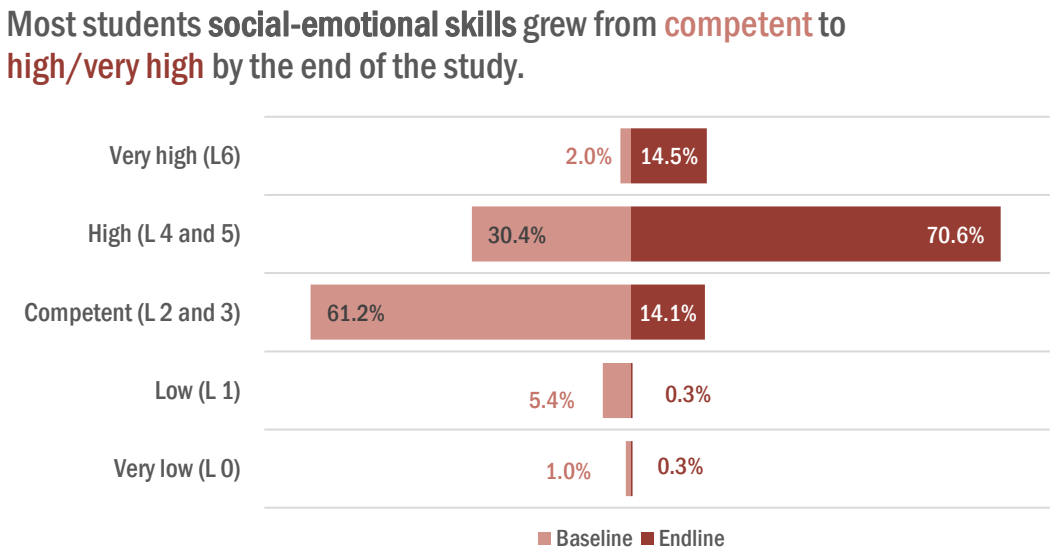
As expected, most Grade 3 students assessed found the literacy items much easier to complete compared with when they were in Grade 1. Performance was distributed across four levels at the end of the study, which is a good result for an assessment used at two different time points. At endline, the highest achieving one third of students found the items generally straightforward although not all skills were demonstrated. There was a good demonstration of difficulty for the remaining two thirds of the sample including enough very simple items to capture the skills of the students with the lowest ability. These results suggest the assessment had a good 'floor' and 'ceiling.'

Social-emotional skills - baseline and endline

Social-emotional results at endline showed in general **all students grew**, with the biggest change being the **increased number of students demonstrating 'high' social-emotional skills** (57% increase), and **large reduction of students demonstrating 'competent' social-emotional skills** (77% decrease).

Social-emotional skills results were described across six levels from very low to very high (see Figure 21). At the end of the study, most students (85%) demonstrated skills considered high to very high. This means these students could propose collaborative solutions to conflicts (high), identify how a proposed interaction is likely to escalate (high), and show insight and empathy in evaluating a character's attitudes and behaviours (very high). Only 14.4% of students demonstrated 'competent' social-emotional skills and very few students demonstrated low or very low social-emotional skills.

Figure 21: Social-emotional skills, baseline and endline, all students



The endline data showed that the social-emotional component of the assessment was generally straightforward for approximately two-thirds of the students with only five items addressing skills that some of these students were unlikely to demonstrate. Most of the items in the assessment were well distributed across the range of ability of the lowest third of students including items to cater for students with the least ability.

Comparing intervention and control groups

Literacy (oral language) – intervention and control groups

Control school students performed higher in literacy, but intervention students grew more between baseline and endline (10.29 score points), which started to **close the gap** in literacy abilities.

Figure 22 shows that at baseline, the control group had an overall higher ability than the intervention group. This was also the case for the lowest ability level. There were more control students in the higher three levels, and fewer students in the lower three levels than the intervention group.

Figure 22: Literacy, intervention and control, baseline

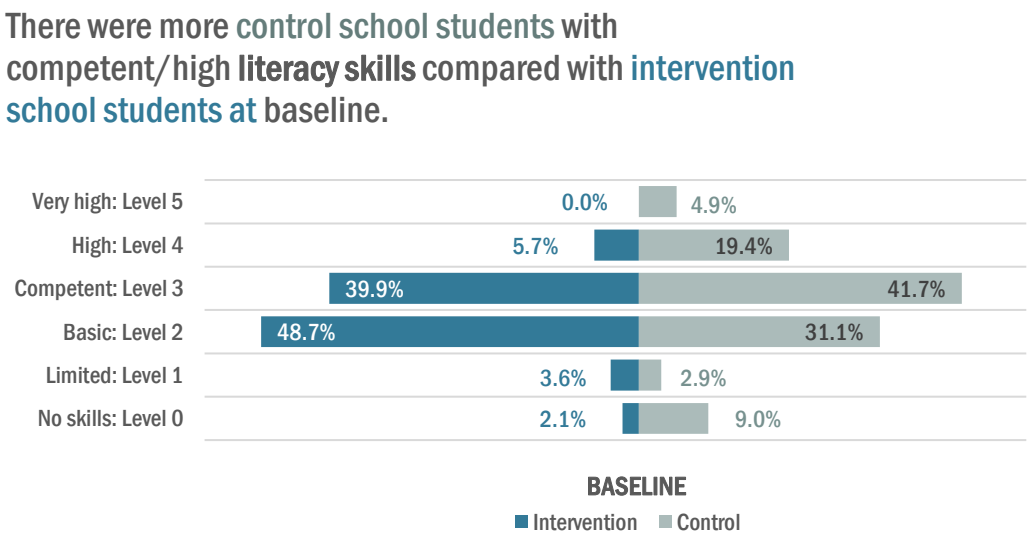
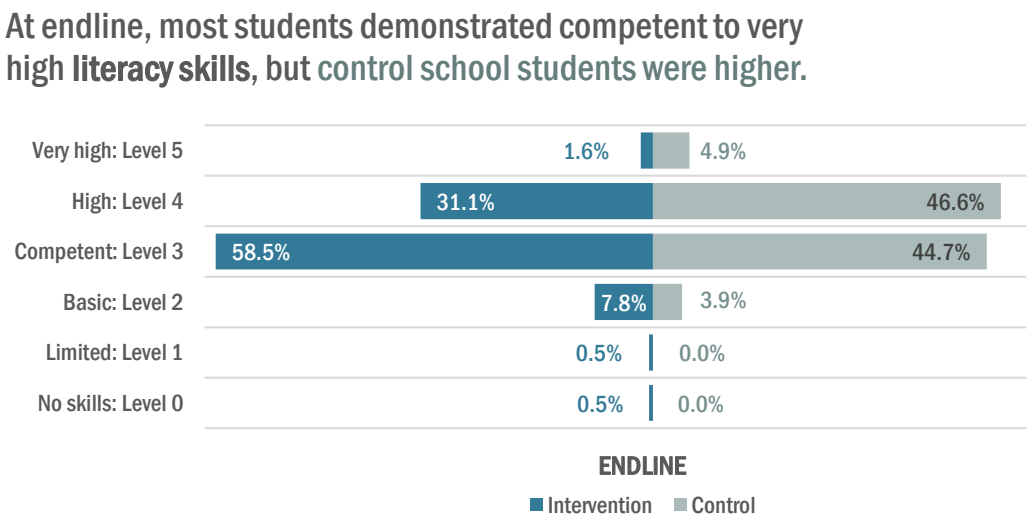


Figure 23 shows that at endline, the control students still had a higher average ability than the intervention students. It is interesting to note that the percentage of control school students in the 'very high' level stayed the same from baseline to endline (4.9%). Nevertheless, there were more control students in Levels 4 and 5 and fewer in Levels 2 and 3 compared with the intervention students. Both groups improved from baseline to endline.

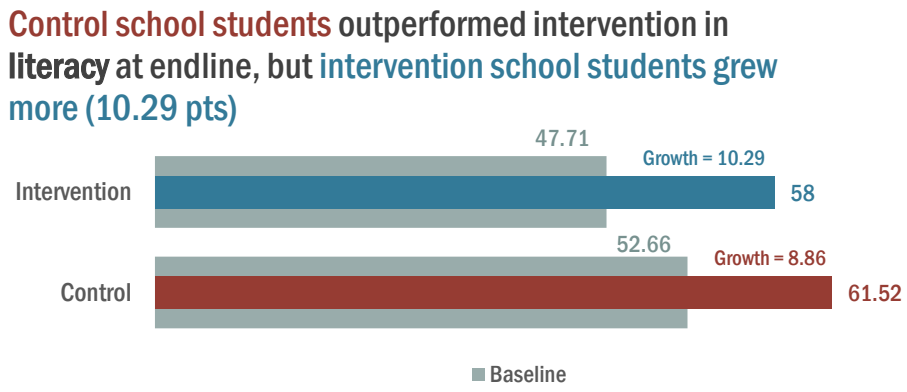
Figure 23: Literacy, intervention and control, endline



Literacy skills growth

The baseline results shown indicate the average literacy ability of the intervention group was 47.71 and control school students was 52.66. At endline, the average score of the intervention group it was 58.00. control group at endline 61.42. As illustrated in Figure 24, the intervention school students demonstrated more literacy growth than the control school students between the two timepoints.

Figure 24: Literacy growth, intervention and control

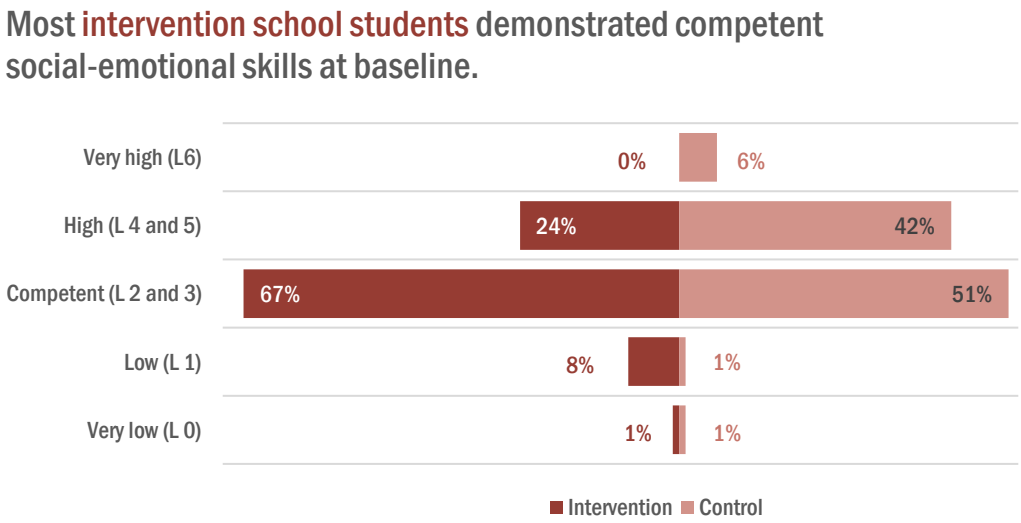


Social-emotional skills – intervention and control groups

Control school students outperformed intervention school students on social-emotional skills, but intervention school students grew more between baseline and endline. This created the effect of **closing the gap** between the two groups.

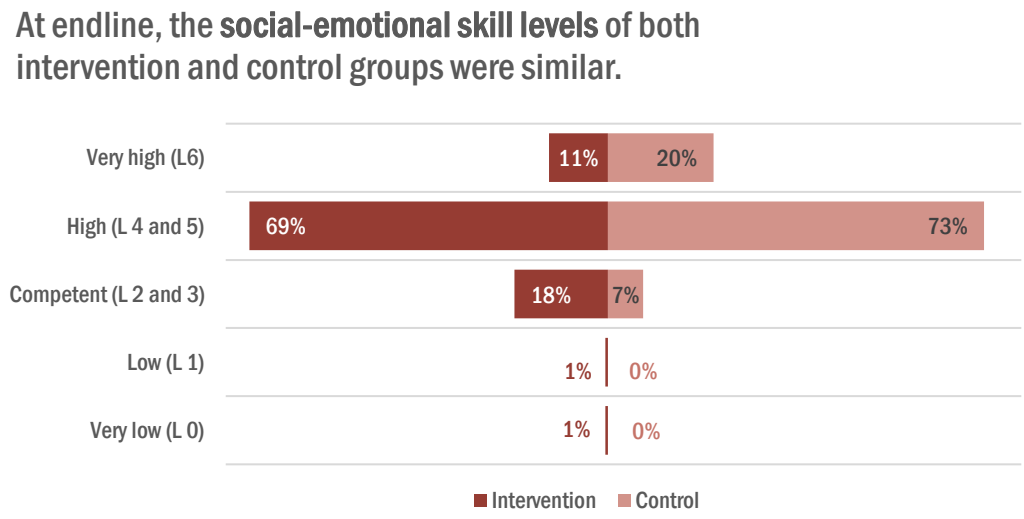
Figure 25 shows that at baseline both the intervention and control groups had a wide distribution of social-emotional skills, but the control group had more students in the higher levels and fewer students in the lower levels compared with the intervention group.

Figure 25: Social-emotional skills, intervention and control groups, baseline



As shown in Figure 26, at endline the control group had higher levels of social-emotional skills with more students in levels 4, 5 and 6 and fewer students in the lower levels compared with the intervention group. However, the overall pattern of improvement was similar between both groups.

Figure 26: Social-emotional skills, intervention and control groups, endline



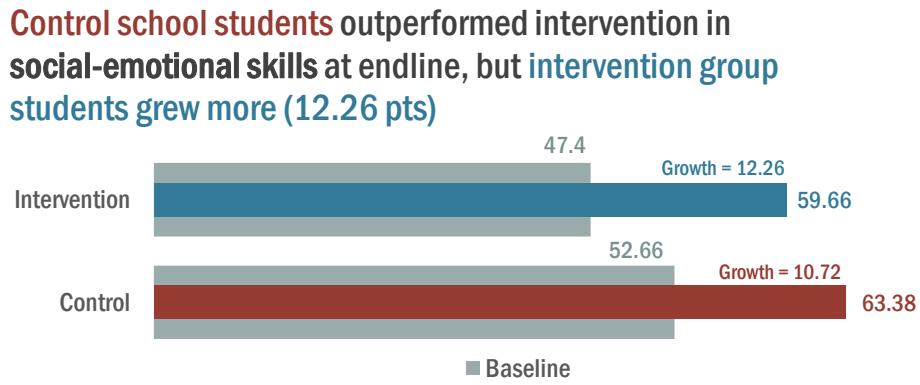
Other data including endline teacher interviews offered supporting evidence of children's social-emotional growth by the end of the study. One teacher said:

> *They now take the initiative, they help each other, you ask a child a question, he doesn't know, I see in the chat room there is help, other students are already helping, giving suggestions. That is, they empathise with each other, even when questions arise: When a student says, "I don't understand this material," they take the initiative: "I'll call you after class and help you. I've already done it and I know how to do it."*

Social-emotional skills growth

At baseline, the average score for social-emotional skills of the intervention group was 47.4 and control group was 52.66. The average score of the control group was 63.38 and for the intervention group it was 59.66. As illustrated in Figure 27, the intervention school students demonstrated more social-emotional growth than the control school students between the two timepoints.

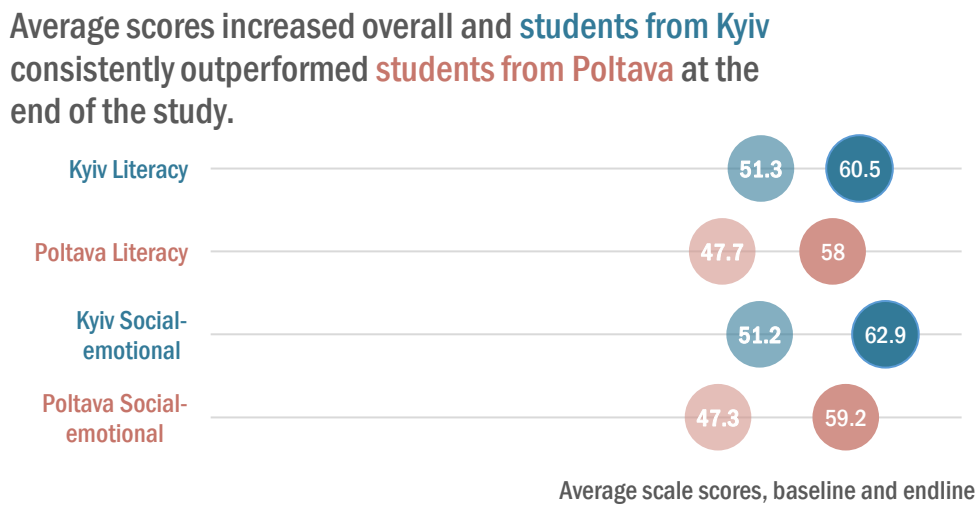
Figure 27: Social-emotional skills growth, intervention and control



Comparing results by region

Presented here are the results of the literacy and social-emotional items by region. Due to the small sample size at endline, no further disaggregation can be made. As illustrated in Figure 28, students from Kyiv performed higher than students in Poltava in both literacy and social-emotional skills.

Figure 28: Results by domain and region



Difference in socio-economic status

To understand the results, it is important to acknowledge that the intervention and control schools were not perfectly matched regarding socio-economic status. As shown below, the intervention school group included far more low SES students than the control group. The results must be considered in light of these differences. The positive correlation between SES and student achievement has been confirmed by existing research (OECD, 2016; Marks, 2017; Filmer & Pritchett, 1999). Hattie (2018) found that SES has a moderate effect on student achievement in his meta-analysis of studies analysing the factors affecting student learning., Accordingly, we do not know whether these results would be replicated if the groups were equally matched on SES composition.

Figure 29: SES at baseline

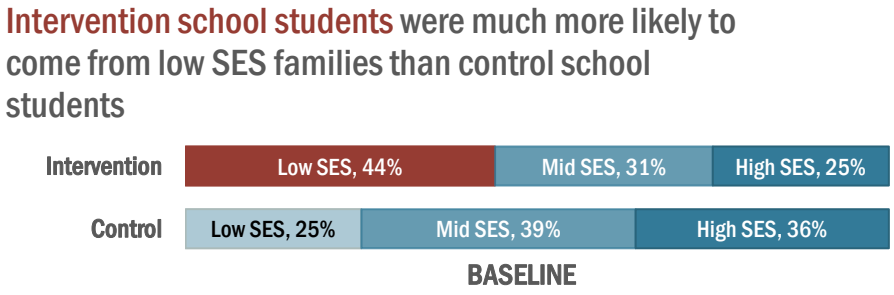


Figure 30: SES at endline

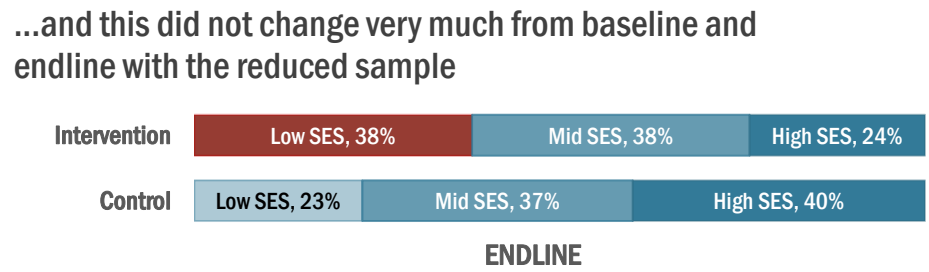
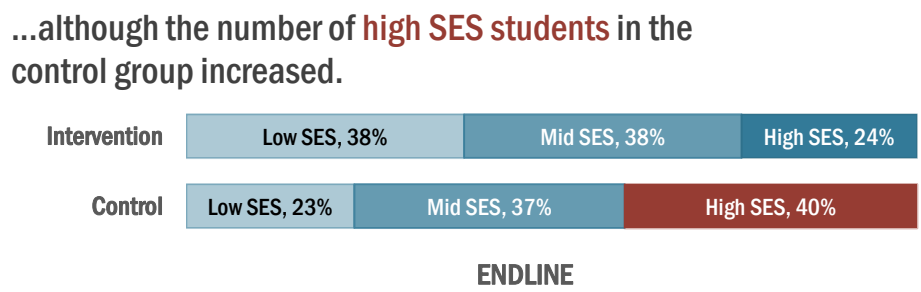


Figure 31: High SES in control group, endline



Teacher perspectives on student growth

At the end of the study, teachers were asked about whether they had noticed any benefits of LTP for students. Teachers provided detailed descriptions of children's growth that elaborate the findings of the assessment. These included:

1. Improved **social-emotional skills** such as demonstrating increased empathy toward each other and the teacher, confidence to try, ask each other, make mistakes, and persist when struggling, motivation to complete work, agentic learning such as seeking out information and reaching solutions themselves, goal setting, and problem solving.

2. Improved **literacy skills (oral expressive and receptive communication)** such as collaborative communication skills to work effectively in teams, facilitating classroom discussions, express thoughts, comment on ideas, listen to others, collaborate on projects or problems, creativity through generating ideas, composing stories.

Several teachers referred to the growth of particular **children who had previously struggled** and made gains via LTP in self-management, that is, regulate their emotions, manage stress, and control impulses. One teacher said:

> *When they work in pairs, they learn to communicate. This will be a lifelong experience for them, how to communicate with other people. Even if they come to work in a large organisation with lots of people, they will already have experience of dealing with people. They learn to control their emotions, for example I had a boy in my class who moved from the Donetsk region and had big problems with his behaviour….it was impossible [for him] to attend classes. Little by little, through conversations, games, work in pairs and teams, he learned and calmed down, realised what he could do and what he could not do…He found the mistakes himself and corrected them, even when he did something wrong, he came and apologised.*

Learners' ability to **work flexibly and supportively in teams** was cited as a growth area. Several teachers referred to children's preference for teamwork and capacity to initiate positive interactions with each other. One teacher said:

> *I often hear that in high school, children have a big problem with working in groups and they refuse: no, I won't do it. I can already see the advantages of our study. When there is an air raid, we go down to the shelter, we have tables there and have to work in groups. They just sit down without looking up, I give them a task and they work with whomever they sit down with. There are no problems now. They always try to help each other, even when they don't need to. They have developed such a team skill of social interaction. I say, 'can I help?', and they say… 'No, you don't need to'.*

Summary

The growth of almost all students in literacy and social-emotional skills from Grade 1 to 3 is a positive result from a unique dataset. Data collection commenced in late 2020 when COVID-19 restrictions and illness were severely impacting teacher and student attendance at all schools. Very few studies have quantified the scale of learning loss caused by the pandemic without the use of predictive models. Donnelly and Patrinos identified only eight such studies (2021). The findings from this assessment adds to our collective understanding of learning during the pandemic, particularly the development of social-emotional skills such as conflict resolution, problem solving, self-regulation and empathy. There is very little empirical research regarding the development of these skills in Ukraine, much less during the pandemic, and no evidence on the impact of the invasion on children's social-emotional learning. Researchers have found COVID-19 to negatively impact children's social-emotional skills development (Rodriguez-Monge, Isabela & Chiapelli, 2023). Further, evidence regarding the impact of war on children's learning describes how trauma interferes with cognitive processing (Diab & Schultz, 2021). But here we see growth despite immense challenges. While the context is specific and sample is small, there is potentially a finding here about how playful learning designed to foster positive relationships and experiences may have been a protective factor supporting children to continue to grow. One teacher interviewed confirmed that playful pedagogies have supported children's social-emotional health and wellbeing at this extremely difficult time:

> *Speaking specifically about personal [attributes], especially nowadays when children's emotional state is not very balanced, [learning through play] allows us not only to learn but also to balance the child's emotional state, which is very important.*

5. Discussion

The scoping study that predicated this research argued for a series of new directions to address opportunities and gaps in LTP research (Parker & Thomsen 2019). It described the need for more evidence from low and low to middle income countries and also:

- Good practice examples of learning through play
- Studies of learning through play in key transitions, namely the first three years of school
- Studies measuring the impact of learning through play on non-cognitive including social-emotional, physical, and creative skills.
- Studies that contribute to understanding the impact of system-level reforms

This study offers evidence across all areas.

Good practice in learning through play

The persistent lack of clarity about what good practice looks like when implementing LTP presents a major hurdle for practitioners who are trying to develop their practice. While there are numerous frameworks that describe the different elements of LTP (e.g., characteristics of LTP experiences, forms of play, spectrum of facilitation practices, etc.), they do not provide adequate guidance to support teachers who are trying to learn and develop their use of LTP. Teachers are still left wondering how to apply this knowledge in their daily decision-making and have no way of knowing if they are "doing it right" or what the next steps might be. It is akin to providing a beginning cook with a list of ingredients, but no finished product to aim for and no recipe to follow. At the core of this problem is a continued lack of consideration of **teachers as learners** going through a process of learning. What they might need as they strive to learn about LTP, and how those needs might change over time. To understand what this learning process looks like, we need to work with teachers as they learn to implement LTP within their classrooms. The continuum of understanding

developed in this study provides a first step in mapping out the development of understanding as teachers move from surface level applications through to deeper and more transformational approaches to LTP. This provides a promising opportunity to think about existing frameworks through the lens of the developing practitioner with the aim of making them more meaningful and useful to teachers as they learn.

The continuum also provides a promising pathway for expanding our thinking and approaches to evaluating impact and teacher change in relation to LTP interventions. Shifting our focus to the process of learning will help us to not only see change in all of its forms (practice, understanding, feeling), but also to more accurately identify what support may be needed along the way.

Learning through play in key transitions

At baseline, the teachers identified children as the most significant barrier to implementing LTP. Of concern was the capacity of children to successfully learn through play. This included doubts about the children's ability to manage conflict, manage and regulate their emotions, and make decisions about their learning. At the core of this challenge was a concern about children's *readiness for school* rather than their readiness for learning through play. The idea that children need to become ready to conform to the requirements of daily classroom life is not new, especially in countries where traditional approaches are dominant (OECD, 2006), as was the case in Ukraine at the beginning of the NUS reform. While many of the teachers in our study did experience a change in their beliefs about the capacity of children and grew to appreciate the value of LTP for supporting children to develop holistic skills, it remains in question whether this experience will result in a different view of children's readiness when the teachers move back to Grade 1 with a new group of children.
At endline, as teachers were preparing children to transition to the next phase of schooling, some expressed doubts about the system's readiness for children who have experienced LTP as a key approach used in the classroom. Of concern was the perceived lack of understanding and use of LTP beyond the early primary years, as well as concerns that the children's agency and voice may not be welcomed, let alone supported, in those classrooms.

Teacher responses at baseline and endline highlight the challenge of LTP in key transitions periods. The challenge requires two fundamental shifts in thinking if we are to facilitate positive and supportive transitioning as children move through the education system. Firstly, a shift in thinking is required to move the focus from making children ready for school to ***making schools ready for children***. Secondly, a shift from seeing LTP as something relevant only for early childhood towards seeing LTP as something that evolves and is valuable for learners of all ages may help pave the way for LTP beyond the early years of schooling.

Measuring the impact of learning through play on holistic skills development

In this study we learned it was possible to develop and administer an assessment of children's receptive and expressive oral language and social-emotional skills that posed open-ended questions, was enjoyable and engaging to complete, and was a reliable measure of these skills. The study provided valuable evidence of children's achievement and growth in these critical domains during a time when it was most needed. The key finding that almost all children grew between baseline and endline in literacy and social-emotional skills is important evidence for all stakeholders. The finding that intervention school children grew more than control schools, to be interpreted with caution, is a promising result for LTP research.

Prior research has found the impact of playful pedagogies is more likely to be measured in terms of literacy and numeracy gains, as these measures are more readily available, and these domains are

prioritised in education (Parker & Thomsen, 2019). However, good quality intervention studies measure the skills *most likely to be influenced by the intervention* using well targeted, valid, and reliable tools. Receptive and expressive oral language and socio-emotional skills are the skills most likely to be influenced in the early years of LTP practice. Oral language skills assessments are less likely to have a 'floor', that is a basic level of skill required to access the assessment enabling greater participation and more accurate measurement.

This study collected a range of data from teachers, school leaders, parents and children using a mixed methods approach enabling us to cross check and validate data from different sources. The result was a rich and detailed database that supported high-level and practical understandings of not only what changed but how and for whom, and what can be learned from this. LTP teaching practice is complex, integrative, context specific, and involves the head and heart. It follows that research methods investigating LTP are correspondingly diverse.

Understanding the impact of system level reforms

This study contributes to the evidence base regarding the implementation and impact of the Ministry of Education and Science of Ukraine's New Ukrainian School reform. The reform describes a key shift toward developing competencies over transmitting content and the findings here illuminate how this is possible, via a blended face-to-face and online teacher professional learning program that included: 1-1 coaching, sufficient time to teachers to change and apply new learning, opportunities for guided reflection, and social support networks. These features contributed to changes in teacher thinking, feeling and practice, and are important to consider when designing new programs in the NUS context.

Teachers and Pedagogy partners reported that LTP had supported children to develop their skills as agentic learners, communicating with teachers, each other, and parents more frequently, asking more open-ended questions and in general demonstrating more curiosity about everything. As stated above, teachers and Pedagogy Partners expressed concern about how these children's new skills would continue to evolve in later years if middle school teachers were not experienced LTP facilitators. In this sense, these children are in the position of driving the continuation of the reform. One teacher said:

> *In our classes, children have learnt to express their own thoughts, comment, listen to others, and collaborate on a project or a problem. When we held an event dedicated to safety, representatives of international organisations came to our school and said: "We have never met children like these! We have held events in all primary schools." …That is, [the children] answered open-ended questions without hesitation, [presented] evidence, and formed a conclusion.*

The study illustrated how teachers had diverse needs and starting points regarding LTP practice. Ministry of Education and Science partners including pre- and in-service education providers would be well placed to consider how professional learning programs could be tailored to meet teachers needs and starting points to maximise meaning, engagement and success. Further pre- and in-service training designed to foster LTP must also be playful in nature, so teachers learn in the same way expected of their students.

Finally, this study offers the Ministry of Education and Science of Ukraine insights into the process and outcomes of change for a group of students and teachers before and after extreme events, namely the global pandemic and full-scale invasion of Ukraine. The war has prompted reassessment of priorities in teaching and learning to focus on mental health, trauma informed practice and

wellbeing. This study illuminates specific aspects of these areas including the development of social emotional skills including self-regulation, empathy and collaboration which are important in the context of trauma, stress, and uncertainty. Further, in sharing the impact of these extreme events, teachers and school leaders have highlighted the adaptability and resilience of the education system and people. A challenge is not something to eliminate, but rather something that we continue to engage with as a natural part of learning and growing. Understanding this may help us to think differently about how we support teachers to become the drivers of change the NUS positions them to be. One teacher said:

> *[Challenge] is a tool to prepare children for the future, they will become people in society, and I am a teacher who knows how to teach children. The challenges change every year. They need to know how to change and how to deal with these changes. I have to prepare them emotionally and socially…I can pass on this ability to be confident to children. For me it's about success, showing my skills and not being afraid, pushing for new things, opening up new opportunities, future success and growth as a teacher."*

References

Adams, R., Wu, M., Cloney, D., & Wilson, M. (2020). ACER ConQuest: Generalised Item Response Modelling Software. https://www.acer.org/au/conquest

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer Netherlands. https://doi.org/10.1007/978-94-007-6857-4

Bubikova-Moan, J., Næss Hjetland, H., & Wollscheid, S. (2019). ECE teachers' views on play-based learning: A systematic review. *European Early Childhood Education Research Journal*, *27*(6), 776-800. https://doi.org/10.1080/1350293x.2019.1678717

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7

Cloney, D., & Adams, R. J. (2021). *Conquestr version 0.9.95*. https://cran.r-project.org/package=conquestr

Dean Jr., D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education*, *91*(3), 384-397. https://doi.org/10.1002/sce.20194

Diab, S. Y., and Schultz, J. (2021). Factors contributing to student academic underachievement in war and conflict: A multilevel qualitative study. *Teaching & Teacher Education,* 97, pp.103211)

Donnell, L. A., & Gettinger, M. (2015). Elementary school teachers' acceptability of school reform: Contribution of belief congruence, self-efficacy, and professional development. *Teaching and Teacher education*, 51, 47-57. https://doi.org/10.1016/j.tate.2015.06.003

Donnelly, R. & Patrinos, H. A., (2021). Learning loss during COVID-19: An early systematic review. *Prospects*, 51:601–609 https://doi.org/10.1007/s11125-021-09582-6

Filmer, D., & Pritchett, L. (1999). The effect of household wealth on educational attainment: Evidence from 35 countries. *Population and Development Review*, *25*(1), 85–120.

Gray, C., & Ryan, A. (2016). Aistear vis-à-vis the Primary Curriculum: the experiences of early years teachers in Ireland. *International Journal of Early Years Education*, *24*(2), 188-205. https://doi.org/10.1080/09669760.2016.1155973

Hattie, J. (2018). Visible Learning. www.visiblelearningplus.com [Website accessed March 2019].

Marks, G. N. (2017). Is SES really that important for educational outcomes in Australia? A review and some recent evidence. *Australian Educational Researcher, 44*, 191–211.

Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47(2), 149–174. https://doi.org/10.1007/BF02296272

Meiers, M., Khoo, S., Rowe, K., Stephanou, A., Anderson, P., & Nolan K. (2006). ACER Research Monograph No. 61 Growth in Literacy and Numeracy in the First Three Years of School. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1000&context=acer_monographs

Merriam, S. B., & Leahy, B. (2005). Learning transfer: A review of the research in adult education and training. *PAACE Journal of lifelong learning*, *14*(1), 1-24.

OECD (2006). *Starting Strong II: Early childhood education and care.* Paris: OECD. https://doi.org/10.1787/9789264035461-en

Organisation for Economic Cooperation and Development. (2016). *PISA 2015 results: Vol I – Excellence and Equity in Education*.

Parker, R., & Thomsen, B. (2019). Learning through play at school: A study of playful integrated pedagogies that foster children's holistic skills development in the primary school classroom. https://www.legofoundation.com/media/1740/learning-through-play-school.pdf

Parker, R., Thomsen, B. S., & Berry, A. (2022). Learning Through Play at School – A Framework for Policy and Practice. *Frontiers in Education*, 7, 751801. https://doi.org/10.3389/feduc.2022.751801

Rodriguez-Monge, M., Isabela, I-P., & Chiappelli, F. (2023). CoViD-19 effects on social-emotional development: impact of early intervention. *Bioinformation.* 19(9): 889-892 (2023)

Roumell, E. A. (2019). Priming adult learners for learning transfer: Beyond content and delivery. *Adult Learning*, *30*(1), 15-22.

Scharff, L., Draeger, J., Verpoorten, D., Devlin, M., Dvorakova, L. S., Lodge, J. M., & Smith, S. V. (2017). Exploring metacognition as a support for learning transfer. *Teaching and Learning Inquiry*, *5*(1).

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. IERI Monograph Series, 2, 9–36.

Wu, M. L. (1997). The development and application of a fit test for use with marginal maximum likelihood estimation and generalised item response models. http://hdl.handle.net/11343/57769

Zosh, J. M., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., Solis, S. L., & Whitebread, D. (2017). *Learning through play: A review of the evidence*. Retrieved from the LEGO Foundation website: https://www.legofoundation.com/media/1063/learning-through-play_web.pdf

Annex 1 Detailed description of methodology

This annex includes detailed descriptions of data collection methods, achieved samples and study participants at baseline and endline, and fieldwork, quality assurance and data management tasks.

Data collection methods

Table 2 summarises the instruments used during the study.

Table 2: Data collection methods

METHOD	FOCUS
Teacher interview	Teacher beliefs and attitudes towards LTP, their experience of LTP in the classroom (e.g., examples from practice, perceived barriers/enablers, goals for this experience), their beliefs about teaching and learning
Parent/carer questionnaire (background)	Short background questionnaire to gather information about student SES including parent/carer education and occupation.
Classroom observation protocol (video recorded)	Qualitative research instrument, for field notes regarding quality of the LTP experience, facilitation of learning, student engagement in the experience
Teacher reflections and planning documents (to accompany video recordings)	Context for the LTP experience, connection to curriculum, goals for learning, intentions for the experience/lesson Reflection on practice, reflection on student learning, goal setting
Pedagogy Partner observations and reflections (to accompany video recordings)	Observations of LTP experience, post-observation discussion with teacher, teacher practice as observed by Pedagogy Partner, teacher thinking/beliefs as revealed during discussions
Student assessment instrument	Literacy development, social-emotional development (including collaboration and conflict resolution).
School leader questionnaire	Impact of COVID-19 and invasion of Ukraine by Russian armed forces on student access to learning, including questions on duration and number of school closures, damage to family homes, school buildings, infrastructure.

Qualitative methods

Semi-structured endline interviews with teachers and pedagogy partners

Semi-structured interviews captured the meanings that participants made of their experiences in their own words. A semi-structured format was chosen to ensure that major touch points were addressed within each interview, providing a degree of comparability across interviews, while remaining sufficiently open to allow the interviewer to personalise the interview by asking probing follow up questions based on the response given. Interviews were conducted online due to the challenges of travel arising from the war.

Of the 30 teachers participating in the professional learning program, 29 consented to the baseline interview and 27 to the endline interview with their Pedagogy Partner. Key questions for teachers at endline were: describe typical ways of using LTP in lessons and why, challenges to implementing LTP, enablers to implementing LTP, describe what LTP means to you now and how is this different from baseline, how has your thinking and feeling about LTP changed, describe benefits for students, describe challenges for students. Teacher interview protocols were designed to mirror the baseline so responses could be compared and changes between time points identified. Key questions for

Pedagogy Partner interviews were: challenges teachers faced when learning about implementing LTP and changes over time, enablers for teachers when learning about implementing LTP and why, challenges faced by PPs when working with teachers and why, rewards of working with teachers and high points, recommendations for the TIPP – keep, change, remove or add.

Interviews were audio recorded, transcribed in Ukrainian and then translated into English for analysis in NVivo.

Analysis

The analysis followed a general thematic approach with the research questions providing the focus for the initial analysis. Initial categories were formed based on responses to interview questions. As an example, an initial category called *Enablers* was formed for teachers' responses to the question, '*What supports you to implement LTP*?' From there, codes from the baseline were used as a framework to identify existing enablers, and new codes were created for new sub-categories. For example, in relation to the category *Enablers*, a new sub-category was formed called 'Positive teacher-child relationships'. The framework established by the lead analyst at baseline was used to sort, code and analyse endline data.

Qualitative video observations

The purpose of qualitative video observations was to develop an understanding of what strategies, teaching practices, and classroom interactions were being used by teachers at the beginning and during the Teacher Innovative Play Program. Indirect video observations allowed the research team to identify what teachers do in the classroom, as opposed to what they think and say about their classroom practice. This was achieved using video data, to investigate the nature of genuine teaching and learning events.

Key observation foci included LEGO's characteristics of play, opportunities for agency (voice and/or choice), types of play, interactions, observers perceived barriers, and enablers to implementing LTP.

Analysis

The teachers' learning through play classroom videos (n=90) were recorded during a planned visit by the regional Pedagogy Partner, who for the most part were nonparticipant observers. Videos were captured prior to February 2022 using a tablet device and uploaded to the LEGO Foundation SharePoint for analysis. Classroom videos were not edited or translated into English. The analysis involved a qualitative exploratory approach that considered interactions, experiences, and opportunities related to learning through play. While viewing the videos, lessons were segmented into separate play events. Events represented parts of the lesson that had a specific focus, they varied in duration and ended when the focus of the lesson changed, e.g., the children listening to the teacher talk moved to the children working collaboratively in pairs. Across each event, instances of the characteristics of play, opportunities for agency, and types of play were noted.

Quantitative methods

Questionnaires

The following questionnaires were administered during the study to gather data on home and school contexts for learning:

- A short parent questionnaire administered to all parents of children participating in the assessment from control and intervention schools, to provide a measure of SES in order to better understand the sample and how well intervention and control schools were matched.

- A short school leader questionnaire, to gather school, classroom and community context data largely pertaining to the combined and separate impact of COVID-19 and the war on student learning and learning contexts.

Student assessment

The LTP LSSA instrument was designed to measure a range of skills developed through LTP experiences in a short, one-to-one assessment. Each student responded to the same set of tasks and the administrator scored their responses on-the-spot using a consistent approach. The instrument comprised 44 questions at endline including three practice items and the average administration time was 25-30 mins minutes per student. The endline instrument also included questions regarding time, place and any interruptions to the assessment caused by air raid alerts and ensuing classroom evaluations.

The LTP LSSA measured key elements of early literacy, social and emotional self-awareness, and conflict resolution strategies. Expressive and receptive oral language was the main literacy focus. To assess speaking and creative thinking, students engaged in imaginative iterative strategies for problem solving including describing two characters and developing a story based on a sequence of three pictures. To assess listening comprehension, students heard a short description of a classroom activity and answered questions about the literal and implied meaning.

Socio-emotional skills were assessed by asking students about what causes some familiar emotions for them and how they try to manage these emotions. Conflict resolution strategies were assessed by presenting an illustrated, conflict scenario explaining the social problem the characters have and asking students to show understanding of the perspective of each of the characters and to critique an unlikely strategy for resolving the conflict and suggest better alternatives.

The administrator also rated the extent to which students actively engaged in the tasks and the students rated themselves about how challenging they found the tasks and how much they enjoyed them.

Domains

There were two components to the literacy assessment and both components had a range of difficulty. These were: 1. Expressive oral language and 2. Listening comprehension (receptive oral language). The skills assessed by the social and emotional item set included empathy, self-awareness and self-regulation, conflict management, teacher ratings of students' attitudes and persistence, and student ratings of their enjoyment and perceptions of challenge. A detailed description of the considerations informing the design of the assessment can be found in the baseline Report dated 22 May 2022.

Changes to the LTP-LSSA at endline

Seven items were deleted from the baseline assessment and ten new items added, to ensure that the endline assessment was sufficiently challenging to capture the full range of ability of students at Grade 3. Two new items were included to capture the language the children used to respond to the questions. These changes were made based on analysis of the baseline results.

Analysis

A technical report of the instrument's reliability, item level statistics, item functionality, targeting, differential item functioning and balance is located at Annex 3. In summary, the report states that the instrument was well targeted and a very good measure of the skills it sought to assess and the range of abilities of student participants.

Sampling

School sampling

The target population was all students (and their teachers), participating in Grade 1 in 2020 in municipal schools in the Poltava, Kharkiv, Zaporizhzhia, Dnipropetrovsk and Kyiv regions of Ukraine. Four schools per oblast were selected including 1xcontrol and 1x intervention in a regional location and 1x control and 1x intervention in an urban location. School leaders were interviewed to ensure their willingness to participate in the study. From this, a sampling frame was provided by the Ministry of Education and Science (MOES) (last updated 5 September 2019), with some exclusions applied by the LEGO Foundation (e.g. schools without primary classes, private schools, small schools). This served as the achieved population. Ten intervention and ten control schools were selected to participate in the study. A full description of the approach to sampling can be found in the Baseline Report.

Student sampling

The initial sampling approach used two-stage sampling: first schools were sampled, then students within schools. Up to 26 students in up to four Grade 1 classrooms at each selected school was invited to participate at baseline only. If there were more than four classrooms eligible, they were chosen at random. If there were less than 3, all classes were selected. The target sample was a maximum of 1,500, minimum of 1,398 students selected from 20 schools. The study design was an approximately balanced design, which is an approximately equal number of students are selected from ten intervention, and ten control schools. At endline, all available students from selected schools who participated at baseline were included in the assessment.

Participants

The final achieved sample at baseline included 1,465 students, 1,442 parents, and 58 teachers. The final achieved sample at endline included 27 teachers from 10 schools and 296 students.

Fieldwork

Operational procedures

The LTP field operations procedures including roles, responsibilities, and resources, were detailed in the Assessment Administrator and Assessment Coordinator manuals. These manuals provided clear instructions for the successful preparation and implementation of the LTP assessment. The source English versions underwent an adaptation process by UERA in consultation with ACER to ensure the standardised procedures were appropriate to the Ukrainian context, before undergoing translation.

Administration training program

ACER adopted a Train-the-Trainer model for the Assessment Administrator training. ACER staff delivered a one-day remote refresher training to UERA trainers, on April 3, 2023.

The purpose of the training was for administrators to develop a deep understanding of the LTP LSSA and the associated field operations. In addition, it aimed to train participants on using the assessment as intended – accurately and consistently and follow operations processes. Finally, it aimed to support UERA on how to train others in how to use the assessment accurately and consistently and how to follow field operations processes.

Assessment implementation

At endline, Assessment Administrators worked in teams of two in both regions, with each team being led by a Team Leader who undertook all the same tasks as the Assessment Administrators, in addition to liaison and coordination duties. All communication from the UERA head office to

Assessment Administrators in the field went through the team leader. The team leader was also responsible for being the contact point for the schools during the assessment window.

Table 3: Test administration schedule

Region	Assessment start	Assessment end
Kyiv	17 April	31 May 2023
Poltava	17 April	16 May 2023

Safety and security planning

In September 2022, the study was relaunched and a revised grant agreement was signed on 19 December which outlined a way forward to answer the research questions provided the team had access to Pedagogy Partners, Teachers, and some students. This new work program was one of three scenarios developed.

To proceed with baseline data collection, a detailed planning process was undertaken to minimise all risks to people, property, reputational risks for all partners, and the study. This work commenced in December 2022 and was completed in April 2023. A Safety and Security Adviser, Dr Pavlo Artomov was recruited, and a field operations risk register was created. Dr Artomov liaised closely with UERA on all aspects of the activity design to assess the risk. Pavlo drew on his experience as a Field Manager for the International Rescue Committee (IRC) in Kharkiv, and training with IRC on safety planning, staff care, and personal safety and security.

A collective decision was made among partners that the study could not proceed with very high-risk rankings with risk effects of injury or death and likelihood greater than 'possible' in the scale of 'unlikely' to 'almost certain'. Risks were minimised through an area analysis resulting in the selection of sample schools in districts with lower likelihood of critical incidents. Schools were chosen based on safe access, access to bunkers, and whether students were learning onsite at these schools. All children who participated were attending school on the day of assessment.

These preconditions excluded most of the school sample. Endline data collection proceeded in 4 schools in Kyiv and three schools in Poltava only. The Safety and Security Adviser developed a detailed Security Management Plan, final risk register, and Infrastructure Premise Assessment which was finalised on March 23 and approved by all partners (ACER, UERA & LEGO Foundation) at the executive level on March 24, 2023. Safety and security training was provided to all assessment administrators on April 6 and assessment commenced on April 17 and was complete on May 31, 2023. There were no critical incidents reported during the fieldwork.

Note that student assessment was the only endline data collection method impacted by the war. All other methods were successfully administered through remote technology and therefore did not involve risks caused by travel.

Data management

The data management plan was developed by the ACER Data Management Team and agreed with the LTP at School team prior to the invasion on Feb 23, 22 and updated in April 2023. ACER delivered a refresher data management training program on June 1, 2023, to train the UERA data manager in:

1. Collecting student and teacher lists from schools
2. Preparing student and teacher tracking forms
3. Preparing materials for the assessment

4. Preparing materials for data entry when they are returned by Assessment Administrators after the assessment
5. Using ACER Maple data management software to
    - Enter;
    - Validate; and
    - Submit tracking and response data.
6. Participate in data cleaning after the data is submitted to ACER.

Data entry was due for completion on 30 July but delayed to 16 August 2023 due to issues with configuring the data entry software.

Limitations

Sampling

This sample is not representative of a predefined population and the probability of selection for both schools and students could not be known. Therefore, no student weights were computed. In addition, sample descriptives (such as means or percentages) are not estimates of population statistics, and standard errors to estimate the precision of the population estimates are not relevant. Therefore, no tests of statistical significance can be conducted using this approach; differences in mean performance for groups based on the variables can be reported using descriptive, rather than inferential, statistics.

The implication here is that the performance of different groups in the sample can be described, but inferences from the performance of those groups cannot be made about the larger population. Generalisations can be extended to schools and students with the same characteristics as those studied here.

Attribution of the impact of the TIPP

Impact on teachers

As reported by teachers, some were familiar with the LTP as described in the TIPP as it aligned with the professional learning they had received as part of the NUS rollout. In addition, 8 schools were NUS pilot schools (including 4 intervention and 4 control) and 5 had already received training by the LEGO Foundation. Accordingly, intervention schools were not complete beginners regarding LTP. It is impossible to differentiate between the impact of their prior training and the impact of the TIPP.

Impact on students

After the relaunch in September 2022, purpose of the student assessment was no longer to investigate the impact of the LTP intervention on children's literacy and SEL, but rather to compare sampled students' literacy and SEL skills between two time points; prior to and after the invasion of Ukraine by Russian forces. While we present the literacy and social-emotional skills results disaggregated by intervention and control groups in this report, we acknowledge that learning context has changed considerably for children in the study. It is very difficult to reliably compare the results of these groups as we do not know how the invasion impacted the different students and families sampled. Accordingly, these results should be interpreted with caution.

Annex 2 Map of Research Study Questions and Methods (revised September 2022)

RESEARCH QUESTION	SUB QUESTIONS	WHAT DO WE NEED TO KNOW?	INSTRUMENTS	WHAT WILL WE COLLECT?	PARTICIPANTS & APPROX.IMATE NUMBERS	FREQUENCY
RQ1: **What are the barriers and enablers that limit and/or support effective implementation of LTP in intervention school classrooms?**	What barriers (e.g. teacher attitudes and values, training) limit effective implementation of LTP in intervention classrooms?	What are the specific barriers that teachers encounter? What are the specific barriers that school leaders encounter? Do coaches observe barriers?	Teacher interview	Detail on teacher reported barriers	30 teachers max (approx. 700 students)	**Twice**: Pre and post intervention
			School leader interview	Detail on school leader reported barriers	10 School leaders	**Once only,** at baseline
			Pedagogy Partner reflection protocol	Reflections on teacher attitudes and practices that might be barriers	5 Pedagogy Partners working with up to 30 teachers	**Three times in total** during the intervention period
			Teacher reflection protocol	Teacher perceptions of barriers	30 teachers max (approx. 700 students)	**Three times in total** during the intervention period
			Policy and curriculum document review framework	Detail on the education system in Ukraine (e.g. curriculum reform, education policy) to identify potential system level barriers	N/A	Throughout
			School leader questionnaire	Information about the impact of COVID-19 on schools, teachers and students, and families, and the impact of the invasion on schools, teachers, students and families.	20 school leaders	**Once at endline**
			Teacher questionnaire	Teacher attitudes and beliefs about teaching and learning	Max 60 teachers (approx. 1400 students)	**Once at baseline**
	What enablers support effective LTP integration in intervention	What are the specific enablers that teachers encounter? What are the specific	Teacher interview	Detail on teacher perceptions of enablers	30 teachers max (approx. 700 students)	**Twice**: Pre and post intervention
			School leader interview	Detail on school leader perceptions of enablers	10 School leaders	**Once only, baseline**

	classrooms?	enablers that school leaders encounter? Do coaches observe enablers?	Pedagogy Partner reflection protocol	Reflections on teacher attitudes and practices that might be enablers	5 Pedagogy Partners working with up to 30 teachers	**Three times in total**
			Teacher reflection protocol	Teacher perceptions of enablers	30 teachers max (approx. 700 students)	**Three times in total** during the intervention period
			Policy and curriculum document review framework	Detail on the education system in Ukraine (e.g. curriculum reform, education policy) to identify potential system level enablers	N/A	**Throughout**
			Teacher questionnaire	Teacher attitudes and beliefs about teaching and learning	Max 60 teachers (approx. 1400 students)	**Once only, baseline**
RQ2: **How do teachers in intervention schools implement LTP, and adjust their classroom practices, to promote learners' literacy and social and emotional development?**	What practices do teachers demonstrate at the beginning of the study?	Teacher's practices in relation to implementing LTP: What is the starting point?	Teacher interview	Teacher perceptions/self-reported practices	30 teachers max (approx. 700 students)	**Once at endline, once at baseline**
			Teacher reflection protocol	Teacher perceptions/self-reported practices	30 teachers max (approx. 700 students)	**Three times in total** during intervention period
			Pedagogy Partner reflection protocol	Pedagogy Partner perceptions/observed teacher practices	5 Pedagogy Partners working with up to 30 teachers	**Three times in total** during intervention period
			Classroom video review	Researcher observed teacher practices	30 teachers max.	**Three times in total** during intervention period
	What practices do teachers	Do teacher practices in relation to LTP change over time?	Teacher interview	Teacher perceptions/self-reported practices	30 teachers max.	**Once at baseline, once at endline**

	demonstrate during the intervention?		Teacher reflection protocol	Teacher perceptions/self-reported practices	30 teachers max (approx. 700 students)	**Three times in total** during intervention period
			Pedagogy Partner reflection protocol	Pedagogy Partner perceptions/observed teacher practices	5 Pedagogy Partners working with up to 30 teachers.	**Three times in total** during intervention period
			Classroom video review	Researcher observed teacher practices	30 teachers max.	**Three times in total**
RQ3: **How do children's literacy and social emotional skills compare between testing time points including prior to and during the invasion of Ukraine by Russian armed forces?**	When comparing the control and intervention groups, is there a difference between children's literacy, and social and emotional skills including collaboration, at the start of intervention?		Student assessment instrument (designed by ACER)	Literacy development (receptive and expressive oral language) Social-emotional skill development	At baseline: approx. 700 intervention students and approx. 700 control students. At endline, students in Kyiv and Poltava.	**Once at baseline and once at endline**
			Parents and carers questionnaire	Information on family socio economic status to understand how well the intervention and control schools were matched.	At baseline, approximately 1465 parents and carers	**Once only,** at baseline
			School leader questionnaire	Information about the impact of COVID-19 on schools, teachers and students, and families, and the impact of the invasion on schools, teachers, students, and families.	20 school leaders	**Once only,** at endline
	When comparing children learning in different locations in Ukraine, is there a difference between children's literacy and SE skills?		Student assessment instrument (designed by ACER)	Literacy development (receptive and expressive oral language) Social-emotional skills development	At Baseline: approx. 700 intervention students and approx. 700 control students At endline, students in Kyiv and Poltava.	**Once at baseline and once at endline**

Annex 3 LTP-LSSA - Endline Technical Report

Literacy skills described scale

Table 4 shows the percentage of students in each level for baseline and endline for the 296 students who participated in both assessment rounds. Adjustments to percentages were made to equate both rounds, which is why these results differ slightly from the baseline report. However, the distribution across the levels for this small sub-group from baseline still closely matches the distribution for the whole baseline cohort, indicating that when they were at baseline, these 296 students had a similar range of literacy abilities to the original larger cohort.

Here we also describe the levels of literacy skill based on the items located in these levels. The students located in each level can demonstrate the skills in the levels below them. They are still consolidating the skills in their level with a fifty percent probability of demonstrating these skills. A few small changes were made to the skill descriptions since baseline. New descriptions, based on the new items are shown in red font and descriptions that have been moved are shown in *italics.* The reasons for these changes are provided after the table.

Table 4: Literacy (oral language) skills described scale

Level No.	baseline %	endline %	Level name	Literacy (oral language) skills demonstrated
5	**1.7**	**2.7**	**Very High**	*Baseline Level 5 descriptions have been moved to Level 4 based on more reliable endline data for these levels*
4	**10.5**	**36.5**	**High**	Imaginatively generates and extends descriptions of characters and their actions, implying ideas beyond the images, connecting sequences coherently by giving detail about causality and showing flexibility in adapting to changed images. Gives 3 different synonyms for a familiar adjective.
3	**40.5**	**53.7**	**Competent**	Links features in describing characters, includes details of images with some elaboration and establishes clear connections across a sequence of images and variations. Listens to a descriptive text and answers 5 out of 5 questions (more skilled students). Listens to a descriptive text and answers 3 out of 5 questions (less skilled students). Gives 1-2 different synonyms for a familiar adjective.
2	**42.6**	**6.4**	**Basic**	Lists a few disconnected features of characters, gives literal, incomplete descriptions of images making simple, limited connections across a sequence of images. Speech consistently clear. Listens to a descriptive text and locates one piece of literal information from the beginning of the text.
1	**3.4**	**0.3**	**Limited**	Gives simple labels to things or actions without forming sentences, and makes minimal, or ineffective links across a sequence of images.
0	**1.4**	**0.3**	**No skills**	No literacy skills demonstrated

Descriptions in italics indicate skills that have changed levels from baseline to endline.
Descriptions in red font are for new skills introduced at endline.

Changes to the described levels of skill

The endline data was equated onto the original baseline Literacy scale. Ideally, the items and level descriptions should stay in the same locations, with student ability improving. However, there were two changes made to the level descriptions:

- **Level 5 descriptions moved to Level 4**: At baseline there were a tiny percentage (1.9%) of very high achieving students who were estimated to be working in Level 5, but this estimate was unreliable because the percentage of students at this level was so small. The endline data with more students at the highest levels, shows that the skills previously described as Level 5, are more reliably located in Level 4.
- **New item skills added**: Two new literacy items requiring students to generate synonyms were included in the endline assessment. The descriptions of these skills have been added to the levels.

Literacy results comparing intervention and control groups

At baseline, the average literacy scale scores for the control and intervention groups were 52.66 and 47.71, respectively (see Table 5). At endline, the average literacy scale scores for the control and intervention groups were 61.52 and 58.00, respectively. The control group had higher average achievement in literacy compared with the control group at both timepoints. However, the results show that the magnitude of the differences between groups decreased from baseline to endline (intervention group closed the gap slightly), but the effect sizes were very similar.

Table 5: Differences in literacy scores between groups at each timepoint

	N	Mean at baseline	Mean at endline
Control	103	52.66	61.52
Intervention	193	47.71	58.00
Difference (effect size)		4.95 (0.51)	3.52 (0.49)

The intervention group grew on average by 10.29 score points with an effect size of 1.30. This was more than the control group which grew on average by 8.86 score points with an effect size of 0.84. Whilst the intervention group ultimately didn't quite catch up to the control group they are beginning to close the gap (Table 6).

Table 6: Score improvement in Literacy for each group

	N	Mean at baseline	Mean at endline	Mean of Improvement	Std. Deviation of Improvement	Effect size
Control	103	52.66	61.52	8.86	10.52	0.84
Intervention	193	47.71	58.00	10.29	7.89	1.30

Socio-emotional skills described scale

Table 7 shows the percentage of students in each level for baseline and endline for the 296 students who participated in both assessment rounds. Adjustments to percentages were made to equate

both rounds, which is why these results differ slightly from the baseline report. However, the distribution across the levels for this small sub-group from baseline still closely matches the distribution for the whole baseline cohort, indicating that the 296 students have a similar range of social and emotional skills to the original full cohort.

Table 7 describes the levels of social and emotional skill based on the items located in these levels. The students located in each level can demonstrate the skills in the levels below them. They are still consolidating the skills in their level with a fifty percent probability of demonstrating these skills. A few small changes were made to the skill descriptions. New descriptions, based on the new items used in Round 2 (endline) are shown in red font and descriptions that have been moved are shown in *italics*. The reasons for these changes are provided after the table.

Table 7: Socio-emotional skills described scale

Level no.	baseline %	endline %	Level name	Skills demonstrated
6	**2.0**	**14.5**	**Very High**	Shows insight and empathy in evaluating a character's attitudes and behaviour as including both reasonable and unreasonable elements. *Rates a challenging task as very easy for them (was L5)*
5	**6.4**	**37.8**	**High**	Identifies how a proposed intervention is likely to escalate a simple conflict scenario showing understanding of ineffective strategies. Shows understanding of different perspectives by generating a solution that works for both characters.
4	**24.0**	**32.8**		Proposes a collaborative solution to a conflict that focuses on the characters resolving it themselves. Identifies how characters might respond to a proposed intervention showing empathy for familiar, different perspectives.
3	**38.2**	**11.1**	**Competent**	Suggests two appropriate responses that show empathy for a character experiencing sadness or anger and describes two different, plausible feelings a character might have when disappointed. Describes two different strategies to self-regulate their feelings of sadness. *Describes two different strategies to self-regulate their feelings of anger (was L4)* Provides two strategies to resolve a simple conflict. Values persistence in the face of disappointment. Recognises when an accusation is unfair. Evaluates a character's response to a conflict scenario from one perspective only. Proposes a solution to a conflict that focuses on an external solution. Gives a reason to support why a challenging task was easy for them. Expresses a highly positive attitude to inventing character descriptions. Engagement with story tasks and conflict scenario is rated as enthusiastic.
2	**23**	**3.0**		Describes one strategy to self-regulate their feelings of anger and sadness. Identifies the perspective of both characters in a simple conflict scenario but suggests a one-sided solution or suggests others who might help. Considers one perspective only when rating an accusation as fair or not.

				Shows limited empathy with one appropriate response only to an angry character and identifies one plausible feeling only for a disappointed character. Rates a challenging task as easy for themselves with no reason. Expresses a highly positive attitude to story sequence tasks and solving a conflict. Is slightly positive about a challenge.
1	**5.4**	**0.3**	**Low** **Very low**	*Identifies what might cause themself to feel angry* (was L2). Identifies what might cause themself to feel sad with no self-regulation strategy. Identifies one explicitly stated familiar perspective only in a simple conflict scenario. Shows limited empathy with one appropriate response only to a sad character. Expresses a slightly positive attitude to story tasks & solving conflicts.
0	**1.0**	0.3		Identifies what might cause themself to feel happy. *Engagement with story tasks and conflict scenario is rated as compliant (was L2).* Expresses a slightly negative attitude to the story tasks, dealing with a challenge or solving a problem

Descriptions in italics indicate skills that have changed levels from baseline to endline
Descriptions in red font are for skills introduced at endline

Changes to the described levels of skill

The endline data were equated onto the original baseline Social-emotional scale. Ideally, the items and level descriptions should stay in the same locations, with student ability improving. However, there were two changes:

- A few descriptions changed levels
- New item skills added

A few descriptions were moved. A few items had a substantial change in their scale location from baseline to endline, so the descriptions of these skills were moved to a new level. The relative difficulty of items on a scale should be stable so items stay in the same level. Some movement in the item scale locations in a longitudinal study is to be expected, especially when a scale such as the social and emotional scale combines a wide range of skills.

Some new items added. Five new items were added to the assessment. These were not as difficult as had been hoped. Any skills addressed by these items that were not included in the original level descriptions have been added in red font.

Social-emotional results comparing intervention and control groups

At baseline, the average social-emotional skills scale scores for the control and intervention groups were 52.56 and 47.40, respectively (see Table 8). At endline, the average social-emotional skills scale scores for the control and intervention groups were 63.38 and 59.66, respectively. The control group had higher average achievement in social-emotional skills compared with the intervention group at both timepoints. However, the results show that the magnitude of the differences between groups decreased from baseline to endline. In addition, the effect size difference from baseline to endline also indicate that there was a noteworthy closing of the gap in social-emotional skills results between groups.

Table 8: Differences in social-emotional skill scores between groups at each timepoint

	N	Mean at baseline	Mean at endline
Control	103	52.56	63.38
Intervention	193	47.40	59.66
Difference (effect size)		5.16 (0.57)	3.72 (0.42)

The intervention group grew on average by 12.27 score points with an effect size of 1.36. This was more than the control group which grew on average by 10.82 score points with an effect size of 1.09. Whilst the intervention group ultimately didn't quite catch up to the control group they were beginning to close the gap (see Table 9).

Table 9: Score improvement in social-emotional skills for each group

	N	Mean at baseline	Mean at endline	Mean of Improvement	Std. Deviation of Improvement	Effect size
Control	103	52.56	63.38	10.82	9.93	1.09
Intervention	193	47.40	59.66	12.27	9.00	1.36

Psychometric analysis

To directly compare the results of students from baseline to endline, student estimates of ability need to be placed on a common metric. This is done using a method called test equating. To do this, new test forms (endline) needed to be equated to the original test forms (baseline). Typically, the most used design in IRT Rasch equating is the common item method, which assumes that there is item parameter invariance across different test forms. Parameter invariance is a property of IRT models in which parameter estimates remain invariant irrespective of the sample from the population used to obtain the item estimates and, conversely, student ability estimates are not influenced by the selection (sample) of test items that are used to generate the ability estimates. A key requirement for equating and linking using the common item method is that the statistical properties of the common items remain stable across different populations and test forms. In particular, the below paragraph describes the equating process applied for Literacy and Socio-Emotional test forms from endline to baseline.

First, items at endline were calibrated separately, where two unidimensional 1 parameter logistic (1PL) partial credit models (PCMs; Masters, 1982) for literacy and social-emotional skills separately were fit to the data using ACER ConQuest version 5 (Adams et al., 2020). The item difficulty estimates of the common items were compared with their estimates from baseline calibrations. Any items that showed substantial differential item functioning (DIF) between the two calibrations (difference in relative difficulty greater than 0.5 logits and statistical significance at 0.05) were not used to produce the equating shift. There were 11 common items in Literacy and 17 common items in Socio-Emotional between baseline and endline. The process described above resulted in selection of the final common item set which consisted of 9 Literacy items and 13 Socio-Emotional items for computing the test equating shift (average of baseline item difficulty estimates – average of endline

item difficulty estimates). The test equating shift was 0.427 and 0.423 (logits) for Literacy and Socio-Emotional, respectively. Item difficulty estimates and student ability estimates (in logits) from the calibrations at endline were adjusted by the corresponding equating shift so they were then placed on the previous latent scales from baseline.

Using ConQuest and the conquest library (Cloney & Adams, 2021) in R (R Core Team, 2020), test- and item-level statistics are provided below to illustrate how well the assessment functions as a whole (e.g. test reliability) and whether all of the items included in the assessment are adequate in terms of facility, item fit, item functioning and targeting.

Psychometric results

Test reliability

Estimates of expected a-posteriori/plausible value (EAP/PV) reliability are used to assess the reliability of the literacy and social-emotional skills scales. EAP/PV estimates range from zero to one, with values above 0.8 indicating good reliability. The EAP/PV reliability estimates for the literacy and social-emotional skills scales are 0.835 and 0.788, respectively.

Item facility

Item facility (percent correct) provides an indication of whether items were either too easy (>90%) or too difficult (<10%) for the test takers. For partial credit items, facility refers to the percent of the sample who were assigned the possible scores (response categories) for each item.

All of the literacy items which have five response categories (LTP001, LTP002, LTP003, LTP005, LTP006 and LTP008) had the two lowest scored response categories with less than 10% of the test-takers assigned to them. The lowest scoring category for LTP027 also had less than 10% of test takers assigned to it. Twelve (out of 25) social-emotional skills items (LTP004, LTP018, LTP019, LTP021, LTP024, LTP025, LTP035, LTP036, LTP037, LTP038, LTP042 and LTP044) had at least one response category with less than 10% of test-takers assigned to them, most of which was the lowest category.

Item-rest correlation

Item-rest correlations can be used to indicate how well the vector of item responses for each individual item is correlated with the sum of item responses across the rest of the items included in the assessment. In general, item-rest values >0.2 are considered adequate. All but one literacy (LTP030) and two social-emotional skills (LTP009 and 0.18) items were marginally (0.18) outside the generally accepted adequate range (see Table 10).

Item fit

Item fit statistics (weighted infit mean square (MNSQ: Wu, 1997)) are used to determine the level to which the model fits the data. In general, MNSQ values >1.2 are considered poorly fitting (and under discriminating) items. All but one social-emotional skills item (LTP037) fit within the generally accepted range (see Table 10). LTP037 was only marginally outside this range (1.21).

Table 10: Item-rest correlation estimates for literacy and social-emotional skills

Domain	Item	Item-Rest	Item fit (MNSQ)
Literacy	LTP001	0.57	0.96
	LTP002	0.54	0.98
	LTP003	0.37	1.03
	LTP005	0.67	0.79
	LTP006	0.68	0.81

	LTP008	0.58	0.95
	LTP027	0.35	1.04
	LTP028	0.33	1.05
	LTP029	0.24	1.13
	LTP030	0.18	1.18
	LTP031	0.37	1.00
	LTP045	0.52	1.03
	LTP046	0.48	1.09
Socio-emotional	LTP004	0.42	0.95
	LTP007	0.21	1.08
	LTP009	0.18	1.13
	LTP010	0.41	0.94
	LTP018	0.22	0.97
	LTP019	0.27	0.95
	LTP020	0.43	0.91
	LTP021	0.28	0.94
	LTP022	0.25	1.02
	LTP023	0.34	0.99
	LTP024	0.33	0.86
	LTP025	0.27	0.94
	LTP026	0.31	1.00
	LTP032	0.25	0.99
	LTP033	0.23	1.00
	LTP034	0.43	0.95
	LTP035	0.42	1.05
	LTP036	0.40	1.11
	LTP037	0.18	1.21
	LTP038	0.55	0.85
	LTP040	0.27	1.01
	LTP041	0.37	0.99
	LTP042	0.43	0.99
	LTP043	0.52	0.89
	LTP044	0.53	1.05

Test targeting

Item-person maps are a useful way of visually examining the targeting of an assessment by showing the alignment of items (in terms of item difficulty) and persons (in terms of ability) on the same scale. A well-targeted test is one where the distribution of child abilities aligns well with the items in the test. If the distribution of child abilities is higher on the scale than the majority of items then the test is generally considered too easy. Conversely, if the distribution of child abilities is lower on the scale than most items then the test could be considered too difficult. The literacy item thresholds provide generally adequate coverage of the span of the distribution of abilities for this sample of children. There are a few item thresholds located in the upper half which provide better targeted opportunities for higher ability children to illustrate their skills. For social-emotional skills there are more limited opportunities for higher ability children to illustrate their skills with fewer item thresholds located towards the upper half of the ability distribution. There are also several item thresholds located at or below the lowest ability children.

Figure 32: Map of literacy item difficulty and student ability on the same scale

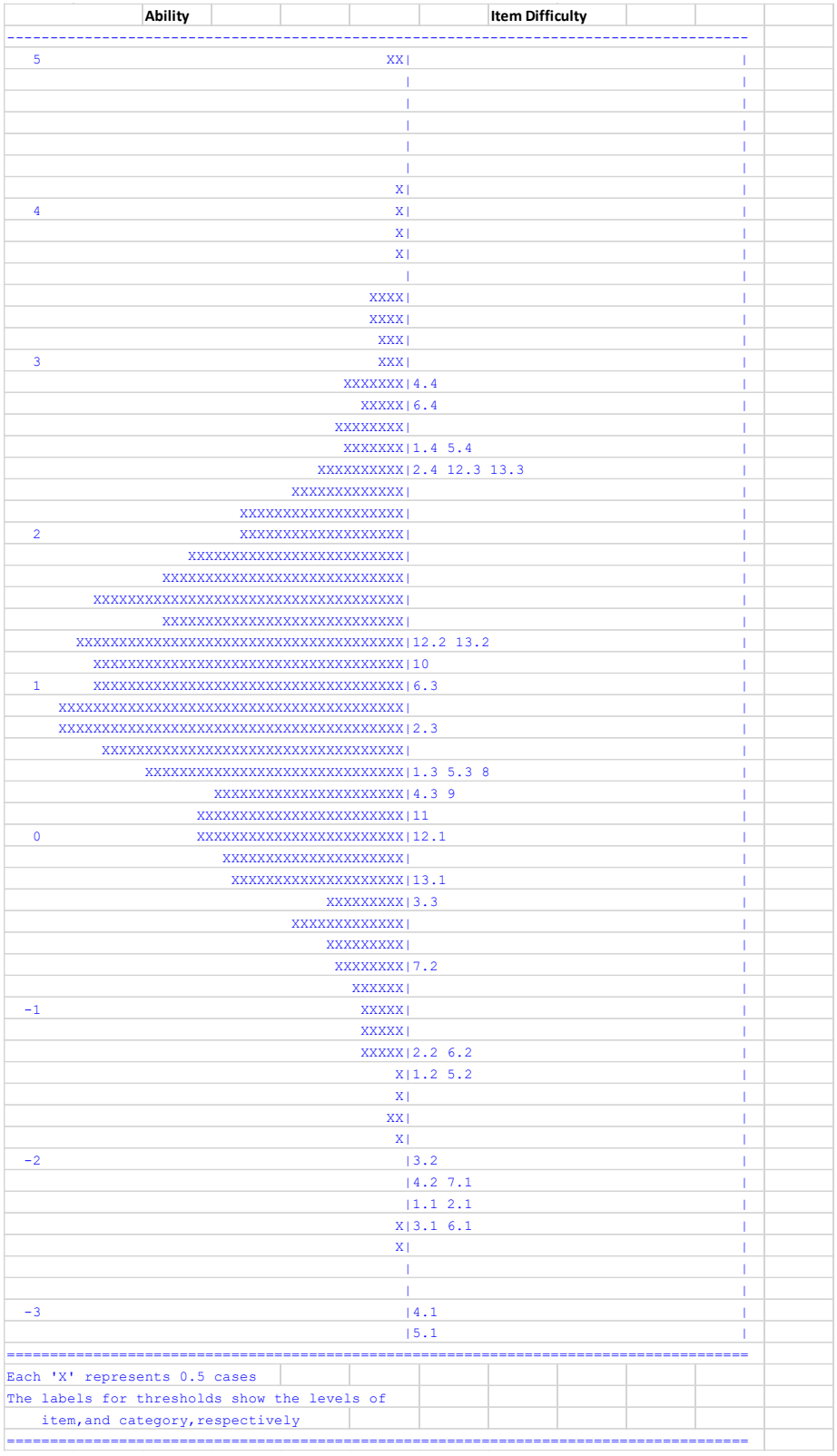
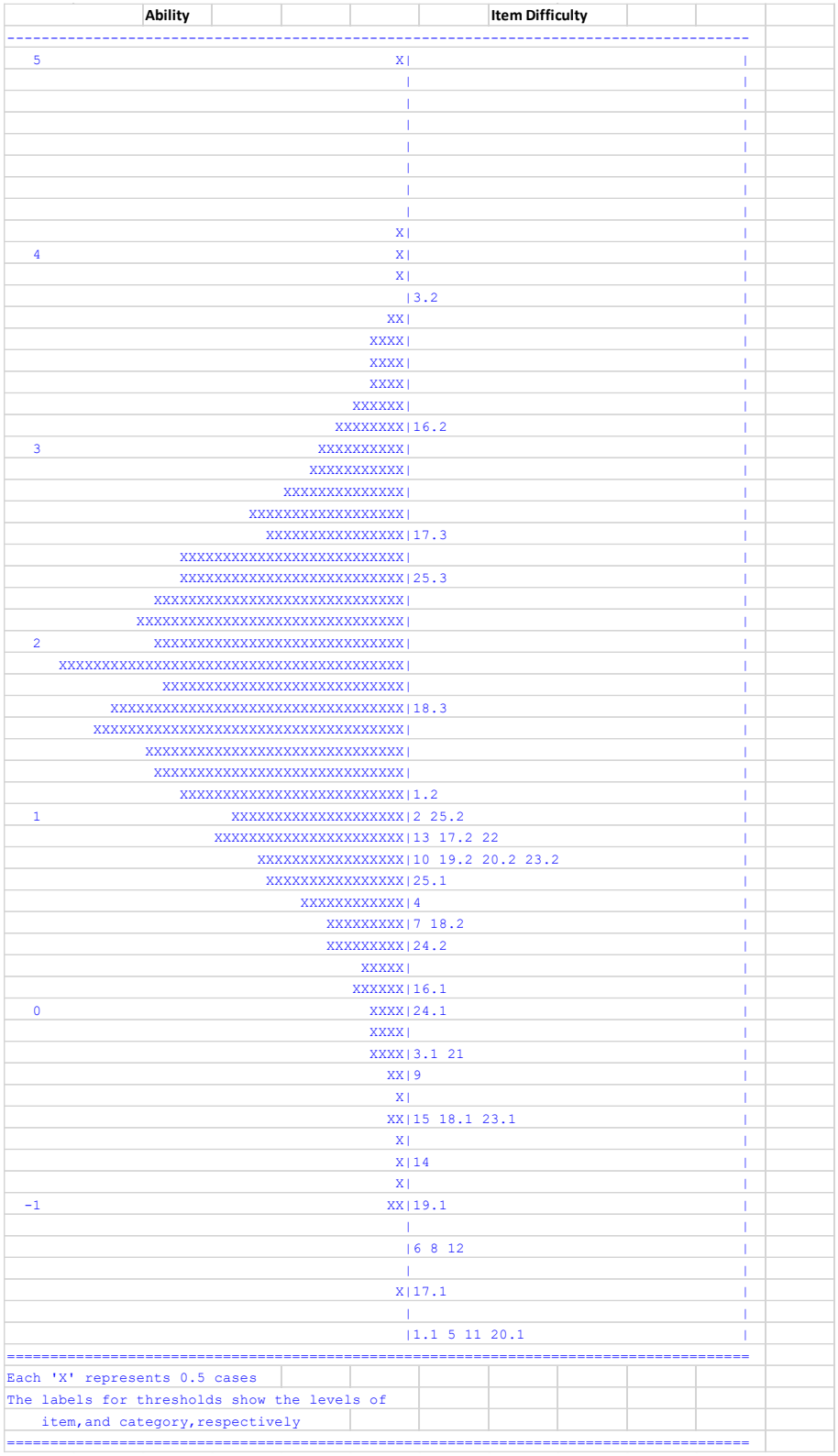


Figure 33: Map of social-emotional skills item difficulty and student ability on the same scale



Differential item functioning (DIF)

Differential item functioning (DIF) is typically a result of items functioning differently (i.e., more or less difficult) for one subgroup of the population compared with other(s), holding ability constant (Boone et al., 2014), resulting in measurement bias. DIF analyses were undertaken for gender and intervention group membership (i.e., intervention vs control). To examine DIF between these

comparison groups separate calibrations were run for each separately. Then, they were compared to determine if the magnitude of the differences were sufficiently large enough to be considered to be exhibiting DIF. The criteria used for this was a difference between item difficulty estimates >0.5 and statistical significance at the 0.05 level. Scatter plots were also drawn to illustrate the magnitude of the difference between the sub-groups being compared. When no DIF is present, the data points will fall along the unit line, whereas increasingly large differences will fall further from the line.

Two literacy items (LTP028 and LTP030) and five social-emotional items (LTP018, LTP024 LTP032, LTP038 and LTP043) were shown to exhibit gender DIF (also see Figure 34 and Figure 35). One literacy item (LTP027) and six social-emotional items (LTP007, LTP018 LTP020, LTP022, LTP025 and LTP041) were shown to exhibit intervention group membership DIF (see Figure 36 and Figure 37).

Whilst some DIF was present, no item treatments were applied. Future work should explore the presence of DIF further (inc. other sub-groups (e.g., SES)) and determine whether the issues are related to language/translation, field administration or item fit. If significant levels of DIF are shown steps can be taken to account for it, such as deleting items or freeing the parameters for different sub-groups of the population.

Figure 34: Gender DIF plot for literacy

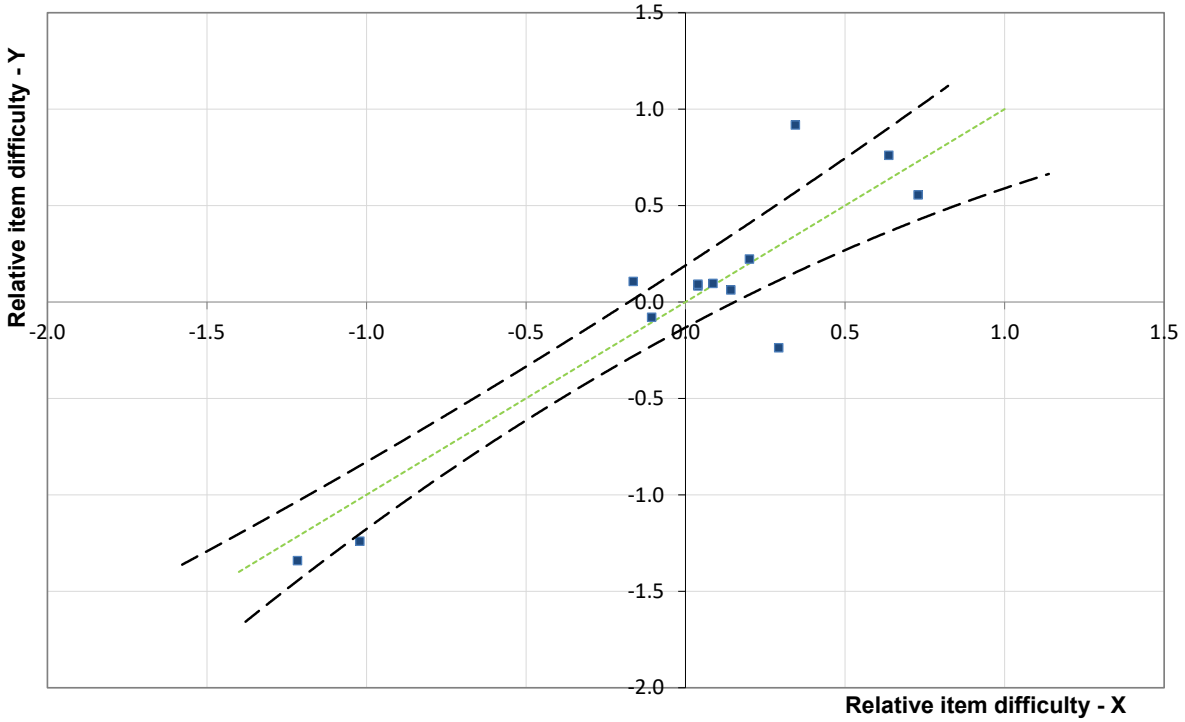


Figure 35: Gender DIF plot for social-emotional skills

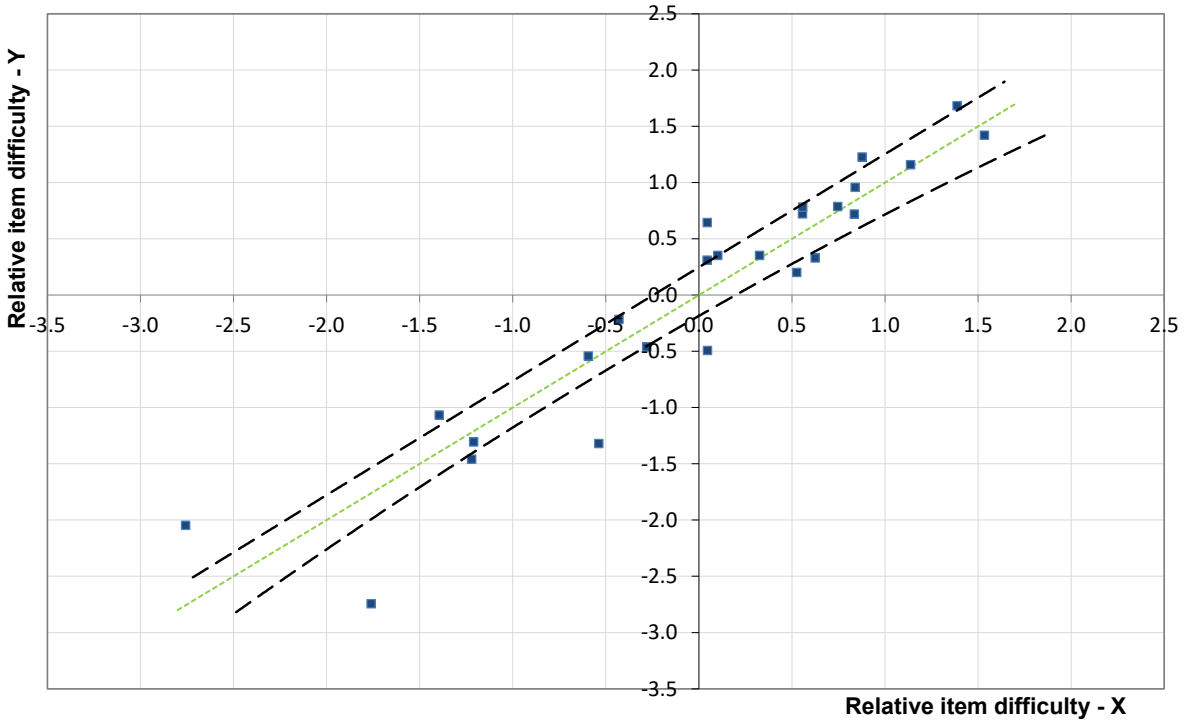


Figure 36: Intervention group membership DIF plot for literacy

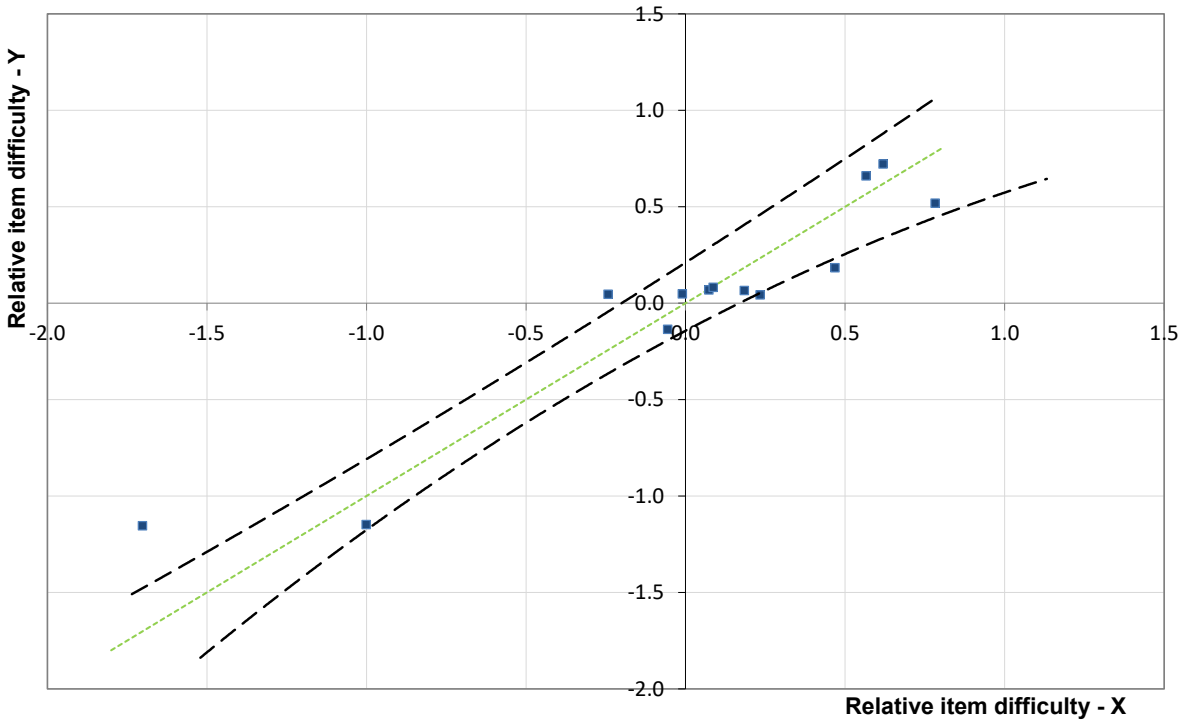
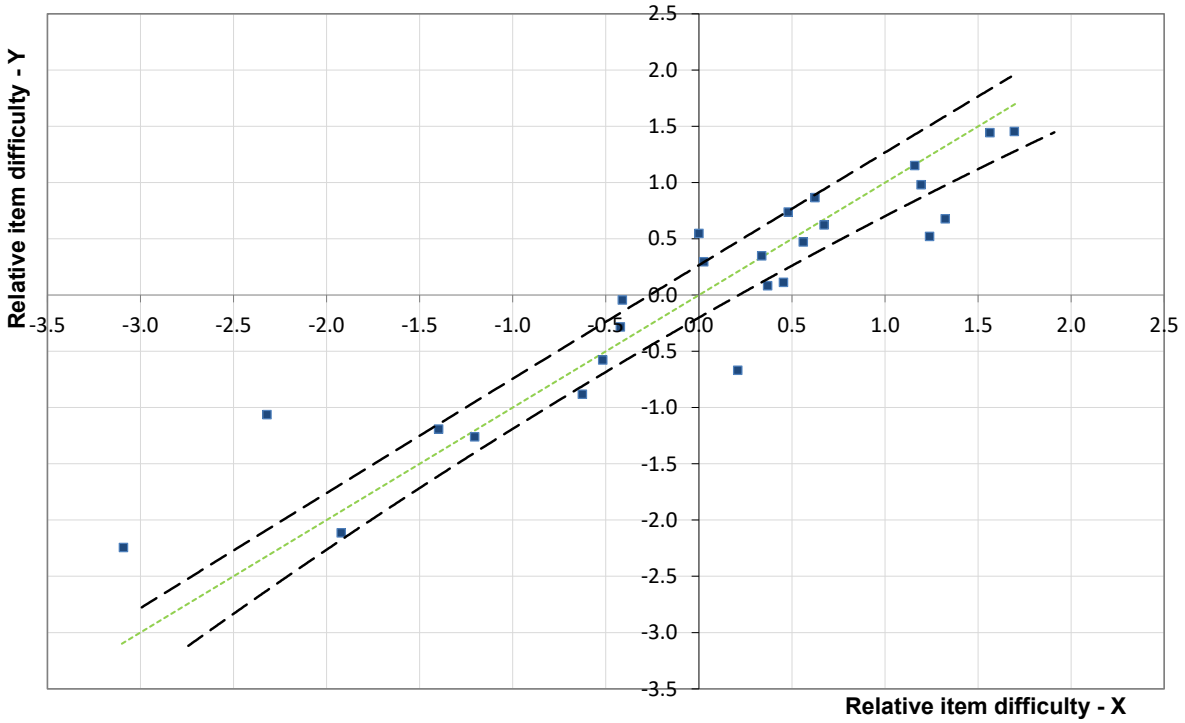


Figure 37: Intervention group membership DIF plot for social-emotional skills



Longitudinal results

The endline results are a longitudinal comparison of the same students who also participated at baseline. This was a group of 296 students composed of 193 intervention students and 103 control students. Baseline was administered when students were in Grade 1 and by endline they were in Grade 3. Sadly, participation at endline by most of the original cohort of students was prevented by the invasion of Ukraine by Russian armed forces. This means that the students who participated at baseline only, were omitted from the analyses presented in this report.

Reliable interpretations of the results for the intervention group compared with the control group are compromised by the small number in each group, particularly the control group with only 103 students. The results suggest that on average the intervention group improved more than the control group from baseline to endline for both literacy and social-emotional skills. The intervention group started at a lower point and did not catch up to the control group, but it did start to close the gap. While these results indicate a trend, they are not statistically reliable. Small numbers mean the comparison is less likely to be fair. The results are easily influenced by possible greater advantages experienced by one group compared with the other such as more opportunities to learn, fewer disruptions, more supportive school culture or fewer low ability students that larger numbers would ameliorate. Larger numbers make it more likely both groups would have a similar share of advantages and disadvantages in terms of supporting learning.

It is also noted that the intervention group had 15 students likely to require extensive learning support compared with only one student in the control group. Twelve of the intervention students likely to require support had limited language experience and two had a functional disability. Each group had one student with cognitive, behavioural, or emotional Special Education Needs. Managing these students places greater demands on the teacher which may also limit the learning opportunities for the other students in the intervention group.

www.acer.org

www.legofoundation.com

www.uera.org.ua