



2023 LEAP CHALLENGE

LEAP Final Deliverables

Project Host:

Aflatoun International



Fellows:

Luca Maria Pesando, Research Fellow

Martin Tomasik, Research Fellow

Simi Lawoyin, Social Entrepreneur Fellow

TABLE OF CONTENTS

EXECUTIVE SUMMARY	4
Introduction: Aflatot Program	4
Organization’s role & strength	4
Need summary	5
REPORT ON EXISTING LITERATURE AND STATE-OF-THE-ART FOR AFLATOUN	7
Introduction	7
Section 1: The role of social and financial literacy in people’s lives	7
Benefits of financial literacy in people’s lives	7
Role of financial literacy in children’s lives	8
Measuring Financial Literacy in Children	10
Associations between Financial Literacy and Other Outcomes	11
Role of families vs peers vs schools	13
Section 2: Do financial education interventions affect social and financial literacy?	14
Effectiveness of financial education interventions	14
Socio-demographic considerations	14
Benefits of financial education interventions	15
Effects of parental involvement	16
Section 3: The Aflatot program and related programs	17
Existing evidence on the impacts of Aflatoun or Aflateen	17
Existing evidence on the impacts of Aflatot (3-6)	19
Exploring other similar programs	22
What can we learn from other programs?	25
Section 4: Recommendations and ideas	28
Testing outcomes for children	28
Parental Involvement	28
Evaluation tools	29
EVALUATIVE RESEARCH	33
Introduction	33
Formative and Summative Assessments	33
Distinguishing Proximal and Distal Outcomes	36
Framing an Evaluative Research Approach for Aflatot: Recommendations	37
Potential Gaps and Opportunities for Further Research	40
Randomized Control Trial (RCT) Design Recommendations	41
Control Group Design	42

Selecting the Right Sample Size	43
Challenges of Designing a RCT for Young Children	44
Framing a RCT Design for Aflatot: Considerations and Recommendations	45
Methodological Considerations	45
Resource Considerations	46
Ethical Considerations	46
Sample RCT Design	48
Concluding Remarks	57
Appendix A	58
References	61

EXECUTIVE SUMMARY

Introduction: Aflatot Program

Aflatot is an early childhood education program for sustainable development, developed by Aflatoun International specifically targeting preschool children. The program aims to develop age-relevant social-emotional skills and financial awareness and features training for facilitators to deliver the program through an active learning methodology.

The Aflatot curriculum is designed for children aged three to six years old and aims to lay the foundation for social and financial literacy before they reach primary school. By balancing social learning and financial concepts, Aflatot empowers children to believe in themselves, understand their rights and responsibilities, practice saving and spending, and even start their own enterprises.

The program includes a curriculum with detailed lesson plans or activities designed for educational settings such as schools, nurseries, and community centers. The program has a duration of 45 hours. In addition to the curriculum, Aflatot provides a family toolkit that offers sessions facilitated by educators, engaging parents with their children to learn together about social and financial content. The toolkit also suggests activities that can be done at home.

Organization's role & strength

Aflatoun International has 17 years of experience in life skills and financial education programming worldwide. The organization's mission is to ensure access to quality, inclusive, child-centered social and financial education for all children and young people, especially the most vulnerable. Aflatoun's overall goal is to socially and economically empower and provide children and young people with the skills, knowledge, and resources necessary to participate and thrive in their communities. Aflatoun, as a social franchise, currently works with and provides technical expertise to a partner network of 345 organisations (including NGOs, CSOs, and governments) that implement social and financial education programmes in more than 100 countries across different regions.

Need summary

Aflatoun was seeking to understand the following in relation to the Aflatot program:

- **Need 1:** What is the short-term and long-term impact of the Aflatot programme on the socio-emotional development, financial literacy, and behaviour of young children, particularly in terms of their self-regulation, delayed gratification, decision-making skills, and understanding of money? How can we further improve the programme's content and delivery?
- **Need 2:** What is the current state of social and financial skills development in early childhood education? What is the impact of social and financial skills development programs targeted at preschool-age children in early childhood education, and what are the research gaps in understanding their effectiveness?

Toward addressing the Project Host's needs, we developed two deliverables:

- **Deliverable 1:** A state-of-the-art literature review and report specifically addressing the questions posed.
- **Deliverable 2:** A research feasibility report to evaluate the implementation methodologies, necessary resources, and design strategies for conducting an effective Randomized Controlled Trial (RCT) on preschool-age children's social and financial skills development.

DELIVERABLE 1:

REPORT ON EXISTING LITERATURE

REPORT ON EXISTING LITERATURE AND STATE-OF-THE-ART FOR AFLATOUN

Introduction

This report provides a state-of-the-art review of the existing literature for financial literacy programs, with a focus on those particularly related to Aflatot—an initiative by Aflatoun promoting social and financial literacy for young children. The report is structured into four broad sections: the first discusses the impact and benefits of financial literacy across different stages of life; the second considers the effectiveness of educational interventions in fostering financial literacy; and the third examines the outcomes associated with the Aflatoun program and compares these with other similar initiatives. In the fourth section, we draw from insights synthesized from the literature to offer recommendations for optimizing financial literacy programs to better serve young children, and to highlight areas where further research is needed.

Section 1: The role of social and financial literacy in people’s lives

Benefits of financial literacy in people’s lives

The role of financial literacy

Financial literacy, referring to knowledge of financial concepts and procedures, has been widely studied in relation to its role and impacts in the lives of all ages. Among adults, extensive literature (beyond the scope of this report) shows that financial literacy is correlated with higher wealth, better financial planning, better retirement preparedness, stock-market participation, and other favorable outcomes. Literature suggests that financial literacy among youth is generally low (Chen & Volpe, 1998; Erner et al., 2016; Lusardi, 2015; Lusardi et al., 2010; Mandell, 2008).

Clarifying the Terminology

“Financial literacy” typically refers to knowledge of financial concepts and procedures, whereas “financial capability” has been utilized to indicate the skills to apply this knowledge, and “financial inclusion” to indicate the opportunity to do so. These potentially different aspects are often conflated with each other.

- *Financial Literacy*: Understanding of financial concepts and procedures.
- *Financial Capability*: Skills to apply financial knowledge.
- *Financial Inclusion*: Opportunity to utilize financial knowledge and skills.

Impact of early financial literacy on long-term outcomes

The literature also explores the long-term impacts and societal benefits of financial literacy, highlighting how early financial decisions influence future financial stability and equality:

- A solid body of research has shown that poor financial decisions early in life might have serious consequences for later-life outcomes (Brown et al., 2016; Xiao et al., 2014). Among others, better financial literacy is associated with positive outcomes such as lower debt load (Lusardi & Tufano, 2015), wider portfolio diversification (Guiso & Jappelli, 2009), active participation in the stock market (van Rooij et al., 2012), higher wealth accumulation (van Rooij et al., 2012), and better retirement planning (Arrondel et al., 2013; Fornero & Monticone, 2011; Lusardi & Mitchell, 2007, 2011b, 2011a).
- Inequality in financial literacy early in life may contribute to inequality in financial assets and behaviours later in life. Due to the far-reaching potential of targeting individuals in schools before they enter the labour market, strengthening financial literacy in schools and promoting financial-education interventions has often been advocated as one sensible strategy.

Role of financial literacy in children’s lives

- McCormick (2009) and Totenhagen et al. (2015) provide two big-picture analyses focused on the importance of financial literacy and financial education among youth. McCormick (2009) claims that financial literacy among youth is essential to help children deal with the struggles of their families and familiarize them with adult financial tasks and responsibilities, especially if they don’t move on to postsecondary education.
 - Children throughout the K–12 grades, including children who differ in ability levels and socioeconomic backgrounds, can learn worthwhile content in personal finance if their teachers use appropriate strategies and materials.

- Children’s understanding of economics and personal finance develops through a series of levels or stages. Nothing about the subject matter *per se* makes personal finance inappropriate for study by children in the early grades.
- Literacy, as the foundation for virtually all other subject areas, needs to be taught from the very earliest ages; this focus on early childhood literacy is known as *emergent literacy*. A Networks Financial Institute report (2006) contends that the core concepts that undergird financial literacy, such as goal setting, intertemporal choice, philanthropic giving, earning, saving, and spending, need to be emphasized heavily and supported from the very earliest grades (McCormick & Godsted, 2006).
- An interview study with 103 first-grade children in Israel attempting to understand the relationship between involvement in saving money (ISM) and finance-related outcomes such as attitudes towards money and access to money (Te’eni-Harari, 2016). The study stresses the key relevance of “involvement” variables for understanding financial literacy outcomes. The author finds that children with a high level of ISM expressed more positive attitudes toward saving, as well as more positive behavior regarding saving. With respect to the factors that might explain the child’s level of ISM, the perceived importance of parental and peer attitudes toward saving money and the child’s access to money played a significant role in determining the child’s level of ISM.
- Gnatuk and Granovsky (2014) suggest that at five years of age children are generally able to understand the concepts summarised here:
 - Saving and Spending:
 - Money has value because you can exchange, trade, or spend it
 - Putting money aside for later spending is called saving
 - You can keep track of money by putting it somewhere safe
 - Saving money is a way to get something you want instead of hoping it will be given to you
 - Shopping:
 - You can make decisions about how and when to spend money before you go to the shops
 - You do not have to spend all your money
 - Good shoppers compare prices and choose items that cost less
 - Working for pay:
 - Different jobs pay different amounts of money
 - Value:
 - The amount of money needed to buy something is called its price
 - Different items have different prices
 - Coins and paper money have different values
 - Some things that seem to be free such as food, really cost money

- On the other hand, the same authors caution that three to five-year-old children are likely to have difficulty understanding planning for the future, or where money comes from.

Measuring Financial Literacy in Children

How is financial literacy best measured among children (e.g. through which proxies)?

- Measurement of financial literacy among children varies extensively by age. Usually, formal financial “tests” are not implemented before the age of 6.
- Research suggests that picture books are an effective and engaging tool for assessing mathematical concepts in young children.
- Van den Heuvel-Panhuizen et al. (2009) conducted a study in the Netherlands with five and six-year-old children to evaluate how picture books can facilitate the comprehension of mathematical concepts in young children. They found that these books were effective in stimulating mathematical thinking by providing a meaningful context for learning mathematics, in a way that traditional lessons alone may struggle to deliver.
- Some recent research stresses that comic strips may be even more effective than picture books, especially at the age of 4 when children are only starting to recognize letters (Chu & Toh, 2020). Comic strips typically require little to no text for the story to develop, allowing the material to be accessible to all children, regardless of reading ability.
- A review by Jay et al. (2022) suggests that it is important to identify a focused set of concepts to include within any measure of financial literacy. When a measure includes too many different concepts, it will be difficult to assess them all within a single testing session as children’s attention spans will impose time limits. On the other hand, assessing too few concepts bears the danger of not allowing for a comprehensive assessment of children’s understanding of the subject matter.
- Jay et al. (2022) discuss the effectiveness of picture and comic books for 4-6 year olds
 - Very effective at stimulating mathematical thinking in context
 - Very effective at showcasing basic geometry concepts
 - Many experimental interventions are designed such that a group of students is exposed to the teacher reading one such books for about one hour per week and the other is not.

Measuring Financial Literacy in Children

The following study is specifically devoted to exploring different ways of measuring financial literacy for different kids of various ages:
<https://maps.org.uk/en/publications/research/2022/measuring-financial-literacy-of-children-age-d-4-to-6-years-design-and-small-scale-test#>

For a full reference on how to measure financial literacy among kids 4-6, see Jay et al. (2022).

- Another effective measure could be outlining a story with a character engaged in daily tasks and chores, giving and receiving money, making decisions about what to spend, etc.
 - Children could be tested in terms of their understanding of various dynamics ongoing in the story, or even by eliciting their suggestions and recommendations in terms of what the story's character should and should not do in various situations. A scoring method can then be devised and implemented through 20-30 mins interviews with children:
 - In some interventions, the scoring method has been designed to be as easy as possible to administer for an interviewer during the interview, while also providing sufficient detail to differentiate between children's levels of financial literacy. For most questions asked, answers were scored on a three-point scale from zero to two points. A score of zero indicates no or limited understanding of a concept, which may mean the child is not familiar with, or shows low levels of knowledge surrounding basic aspects of the concept. For example, a child may receive a score of zero on a question if they merely state that they don't know how to respond/what they would do or, for more objective questions, provide the wrong answer. A score of one indicates some or emerging reasoning of the concept, which may mean the child has some familiarity with the concept but struggles with some underlying mechanisms of that concept, e.g., if they provide the correct answer but struggle to explain why, or if they are able to reason with some key aspects of a concept but not others. For instance, if they rightfully acknowledge that notes have more value than coins but struggle to explain why, or if they understand that they must return something that they have borrowed yet struggle to understand that it does not belong to them. Finally, a score of two indicates good or developed understanding of the concept, where the child illustrates a high level of knowledge and holistic understanding of the concept. The three-point scale allows the interviewer to make a quick decision and to record a score without interrupting the flow of the interview and the associated story.

Associations between Financial Literacy and Other Outcomes

Research explores associations between financial literacy and children's socio-developmental outcomes such as self-regulation, delayed gratification, decision-making skills, executive function, etc.

- The educational and developmental psychology literatures suggest that fostering the development of children's executive functioning before they start school;

promoting beliefs around savings, budgeting, and planning for the future in elementary school and middle school; and nurturing in them self-efficacy and growth mindsets with respect to financial knowledge acquisition through experiential learning in high school are promising approaches to raising children to be financially capable adults (Drever & Else-Quest, 2021) .

- Research is lacking on whether financial literacy is conducive to better self-regulation, better decision-making skills, etc. OR whether these are ways through which financial literacy can be measured among children. The second hypothesis seems to be more heavily supported.
- Drever et al. (2015) and Scheinholtz et al. (2011) review literature from sociology, consumer science, and psychology, uncovering the role of executive function, financial socialization, and experience-based learning in childhood and youth.

Does financial literacy matter more for children’s attitudes or behaviors (or both)?

- Though the evidence is inconclusive, a study on high school students suggests that higher financial literacy is associated with less absenteeism (i.e., behavior), while attitudes towards the importance of schooling are unchanged (Pesando, 2018).
- Summary of evidence from interventions at earlier ages seems to suggest, instead, that attitudes and knowledge are more malleable among young kids, rather than behavior itself (Amagir et al. 2018).
- In terms of financial behavior of students, Amagir et al. (2018) report: “Studies on the impact of financial education programs on student’s financial behavior show diverse results. One out of two studies found positive effects in the intention to save, and five out of six studies found positive effects in self-reported behavior. On the contrary, no effects were found in assessed financial behavior.”

Does financial literacy matter for children’s school readiness, i.e., is there evidence that it makes children “more” school ready?

- There is limited evidence at the level of KG or primary school, but a study focused on PISA data suggests that higher financial literacy increases students’ perceived value of schooling by boosting their time commitment to education (Pesando, 2018).
- Most of the existing literature relates, instead, financial literacy and job-readiness (Cedeño et al., 2021).
- Given the very high correlation between financial literacy and numeracy (Pesando, 2018; Skagerlund et al., 2018) – and numeracy being one component of school readiness – there is reason to believe that higher financial literacy may correlate positively with school readiness.

Role of families vs peers vs schools

With regard to the respective roles of families, peers, and schools (or broader institutions) in providing financial literacy, and which yields the most effective outcomes, the literature suggests:

- Broad reviews on socialization (parents vs peers vs schools) tied to financial literacy can be found in Jorgensen & Savla (2010) and Van Campenhout (2015). These studies explore how each agent contributes to the financial socialization of children. All in all, evidence suggests that schools are more effective than family in providing social and financial literacy, especially at relatively young ages.
- While there is limited research on the genetic determinants of financial literacy, existing scholarship suggests that parental engagement and students' experiences with money matters play a far stronger role than genetic determinants (Moreno-Herrero et al., 2018).
- Just one study at much later stages has identified that 33% of the variation in savings behavior across individuals can be explained by genetic differences (Cronqvist & Siegel, 2015).
- Focusing again on adults, Xu et al. (2017) note that genetic influences of financial distress are highest at the extremes of socioeconomic status (SES). This observation can be partially explained by the fact that neuroticism and cognitive ability are more important predictors of financial distress at low and high levels of SES, respectively.

Opportunities for Further Research

The research highlights the need for integrating financial literacy into school curricula from an early age. The literature also seems to point to the significant role that structured educational settings play in imparting financial literacy, particularly highlighting the effectiveness of schools over family environments, especially at younger ages. Further worthwhile research might include:

- Investigating which school-based structures and educational strategies are most effective for financial literacy.
- Conducting long-term studies to compare the outcomes of financial education in schools versus that received at home.
- Investigating how genetic traits and socioeconomic factors influence financial behaviors.
- Exploring how financial literacy interventions can be customized for different socioeconomic groups to maximize effectiveness.

Section 2: Do financial education interventions affect social and financial literacy?

Effectiveness of financial education interventions

At what age are these kinds of interventions most effective?

- Amagir et al. (2018) provide one of the most useful references in the area of financial education programs for children and adolescents. The reason is twofold: first, this is a systematic review, hence, a summary of a range of papers (60, to be precise); second, the paper not only discusses benefits for children and adolescents, but also reflects on which design and curriculum-related features may be most beneficial. A similar effort focused on adults was conducted by (Fernandes et al., 2014)
 - Overall evidence suggests that school-based financial-education programs can improve children's and adolescents' financial knowledge and attitudes;
 - Most studies evaluate impacts on self-reported attitudes and behaviors, or financial knowledge by means of tests; the few which focus on actual financial behavior show virtually no effect;
 - Focus here is on elementary, middle, high school and college;
 - Majority of studies assessing impact of financial education in primary schools find positive impacts on financial knowledge and attitudes, with no differences by whether the program is integrated into the curriculum or not. In this respect, Berry et al. (2018) is an exception as no impacts on financial knowledge and attitudes is found.
- Overall, there is evidence that interventions are more effective in primary schools relative to KG or preschool. This is also due to an increased facility to identify the proper outcomes and measure them with objective tools.

Socio-demographic considerations

Are there socio-demographic groups of children that do not seem to benefit much from these interventions?

- A consistent finding is that interventions tend to be most beneficial among students who have lowest knowledge at baseline.
- Another consistent finding is that boys and girls seem to benefit equally from these interventions.
 - Danes et al. (2007) showed that the scores for both genders increased after the study (HSFPP) of several concepts, but males outperformed

females. Because of the program, female teens learned significantly more about finances in areas with which they were unfamiliar prior to the program than the males. Furthermore, male teens were confident about making money decisions, and their confidence scores remained higher than those of female teens after the completion of the program. By contrast, female teens increased their confidence about making financial decisions to a larger degree than males. Female teens were also more likely to believe that managing money affected their future before completing the program, and this difference continued to exist after completion.

- Becchetti et al. (2013) found a lower level of financial knowledge among girls in the pre-test, but, after the training, girls had a greater level of progress compared with boys.
- Mandell (2006) however, found that females had more knowledge in the pre-test and learned more from the intervention (seeing a play about savings) than males did.
- More research is needed to uncover whether there are differential treatment effects by race and ethnicity, or by disability status.
- Evidence from the Aflatoun intervention in China (Zhou et al., 2024) shows that this helped improve children's social participation in underdeveloped areas but had little impact in developed areas. Regarding gender, the Aflatoun intervention had a significant treatment effect on girls, with an increase of 0.226 SD in the treatment group compared to the control group, but no significant treatment effect was found in the boy group.
- The various unpublished reports on Aflatot do not really find differential effects on boys vs girls.

Benefits of financial education interventions

What are short-term benefits of these interventions?

- The majority of evidence discussed is short-term, especially for KG education. The multitude of short-term outcomes are discussed throughout this report and summarized in the next bullet points.
- Children's outcomes:
 - Financial outcomes
 - Financial planning
 - Budgeting
 - Financial awareness
 - Attitudes towards saving
 - Financial behavior
 - Financial knowledge
 - Financial interest

- “Social” outcomes
 - Ability to identify emotions
 - Self-efficacy
 - Future orientation
 - Initiative-taking
 - Locus of control
 - Delayed gratification
 - Self-regulation (includes many of the other in this list)
 - Attachment and relationships
 - Respect towards differences and diversity
 - Sensitivity towards differences and diversity
 - Prosocial behavior
- Program understanding and/or satisfaction and/or level of interest in the program

What are long-term benefits of these interventions?

- No real evidence on long-term benefits of interventions in early childhood. Conversely, at later ages, a series of interventions identify persistent effects of social and financial education programs. As reported in Amagir et al. (2018), Bruhn et al. (2013) looked at the long-term impact of the program, 18 months after the start of the program, and found small effects in knowledge ($d = 0.20$), behavior (self-reported and intentions) and attitudes. On the other hand, Mandell (2009), who performed a longitudinal study, found no lasting effects on students’ knowledge, savings behavior, and attitudes toward savings or “thrift” in the post-high school years.

Effects of parental involvement

To which extent do these interventions target parents/families in conjunction with children? To which extent do these interventions suggest that parents’ role and involvement is essential for delivering better outcomes? In other words, to which extent is better financial literacy attributable to the financial education intervention implemented versus to parents’ roles and responsibilities with children?

- The Financial Fitness for Life (FFFL) materials are complemented with a parent guide for each grade level, with activities that reinforce and extend children’s understanding in personal finance (Amagir et al., 2018).
- Some other evidence of parental involvement is provided by Bruhn et al. (2016) in the context of high-school education in Brazil. As part of a financial education program, students also receive take-home exercises, such as creating household budgets with parents, and role-playing assignments. The authors found that parents were significantly more likely to report that their children discussed

financial matters with them at home, and that they volunteered to help organize household budgets.

- Overall, albeit not at the level of pre-K, the few financial education programs that involve parents in the education of their children seem to be effective in increasing the financial literacy of children and adolescents (Batty et al., 2015; Bruhn et al., 2016; Harter & Harter, 2010; Smith et al., 2011). This is in line with Van Camphenout (2015) and Shim et al. (2009), who report that the effectiveness of financial-literacy education could be improved if parental involvement were higher.
- Hence, there is reason to believe this is even more the case before primary school, where reliance on parents is much higher. Totenhagen et al. (2015) address this issue directly in their review. They claim that including parents is vital for preschool children given that the latter are not in school (Holden et al., 2009). For the youngest children, programs may need to be designed for use by parents to utilize with their own children. As children enter school age, parent involvement may become an *accompaniment* to curricula. Most of the programs with a parent component identified were targeted to children and youth at the pre-K through elementary school levels, such as *It's a Habit*, *Money Savvy Kids*, and *I can Save* (Johnson & Sherraden, 2007; O'Neil-Haight, 2010; Schug & Hagedorn, 2005).

Section 3: The Aflatot program and related programs

Existing evidence on the impacts of Aflatoun or Aflateen

Studies have been conducted on the impacts of Aflatoun (targeting primary school children) and Aflateen (targeting children 14+) – this is where most of the evidence is.

- Berry et al. (2018) evaluated the impact of Aflatoun in primary schools in *Ghana*, alongside another intervention that provided financial education only. After nine months, both programs had positive impacts on self-reported savings at school relative to the control group, but there were no statistically significant increases in aggregate savings nor in hypothesized mechanisms such as attitudes, preferences, or knowledge. The lack of short-term effects of these programs on financial behaviors and attitudes indicate that alternative program designs should be evaluated to understand whether and how these outcomes can be influenced among students in this age group.
 - Key outcomes in Aflatoun's theory of change include increased savings (primarily from reductions in expenditures rather than increases in labor

- supply), more favorable attitudes toward savings, and increased financial literacy.
- Aflatoun did not lead children to work more, while the other program (Honest Money Box) did.
 - The authors speculate that the lack of impacts on savings could be a result either of students being too young or of the program's reliance on voluntary enrollment in after-school groups. Perhaps, if integrated into school curricula, the program would have been more effective.
 - The authors also saw the combination of social and financial education as beneficial, and potentially one reason underlying the lack of impacts of Aflatoun on paid labor.
- In the context of *China*, Zhou et al. (2024) adopted a pair-design experiment to evaluate the effectiveness of the Aflatoun Child Social & Financial Education program on students in grades 4-5 of primary school.
 - By using difference-in-differences estimation, the study found that the Aflatoun curriculum significantly improved children's behaviors in regular saving, rational consumption, and social participation. At the same time, unintended adverse negative effects were identified in terms of children's attitudes and personality development. The estimation of heterogeneous effects shows that financial education for children can achieve greater success in improving boys' financial behaviors and in underdeveloped areas.
 - In the context of *Rwanda*, Shephard et al., (2017) claim that financial education should not only be evaluated by measuring the outcomes for students, but also the outcomes for teachers and the methods used by teachers to deliver financial education. As such, they conducted the first randomized controlled trial to analyze the impact of a new version of Aflatoun focusing on both teacher training and life-skills financial education (LSFE). The study thus assesses the effectiveness of a training for teachers on LSFE, active learning methods, and self-efficacy.
 - The theory of change behind the training is that by experiencing and watching their peers use active learning methods (ALM) in the context of LSFE, teachers increase their positive expectancies regarding applying those methods.
 - Targeted to 50 lower-primary and secondary schools
 - The intervention increased teachers' observed use of ALMs as well as the average time on task of the students observed in class. Students in the treatment group also exhibited increased self-reported active learning environments, general self-efficacy, general financial capability, and self-reported savings behavior. Less robust results indicated the intervention increased students' planning attitudes and decreased their self-reported behavioral and cognitive engagement.

- In the context of *Uganda*, Supanantaroek et al. (2017) explored the impact of a social and financial education program developed by Aflatoun on savings attitudes and behavior among primary school children in Uganda.
 - The authors leveraged variation in order of randomization
 - The program lasted 3 months
 - Differently from Berry et al. (2018), the authors found that the intervention increased awareness of money, money recording, and savings attitudes. It also provided some evidence, although less robust, that the intervention increased actual savings.
- In the context of *Tajikistan*, Karimli et al. (2020) assessed the effectiveness of an experiential education program part of Aflateen Plus on a range of measures of agency, including self-efficacy, future orientation, and locus of control among adolescent girls in Tajikistan.
 - Randomized study on 1,221 adolescent girls across 60 public schools
 - The curriculum was comprised of experiential learning in the following seven modules: (1) personal understanding and exploration (participants learn self-reflection and awareness about their own skills, interests, and talents); (2) rights and responsibilities; (3) savings and resource management (adolescents learn how to use resources (e.g., saving money) and participate in a saving system); (4) planning and budgeting (participants learn how to set personal and financial goals, and how to make plans to achieve those goals); (5) social and financial enterprise (participants learn needs assessment, budgeting, and business planning to address the identified needs); (6) education on reproductive health (participants learn options for family planning; prevention of sexually transmitted diseases; and negative outcomes associated with early childbearing); and (7) education on HIV/AIDS (participants learn how HIV can be prevented; how it is transmitted; and possible treatment options).
 - They found significant positive effects on girls' future planning and locus of control, but no effect on self-efficacy.
 - They also found significant intervention effects on attitudes towards saving, financial behavior, and some indicators of health knowledge.

Existing evidence on the impacts of Aflatot (3-6)

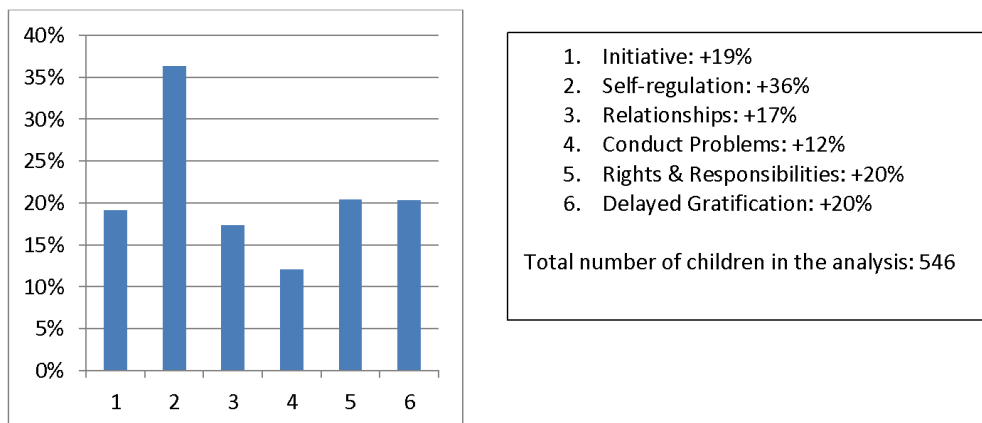
- In Serbia, Croatia, Macedonia, and Albania, Aflatot was integrated into a broader Peace Building programme for early childhood.
 - Children aged 3-6 from diverse backgrounds (e.g., rural, urban, single ethnic, multi-ethnic, single faith, mixed faith groups) participated. The process was monitored monthly and data analysis was conducted.

- In total, 834 children, 331 parents, and 46 teachers participated in the programme.
- Key findings of the pre-post analysis highlighted the following aspects:
 - After the intervention, children were more sensitive and respectful towards cultural and physical differences. The analysis referred in particular to Roma children and children with disabilities.
 - Children were generally more able to identify emotions in themselves or in peers (average increase of 7.7% in the ability to recognize emotions after the intervention).
 - Teachers and parents were asked to what extent they think that young children tend to notice differences in others. The analysis shows an increase in the awareness of teachers and parents towards children's abilities to recognize differences in other children.
- Some evidence from Serbia discusses impacts on training of facilitators, not children or parents.
 - The majority of the trainers agree that the Aflatot training provided them with an understanding of the Aflatoun approach and that it provided them with sufficient understanding on Early Childhood Methodologies.
 - The vast majority felt ready to train other trainers in the Aflatot program and to implement it into their own organisation.
 - 100% agreed that compared to other trainings that they were given, the Aflatot training was more informative.
- Similarly, some evidence from Albania discusses impacts on school staff, not children or parents.
 - Seven classes (210 children) participated in the Aflatot programme. Semi-structured interviews were conducted with their teachers (7). In addition, 4 kindergarten directors were interviewed.
 - Teachers were very pleased with the training they received. They described it as professional and interesting and practical.
 - The respondents were overwhelmingly convinced of the appropriateness of the programme, both contextually as for the different age groups (3-4 and 5-6).
 - In some cases, it was very difficult to engage parents into the programme, although impacts were seemingly stronger when parents were involved.
 - Recommendations mainly addressed the difficulties with involving parents and providing more follow up for teachers.
- Evidence from Kenya through SOS Kenya – Aflatoun report
 - A first important conclusion is that most children show positive behavior related to taking initiative, self-regulation and

attachment/relationships occasionally or frequently after the Aflatot program.

- On average children show negative behavior (such as fighting, inappropriate behavior, offensive language, cry easily, etc.) rarely.
 - No significant difference between boys and girls.
 - Children in KG3 exercise positive behavior related to self-regulation, initiative and attachment/relationship more frequently than those in KG2.
 - Big limitation: no real baseline conducted; through the current design it is not really possible to conclude that benefits observed are due to Aflatot. This is confirmed by the authors of the report: “The program shall conduct a baseline study (if possible for control and intervention groups) in future interventions so that it would be possible to document changes as a result of Aflatot intervention and to adequately learn from the program intervention outcome.
- Evidence from a report on a series of countries where Aflatot was implemented:
 - Children improved in 6 behaviors (taking initiative, self-regulation, attachment and relationships, conduct behavior, rights & responsibilities, and delayed gratification). The biggest improvement was in self-regulation (i.e. listen to others and respect others, control anger, show patience, and share with others).
 - Effectiveness assessed through pre-post averages.
 - Parents’ involvement measured through trainers’ perceptions on parental involvement □ this is not an objective measure.

Graph 1: Overview of results from the Child Observation Tool in Honduras, Ecuador and the Philippines



- Evidence from a similar report on Slovakia highlighted room for improvement in the area of cooperation with families, namely in the forms of communication with

parents and other family members, in using various opportunities for engagement in the kindergarten activities, both directly in the classes and after school.

- The surveyed parents could not very well discern the Aflatoun activities from those of the kindergarten; they evaluated the activities as useful, in particular with regard to the future elementary education of their children. The parents displayed quite good knowledge of the essential financial topics (saving, spending, spending with care); they could give examples, but still this may have been influenced by their actual life, “we live from pay day to pay day,” and the family economic strategies of low-income families. This suggests a need for a more solid approach to involve parents and measure their involvement.

Exploring other similar programs

Are there other “similar” social and financial education programs targeting children?

- Again, most existing literature discusses Aflatoun and related/similar programs targeted to older children. Much less exists on programs analogous to Aflatoun. There are discussions on the importance of social and financial education among preschoolers (Avci, 2022; Birbili & Kontopoulou, 2015), but no evidence of successful interventions.
- The systematic review by Amagir et al. (2018) discusses the Aflatoun program in tandem with another one called “Honest Money Box” (HMB). Both are concerned with investing in one’s own human capital, a particularly crucial endeavor for underdeveloped countries where children and their parents often do not have access to financial services.
- Aflatoun combines a savings program with social education components for children aged 3–18 years executed in 116 countries. “Aflatoun” focuses on personal exploration and children’s rights and responsibilities, while also highlighting the pitfalls of child labor, such as forgoing school to work and the risk of dangerous working conditions.
- The HMB curriculum aims at financial education in a “stricter” sense
- Both Aflatoun and HMB continue as savings clubs after completion of the curriculum (Berry et al., 2018)
 - In terms of teaching approach, for primary and secondary schools the best approach is through “experiential learning;” differently, in college the best approach seems to be to relate teaching styles to specific life events of students, such as life-course markers;
 - An example of experiential learning is provided by the My Classroom Economy (MyCE) program where students actively participate in a “simulated micro-economy” in which they earn “school dollars” that they can use, for instance, to rent their own desks. In this way, students

- experience the impact of their decisions without their teachers teaching specific financial content (Batty et al., 2020);
- Batty et al. (2020) evaluated MyCE and found that students who participated in the MyCE had 6% higher scores in financial knowledge than the control group.
 - Another program that uses hands-on activities to reinforce students' understanding through application and practice, and addresses concepts in a developmentally appropriate manner is the Financial Fitness for Life (FFFL) curriculum.
 - The FFFL material uses various teaching methods like role playing, group discussions, gathering information from the Internet, reading materials, interviewing individuals, drawing pictures, and analyzing case problems. Furthermore, the FFFL materials are complemented with a parent guide for each grade level, with activities that reinforce and extend children's understanding in personal finance.
 - Batty et al. (2015) assessed the FFFL curriculum – alongside MyCE – and found that students significantly gained large positive effects (gain of 11%) in financial knowledge, even after a year. This is a longer timeframe than the one documented by the study of Batty et al. (2020), but MyCE is embedded in other classroom activities, and therefore requires no formal instruction time.
 - Quick overview of similar programs to Aflatoun, as well as findings from studies at the primary-school level in the below figure (note that Collins et al. 2016 is now Batty et al. 2020; Berry et al. 2016 is now Berry et al. 2018):

Table 4. Characteristics and outcomes of studies that assess financial education programs in elementary schools (n=8).

	Target audience	N	Design	Length of program	Program	Outcomes		
						Financial knowledge gains	Financial behavior gains	Attitudes or confidence gains
Batty et al. (2015) USA	Grade 4 and 5	700	RCT	Five weekly lessons of approximately 45 minutes	<i>FFFL</i>	Yes (assessed) (0.77), also after a year	No (self-reported)	Yes, in attitudes
Collins et al. (2016) USA	Grade 4 and 5	2115 (1293 T, 822 C)	RCT	10 weeks	<i>MyCE</i>	Yes (assessed)	Yes (self-reported)	Yes, in attitudes
Berry et al. (2015) Ghana	Grades 5–7	135 schools (average of 40 students per school)	RCT	HMB curriculum: 8 hours and Aflatoun curriculum: 24 hours	<i>Aflatoun and HMB</i>	No (self-reported)	Yes (assessed savings) Cohen's d: Aflatoun=0.078 HMB=0.048	No
O'Neil-Haight (2010) USA	Pre-K -fifth grade	5831	Quasi-NC	Approximately 16 hours	<i>It's A Habit!</i>	Yes (self-reported)		Yes, in attitudes
Roos et al. (2005) South-Africa	4.5–6.5 years	23	Quasi-NC	One week		Yes (assessed)		
Sherraden et al. (2007) USA	Group 1 and 2	149	Multi-method, NC	Four years: Once a week: "after school Club" and one classroom lesson	<i>ICS</i>	Yes (self-reported)	Yes (assessed savings)	Yes, confidence gains
Sherraden et al. (2011) USA	Grade 4	108	Quasi	Four years. Once a week: "after school club" and one half-hour classroom lesson	<i>ICS</i>	Yes (assessed) (0.13)	No (assessed savings)	
Schug and Hagedorn (2005) USA	Grade 2 and 3	316	Quasi-NC	Eight lessons	<i>The Money Savvy Kids</i>	Yes (seven out of 10 items)		Yes, in attitudes

We qualified the (calculated) effect sizes as: small (Cohen's $d \geq 0.20$ and < 0.50), medium (Cohen's $d \geq 0.50$ and < 0.80), and large (Cohen's $d \geq 0.80$) (Cohen, 1988).

Loose Parts media

Wahyuni et al. (2023) claim that a learning model to promote social and financial education in early childhood is most often missing. They found that the financial social education model using loose parts of media effectively develops children's social and financial abilities.

- Loose Parts media is a learning tool in early childhood education institutions that utilizes movable, combinable, re-designable, and "re-purposable" materials. It encourages creativity and offers unlimited learning possibilities for children.
- Loose Parts media is beneficial in teaching children to appreciate materials, engage with nature, and participate in environmental conservation by recycling used goods.
- This is an affordable and adaptable approach to learning that can inspire children's creativity and increase their engagement in play experiences.
- The authors show the effectiveness of this loose parts media approach through a series of interventions in Indonesia (using pre-post comparisons)

Kwartalino Financial Education Program in the Philippines

Some evidence of programs at the preschool level comes from the Philippines. Curugan et al. (2019) assess the impact of the Kwartalino Financial Education Program on the financial knowledge of kindergarten learners.

- The Kwartalino Financial Education Program was specifically designed for kindergarten learners. It has a number of special features such as 1) internationally benchmarked financial lessons and contents; 2) developmentally appropriate instructional materials; 3) varied hands-on and experiential learning activities; and 4) parent involvement.
- Using a quasi-experimental with pre- and post-test design with control and experimental group, the study examines quantitative data to analyze the effect of the program on learners' financial knowledge.
- The researcher developed a 40-item financial knowledge test
- Kindergarten learners who participated in the Kwartalino financial education program scored significantly higher than the control group in the financial knowledge test.
- Results of the study supports the idea that young children can learn money lessons when they are exposed to developmentally appropriate financial education programs.

Smart Money Kit (SMK) in Malaysia

Additional evidence comes from Malaysia. Ramli et al. (2022) examined the impact of a Smart Money Kit (SMK) on the financial interest, financial management behavior, and prosocial level of preschoolers in Malaysia.

- The SMK includes sixteen interactive teaching materials organized as games based on age compatibility for children aged six to twelve that raise awareness among parents and educators on the best financial management practices for children.
- The kit is used in preschools in Malaysia to help actively stimulate the minds of preschoolers.
- This is an RCT, yet with only 50 preschoolers (25 treated; 25 control)
- The focus here is on interest towards financial matters as well as prosocial behavior.
- Results showed a significant effect on the experimental group's level of interest score, a significant effect on the experimental group's level of financial management behavior, and a statistically significant difference in the experimental group's level of prosocial score.
- The authors suggest that SMK can effectively improve children's interests in financial education, management, and prosocial behavior.

What can we learn from other programs?

Where and how have they been implemented?

- Primarily across kindergartens in Eastern Europe and Southeast Asia – most evidence comes from Malaysia and Philippines.

Did they lead to positive outcomes?

- Yes, primarily on self-regulation and emotional skills, much less on actual financial literacy, but this is mostly due to the complexity of measuring financial literacy among kids and the inherent blending that exists between numeracy and financial literacy;
- Child outcomes where some of the strongest impacts are observed: taking initiative, self-regulation, attachment and relationships, conduct behavior, rights & responsibilities, and delayed gratification.

Are positive outcomes observed in some areas more than others?

- Most of these early-life interventions actually focus on staff, teachers, and educators, and not on parents or children.
- Positive outcomes among children observed across interventions include better initiative-taking, self-regulation, empathy towards other children, attachment/relationships.
- Self-regulation emerges as the outcome that is easiest to “shape”

Is there evidence of peer effects among children?

- Not really discussed – important area for future research.

What are some of the challenges that have been encountered?

- Study design: most designs are pre-post average comparisons, while some other have no baseline survey
- Complexity of involving parents (or keeping them involved)
- Complexity of measuring parental outcomes directly – most evidence comes from trainers’ evaluations of parents’ perceptions

Which methods have been used to evaluate/test their effectiveness?

- Primarily pre-post comparisons. No real evidence of RCTs or solid experimental designs, despite available data and study sites.
- Only evidence coming from RCT (Malaysia) is underpowered, with only 50 preschoolers

Were/are those programs scalable?

- Yes, most programs such as Aflatot or similar have been scaled to other contexts
- Doubts (and lack of evidence) remain on the persistence over time of the effects, especially related to financial interest and knowledge. No real designs including consistent longitudinal follow-ups.

Is there any evidence on dosage/intensity related to program effectiveness?

- Kwartalino (*Philippines*, Curugan et al. 2019): seven financial education lessons integrated in learning domains taught in KG class. As such, this was a program embedded into school curriculum for a whole quarter between 2016-17. These lessons were released one day per week and used as integration to “regular” class content. Length of this financial literacy KG integration was around 1.15h.
- Each lesson starts with a story read aloud using a big book as a springboard. These stories are originally written to teach the seven financial lessons integrating at the same time language lessons based on the kinder curriculum. After reading the story, the teacher asks questions related to the story in order to present the financial lesson for the day. The teacher then discusses and explains the financial concepts to the class providing real-life examples and illustrations relevant to the lives of the learners. Then, the teacher provides varied activities to enrich the lesson through games, role play, simulations, and experiential activities.
 - Based on the 7 financial topics included in the program, a specific financial knowledge test was devised, 40-item multiple choice test read out loud to KG learners, each taken in isolation from the rest of the class.
 - This strategy aligns with the review of evaluation tools by Jay et al. (2022), claiming that “Due to children’s levels of reading ability and limitations in ability to follow complex instructions, children generally need to be assessed in one-to-one sessions with a researcher/assessor interviewing one child at a time.”
 - Smart Money Kit (*Malaysia*, Ramli et al. 2022): sixteen interactive teaching materials organized as games. These are basically a series of books with games in them – this is discussed in Ramli et al. (2022) and it involved 3 weeks’ worth of SMK activities, but only 30 minutes per week.
 - In the Slovakia Aflatot report (Ďuríková & Vaněk, 2016), a series of different activities is described; however, it is unclear in all the reports how long each intervention was and how many times each was implemented (versus only once). It appears that these were repeated efforts in classes ongoing for a year. There is no breakdown of length of each component, and it seems this varies in every country.

If parents were/are to be involved, what are some reasonable parental outcomes we should test?

- Most existing evaluations only test parental understanding of the program and parental involvement in the same (or parental relationships with teachers or other staff).
- Other programs just test parents’ agreement of whether children reacted with excitement or not to the program.

- We suggest that other outcomes should be tested, such as parental financial literacy, self-regulation, empathy, as well as a battery of outcomes measuring children’s relationships with parents.
- Another important outcome is parental involvement in kg activities, both at baseline and at follow-up. It may be the case that the program boosted parental involvement even if it was not intended to (e.g., time spent with children on different school-related activities).

Relatedly, if parents were/are to be involved, what are some sensible interventions - or variations of the main intervention - that may make the whole curriculum more participatory?

- Inconclusive evidence. In the Kwartalino program in the Philippines, after the post-test, both the experimental and control group visited a bank. Children were given a copy of the book which is meant to be read at home; a cartoon figure, money locker boxes for savings; a parent invitation to revisit a bank; and a parent questionnaire to determine the effectiveness of the program.
- Parents provided workbooks to reinforce the financial concepts children learn in schools.
- Other options include teacher-made homework activities with parents’ and guardians’ guidance.

Section 4: Recommendations and ideas

Testing outcomes for children

- Most early-life interventions need to test children’s outcomes; so far, evidence is overwhelmingly on staff, teachers, and educators □ there is an urgent need to focus and measure both children and parents’ outcomes.
- Evidence suggests that these early-life programs are quite effective at shaping outcomes such as self-regulation, while less so at shaping financial literacy itself (i.e., the “social” component prevails over the “financial”).
- As such, no real evidence that including explicit “financial lessons” components has any concrete value □ anything experiential and/or hands-on is far more sensible among such young kids.
-

Parental Involvement

- Current structure seems to encourage parental involvement, but it is unclear how: “the curriculum suggests several ways of involving parents and family members in their children’s learning” – existing evidence suggests that this is a key element for the effectiveness of any early childhood program. Evidence

reviewed above gives some suggestions for making parental involvement a “structured” element, but not too much. This is a key area of improvement.

- Several studies show that children learn about finances from parents through intentional instruction, participation, and practice (explicitly) as well as through observation (implicitly). Thus, parents’ role in instilling values of financial literacy education to their children is essential. Parents’ “ignorance” of financial literacy education will only make children learn and understand financial literacy education incorrectly and meaninglessly.
 - As such, parents should be tested “directly,” not through some other people’s (e.g., trainers or evaluators) observations or perceptions, as commonly done.
 - Also, we should not only collect data on parents’ understanding of the program and overall awareness, but also test their financial knowledge directly, as well as children’s relationships with parents and parental involvement in KG activities pre and post intervention. This may provide better hints on the kind of financial socialization occurring at home.
- One additional leg of an intervention that may involve parents more explicitly could include devoting one day to devising more effective strategies parents can leverage to discuss financial matters with kids at home. Like a “financial socialization seminar” with clear takeaways that parents can bring home and implement on a regular basis.

Evaluation tools

- Evaluating impact through classroom observation tools such as the “In-class Aflatot assessment tool” seems appropriate:
 - At the individual level
 - At the classroom level
- It is essential to extend the time period to measure the impact of the program on quality of family involvement in the program activities to at least two years to make sure that the cooperation outcomes are visible and measurable.
 - Relatedly, it is essential to develop tools for quantitative and qualitative measurement of the level and quality of family involvement in the program.
- Potentially, it would be ideal to complement classroom observations by teachers/staff with actual tasks that children have to complete. This would enable scholars and policymakers to evaluate effectiveness of programs through both subjective and objective measures.
- So far, few existing evaluations of Aflatot (which are reports, not actual scholarly outputs) have relied on the Child Observation tool and simple tests for differences in means across groups:
 - Preliminary findings do not properly treat unobserved heterogeneity, potential biases resulting from enumerators’ training, and pre-existing differences among children;

- Expanding all of this and embedding it within an RCT seems the logical and feasible extension;
- As this may not be feasible, some quasi-experimental approaches may be the second-best strategy – potentially doing some matching.
- Essential to include more longitudinal designs to get a better sense of persistence of the effects. As of now, all evidence finds short-term impacts but does not explore long-term sustainability of “effects.”
- In terms of age, it is easier to detect changes in dimensions such as self-regulation, attachment behavior, and initiative among slightly older children.
 - Evidence that age matters and, as such, targeting 3-year-olds may not be the most cost-effective strategy.
 - In an RCT, stratifying analyses by age and grade would seem sensible.
- Currently, all the tools that Aflatoun uses to monitor and track progress of Aflatot are filled by teachers and facilitators, or parents. Aflatoun may consider having actual play-based tests that children may fill.
- Also, despite being called social and financial education, Aflatoun currently does not really test financial literacy as part of Aflatot, or at least not in every country (that have an available report). It only tests social outcomes such as self-regulation and initiative-taking. They may want to consider testing financial outcomes more closely and more universally using similar methodologies.
- One clear piece of evidence is that these early-childhood interventions need to be repeated and spread over time; they cannot be a one-time effort.
- If there is the capacity, these play-based interventions should be introduced as part of the KG curriculum (not just implemented as external interventions). Otherwise, we recommend at least a 45-min play-based session once every other week (or every week, if there is capacity) for at least 3 or 4 months.

Evaluating Social Literacy: The IDELA Scale

Given Aflatot’s overwhelming focus on “social” literacy, the IDELA scale/tool developed by Save the Children and UNICEF (<https://idela-network.org/>) seems appropriate to consider.

- It has been adopted by many countries, which serves the cross-national purpose of Aflatot well
- It measures different aspects of school readiness, including some that Aflatot intends to shape
 - Emergent literacy
 - Emergent numeracy
 - Motor skills (this could probably be dropped, as Aflatot is not interested in this) – it is actually very easy to drop/keep specific modules, so this is another flexible aspect that is key
 - Socio-emotional skills
 - Emotional awareness
 - Empathy
 - Self-awareness

- Solving conflict
- Relationship with peers
- It is recommended for children ages 3.5 to 6, which is exactly Aflatot's target
- It would be relatively easy to insert modules to test intended outcomes at both baseline and follow-up
- It already has a range of existing tools translated in multiple languages, hence no big effort is required
- It is time-effective, as each assessment takes about 35 minutes
- It includes a home environment data collection tool that could get at some of the parental involvement dynamics we are suggesting to look at
- It includes a numeracy sub-item which, in my opinion, could be an indirect proxy of financial literacy, especially in those young kids
- It measures overall school readiness, and we have seen that little research focuses on the extent to which financial literacy matters for school readiness, but this is, in my opinion, perhaps the most policy relevant question when it comes to financial literacy among very young kids.

DELIVERABLE 2:
RESEARCH FEASIBILITY REPORT

EVALUATIVE RESEARCH

Introduction

Drawing from findings in the state-of-the-art review, this report provides a feasibility analysis for conducting evaluative research on the Aflatot program. It delineates the methodologies for framing evaluative research and proposes a sample Randomized Control Trial (RCT) design with methodological considerations and ethical guidelines on conducting research to determine the program's impacts. The recommendations outlined are intended to point Aflatot towards best practices in evaluative research, as well as identify future directions for the program.

Formative and Summative Assessments

Before conducting evaluative research, it is important to clearly define its purpose and objectives. This clarity guides the entire process (See *Appendix A* for further discussion on evaluative research).

A key first step is distinguishing between formative and summative evaluation, and understanding when each type should be applied. Formative assessments aim to monitor and improve the implementation of an intervention program while it is still in progress. They typically involve collecting qualitative and quantitative data to understand participants' experiences, perceptions, and progress. Summative assessments focus on evaluating the overall effectiveness and impact of an intervention program after its implementation has concluded. These assessments primarily utilize quantitative data analysis methods to compare outcomes between treatment and control groups.

Formative Evaluation Questions for Aflatot

Defining clear evaluation questions is essential for guiding the assessment of an educational program developed to promote children's social and financial literacy skills. These questions provide a framework for understanding the program's effectiveness, identifying areas for improvement, and informing decision-making processes. For Aflatot, with its holistic program approach, these questions could address various components of the program, including: implementation quality, the program design, participant response, program modification, and parental involvement.

Below is an exemplary selection of potentially relevant evaluation questions:

- Are the financial literacy curricula and materials being implemented as intended in classrooms? (Implementation 1)
- How are teachers delivering the content, and are they following the recommended instructional strategies? (Implementation 2)
- How well does the curriculum align with the developmental needs and learning abilities of the target age group? (Program Design 1)
- Are the learning activities engaging and interactive, fostering children's interest and understanding of financial concepts? (Program Design 2)
- What are children's perceptions of the financial literacy program? Do they find it enjoyable and relevant to their lives? (Participant Response 1)
- How do children's attitudes and behaviors towards money management change over the course of the program? (Participant Response 2)
- What challenges or obstacles are teachers encountering in implementing the program, and how can these be addressed? (Program Modification 1)
- Based on feedback from teachers and students, what adjustments or enhancements are needed to improve the program's effectiveness? (Program Modification 2)
- How are parents engaged with the program, and what are their perceptions of its impact on their children's financial literacy? (Parental Involvement)
- What strategies can be implemented to enhance parental involvement in the program to further support children's learning at home? (Parental Involvement)

For Further Exploration

Enhancing the Parental Involvement Component

The Family Toolkit

The Aflatot Family Toolkit is an integral component of the program, orienting parents to the program and equipping them to support their children's learning at home. Future studies could explore the direct

effects of the Toolkit on learning outcomes and identify which components are most effective for different demographic groups.

Parents' Perception of the Program

The literature suggests that parents may not fully understand the objectives and benefits of the Aflatot program. Thus, designing specific questions to gauge and subsequently enhance their perception and understanding of the program might be beneficial. These questions can also identify areas where the program may need to improve communication or support to parents.

Parental Financial Literacy

As discovered in the preceding literature review, most financial literacy programs for children focus on evaluating parental engagement. We suggest expanding formative assessments to incorporate other outcomes, including parental financial literacy, self-regulation, empathy, as well as a comprehensive assessment of the dynamics between parents and their children.

Evaluating the Teaching Component

The preceding report on existing literature suggests that some formative evaluation has already been implemented in schools, for instance by asking teachers how satisfied they were with the training that they have received themselves or how appropriate the program was in their own view. Overall, the feedback was quite positive. Also, there is some evidence that children are positively responding to the program. However, there remains an opportunity to formatively evaluate some of the more basic aspects of the program.

Variability in Curriculum Implementation

Most importantly, there seems to be only limited information on whether the program is always implemented as intended. Teachers might or might not stick to the curriculum and, for quality assurance purposes, it would be highly informative to assess the heterogeneity in the quality of implementation together with reasons why teachers eventually did not stick to the curriculum. The identification of specific obstacles for implementing the program is a prerequisite for program improvement.

Contextualization of Teaching Evaluations

As the program is implemented in a variety of contexts, evaluations such as the one mentioned above should reflect this variety as much as possible. The results should not only rely on self-reports but also be conducted with classroom observations.

Impact of Teacher Experience

Finally, an interesting study design would involve selecting teachers that are implementing the program for the first time and comparing them with teachers who have implemented it more often, as well as with those who are implementing it regularly. This would not only show whether experience leads to higher or to lower implementation fidelity, but also allow to identify the obstacles teachers report as a function of their experience.

Addressing these considerations would yield some insights into whether it is necessary to include teacher experience and potentially other contextual variables as a factor in the summative evaluation.

Summative Evaluation Questions

A careful definition of questions that are to be addressed in a summative evaluation is probably even more important than for a formative evaluation. Outlined below are several examples of summative evaluation questions, grouped into outcomes and impact, program objectives, costs and benefits, as well as lessons learned.

- To what extent have children acquired financial literacy skills as a result of participating in the program? (Outcomes and Impacts 1)
- What specific knowledge and skills related to budgeting, saving, and spending have children gained? (Outcomes and Impacts 2)
- Have the program objectives related to financial literacy been achieved? For example, have children demonstrated improved understanding of basic financial concepts? (Program Objectives 1)
- How do children's financial behaviors and decision-making skills compare before and after participating in the program? (Program Objectives 2)
- What are the costs associated with implementing the financial literacy program, including materials, training, and personnel? (Costs and Benefits 1)
- What are the potential long-term benefits of children acquiring financial literacy skills, such as improved financial well-being and decision-making in adulthood? (Costs and Benefits 2)
- What are the key insights and lessons learned from implementing the program? (Lessons Learned 1)
- How can the experiences and outcomes of this program inform the development and implementation of similar initiatives in the future? (Lessons Learned 2)

Distinguishing Proximal and Distal Outcomes

In the evaluation of an educational program aimed at enhancing children's financial literacy skills, it's crucial to differentiate between proximal and distal outcomes. Proximal outcomes are those directly influenced by the program and can be immediately

measured, reflecting short-term changes in knowledge, attitudes, or behaviors. For instance, questions assessing immediate knowledge gained about rational decision-making or budgeting skills acquired during the program align with proximal outcomes. Conversely, distal outcomes represent the program's long-term goals, such as financial stability in adulthood, which may take years to materialize fully. Evaluation questions related to distal outcomes might inquire about participants' financial behaviors and outcomes well beyond the immediate intervention. However, it's advisable for evaluations to prioritize the assessment of proximal outcomes for two main reasons. Firstly, the achievement of proximal outcomes serves as a foundational step toward realizing distal outcomes. If participants fail to demonstrate the desired knowledge or behaviors in the short term, it becomes unlikely that they will exhibit the desired long-term outcomes. Secondly, measuring distal outcomes often requires substantial resources and time, as they may only become evident after a significant temporal distance from the program, potentially delaying the availability of evaluation results. Therefore, by focusing first on assessing and ensuring the achievement of proximal outcomes, evaluators can establish a solid foundation for potential long-term impacts while also obtaining valuable insights into the program's effectiveness in a more timely manner.

As captured previously in the state-of-the-art literature review, current research suggests that existing evaluations fail to address the very proximal outcomes of the program. As outlined above, they tend to focus on the behavioral aspects without addressing emotional and, even more importantly, cognitive outcomes specifically. We recommend measuring these first to ensure that participants are able to demonstrate the knowledge both in the short and in the long term. In both cases, knowledge gains can be assessed by quizzes or related methods right after the instruction, after several lessons, or after several weeks or months. Having a baseline measure to compare with is highly advisable but might bring the challenge of frustrating children with too difficult questions that they are not able to answer yet.

Framing an Evaluative Research Approach for Aflatot: Recommendations

For answering the formative evaluation questions, student self-reports, teacher self-reports and interviews, focus groups, as well as observations in the classroom can complement each other to obtain a comprehensive picture of how the program is perceived. The existing literature suggests that classroom observations coupled with classroom assessment tools are appropriate for young kids, both at the individual and at the group level. Potentially, it is ideal to complement classroom observations by

teachers/staff with actual tasks that children have to complete. We recommend involving an expert in qualitative methodology to prepare the interviews and focus groups as well as to analyze the material obtained.

For the summative evaluation, it is recommended to include a variety of data collection methods that suit the respective evaluation question and that meet high psychometric standards. Whenever existing, tried and tested data collection methods are available they should be preferred over developing them on one's own. However, this will not always be possible, and some methods need to be developed to suit the intervention program. The effort here should not be underestimated as such development requires potentially repeated piloting with sufficiently large samples, psychometric analysis, and revision of the material. Whenever possible, the development of own data collection methods should be guided by a theory. For instance, the instrument to measure socio-emotional skills by Soto and colleagues (2024) was based on a careful consideration on how these skills are structured before attempting to measure them. The authors connected their concept to a successful theory about the structure of personality and this not only helped them to identify the relevant dimensions that needed to be measured, but at the end, it was probably the main reason why the psychometric quality of their instrument was excellent. The subsequent psychometric analysis of a newly developed data collection method should consider aspects such as distribution properties, factorial structure, and validation with similar constructs. We recommend getting support from an expert in test development here who is familiar with the necessary analysis steps. Even seemingly simple quizzes require that the items have optimal difficulty for the specific age (defined as the proportion of participants who answer an item correctly), be unidimensional (in order to be combined to a scale in order to ensure sufficient reliability), and measure the same construct in different subgroups (such as boys and girls or younger and older children). Taken together, the development of a new data collection method requires a lot of conceptual and analytic work that cannot be avoided without compromising its psychometric quality. This is why existing, tried and tested data collection methods should be preferred whenever possible.

As mentioned above, the respective data collection methods of course need to match the summative evaluation question. For testing more proximal outcomes such as knowledge, cognitive self-report methods need to be used. However, they should be complemented with other data collection methods such as conducting behavioral observations or resorting to objective data from records (e.g., school reports). They can not only be used to validate the self-reports but also to scrutinize if knowledge obtained actually translates into behavior. Testing more distal outcomes is most easily achieved

with self-reports that can, however, be subject to all kinds of biases such as socially desirable answering. Whenever possible, more objective measures such as income reports should be used.

The basic requirement for all data collection methods is their objectivity which can be best achieved by taking care that the administration of a method is independent of who administers it in the different contexts. This can become a challenge when, for instance, teachers are conducting the data collection. Some of them might not adhere to the standardized procedures and others might be prone to help their children in quizzes or make them look more successful. Whenever possible, trained external evaluation assistants should be responsible for data collection. This, however, can become very resource intensive and might not always be feasible. If it is only possible to revert to teachers for data collection, they should receive very specific and clear instructions together with an explanation why it is important to stick to the standardized procedures for the sake of objectivity.

Summative evaluation designs typically require the repeated measurement of knowledge or behavior and this might pose a challenge when the data collection method itself changes the trait that is to be measured. When measuring a skill in a challenging situation, for instance, children will learn from the situation and when confronted with it a second time, they will most likely show higher performance. Having a proper control group design allows controlling for that artifact. Still, this could become a problem, if a knowledge question becomes too easy over time, or a test situation is no longer challenging. Then, ceiling effects will make it difficult if not impossible to compare the intervention group with the control group even in a carefully designed randomized control trial.

What has been recommended here for the central outcome variables also holds true for potential control variables. Although here the psychometric challenges are often not that high (when, for instance, measuring the size of the school or a child's age), there might be constructs used as a control variable that require a more careful choice of the data collection method.

Finally, when targeting very young children and selecting or even more so when developing data collection methods suitable for them, it is paramount to consult with an expert in child development (most likely, a developmental psychologist) who is an active researcher and preferably developed data collection methods oneself. Assessing young children is particularly challenging due to their limited attention span, their need for very concrete and vivid stimulus material, their proneness to be distracted quickly, their tendency to give answers in a way they think is expected from them and much more.

Optimally, this expert would be included in the entire design process but at the very least should be asked to review the material intended for use.

Potential Gaps and Opportunities for Further Research

The report on existing literature (regarding Aflatoun’s programs and related programs) suggests that existing summative evaluations have mainly focused on the outcomes and impacts of the program and on the question whether the overall program objectives have been achieved. It also seems that the summative evaluation questions have been defined against the content of the program itself as well as its theory of change. What seems unclear in the existing evaluations is a systematic distinction between affective, behavioral, and cognitive outcomes. It appears that the main focus so far has been on program understanding and behavioral outcomes as assessed by a child observation tool reflecting the intended program outcome and impacts. Assessments of affective as well as cognitive outcomes and impacts seem to be lacking. Affective outcomes could be an important moderator in the program's effectiveness and their assessment could help better understand why the program works better for some children as compared to others. The assessment of cognitive outcomes (e.g., in terms of knowledge gained from the instruction) is crucial to ensure that the program can be effective at all at the behavioral level. If children did not understand what they have been taught, they also cannot implement it in terms of their behavior. This is also essential as the review of the literature seems to suggest that social and financial education programs in early life tend to be more effective at shaping “social” outcomes such as self-regulation, rather than literacy itself. We will come back to this point when distinguishing between proximal and distal outcomes.

Related to the above, another important recommendation relates to the importance of testing children’s outcomes more prominently, as well as their parents’, rather than staff, teachers, and educators. As of now, the literature overwhelmingly prioritizes school staff or, when it does not, it tests program-related outcomes such as individuals’ understanding of the functioning of the program. When it comes to parents, parents should be tested “directly,” not through some other people’s (e.g., trainers or evaluators) observations or perceptions, as commonly done. Also, data should be collected not only in relation to parents’ understanding of the program and overall awareness, but also to test their financial knowledge directly, as well as children’s relationships with parents and parental involvement in KG activities pre and post intervention. This may provide better hints on the kind of financial socialization occurring at home.

What also seems to be underrepresented so far is an analysis of the costs and benefits of the program. The main reason for this probably is that the program objectives are

directed more towards long-term outcomes that require much more time and effort to be measured, including longitudinal designs. This aspect, however, is of utmost importance for convincing the relevant stakeholders to invest into the implementation of the program. From other intervention research we know that programs implemented earlier tend to have a higher return-on-investment, because their implementation is usually cheaper and positive outcomes tend to accumulate over time, if a program is effective. Having such an analysis of costs and benefits, even if it relies on some assumptions that cannot be tested empirically right now, could be an important addition to the summative evaluation framework of the program.

Finally, it is strongly recommended to identify potential moderators of the program's effectiveness and consider summative evaluation questions that explicitly include them. Such moderators can be located at the level of the child, the teacher, or the broader context in which the program is implemented. They can either be identified from the literature or from previous formative evaluations. For instance, one example is the quality of parents' relationship with children, as well as parental involvement in children's activities. This can often provide good indication of children's socialization at home, which is essential to understand social and financial literacy. The purpose of including such moderators is threefold. First, it can help to address the problem of unobserved heterogeneity and hence to reduce error variance when statistically assessing the program's effectiveness. Second, it will allow the project partner to better understand for whom and under what circumstances their program works or does not work as intended. As a consequence of this it will, third, allow to improve the program and potentially offer adapted versions of it for specific subgroups who, for instance, require a more intensive instruction.

Randomized Control Trial (RCT) Design Recommendations

Where summative evaluations are concerned, randomized controlled trials (RCTs) are highly valued as a preferred research methodology for establishing causal relations accurately. Departing from a RCT will always make the interpretation of the findings highly ambiguous and much less convincing for stakeholders. While compromises may be made with regard to the choice of a more convenient control group setting, the sample size, the geographical distribution of the schools involved, or the complexity of the data collection methods, no such compromise should be made either with regard to the reliability and the validity of the measures used.

Control Group Design

Demonstrating the program effects by the means of a RCT evaluation design will always be more credible as compared to other designs, even if a suboptimal control group setting is used, the size of the sample is restricted, and more simple outcome measures are used. Furthermore, there are meta-analyses on intervention research such as the one by O'Mara and colleagues (2006) showing that -- contrary to "conventional wisdom" (p. 197) as the authors state -- effect sizes can even be larger when true random assignment is used.

For a valid interpretation of the program effect, a careful choice of the control group setting is crucial. Pesando (2018) suggests that if a RCT was conducted, evaluators chose a no-treatment control group as a comparison to assess the respective program's effect. Such designs are the easiest to implement but at the same time suffer from not being able to control for non-specific effects. In other words, when differences between

Anecdotal Observation: The Problem with 'No-Control' Group Designs

"From my own experience, in no-control control group designs serious problems with randomization can arise when certain schools are unwilling to serve as a control group. They may demand to be either assigned to the intervention group or drop out of the study completely. Usually, such schools have certain characteristics that are associated with their lack of willingness to participate, such as a high subjective perceived need for this specific intervention or the need to offer their students a broad range of extracurricular activities. In any way, this could become a serious disadvantage that could jeopardize the validity of the entire summative evaluation. Hence, also for this reason no-control group designs cannot be recommended"

Martin Tomasik - LEAP Fellow (2024)

the intervention and the no-treatment control group are found, they might not necessarily be attributed to the program itself. It might be that any intervention totally unrelated to financial literacy skills might have produced the same or a similar effect, because the program effect was, for instance, due to participants' expectations, addition attention they received from the teacher or researcher, or the fact that when participating in the program they could not at the same time participate in deviant activities outside school. For this reason, a placebo or even active control group design is recommended, although finding the right placebo is a challenge in itself.

Everything else being equal, such designs will typically produce smaller effect sizes as compared to a no-treatment control group design. At the same time, it will allow much more valid conclusions to be drawn and hence convince stakeholders (mostly, funders) much more.

Depending on the outcome studied, designs with multiple measurement points can be much more informative than simple pre-post-test-designs. Here, however, a careful selection of the time lag between measurement points is necessary and this can become challenging, because we usually do not know exactly how long some psychological processes need to unfold. Usually, researchers then more or less arbitrarily choose some third or fourth measurement point and this choice is typically based on convenience rather than informed by some theoretical consideration. Studies reviewed by Pesando (2024) chose a third measurement point of 9 or 12 months after the implementation of the program. If the sample size is sufficiently large, one might also think about varying the third measurement occasion between subsamples and hence find out how potential «sleeper effects» build up or what the shape of the «forgetting curve» is.

Designs to identify effective components of the program cannot be recommended at this early stage given that we know very little about what these components might be and how they interact. If there was a meta-analysis available that would identify some relevant moderators, then one could plan a respective study based on empirical evidence. Because there is not and because seriously conducted factorial designs require much more resources, the additional effort would hardly be justifiable.

Selecting the Right Sample Size

Maybe the most important decision besides the choice of a study design is related to the sample size that is used in a summative evaluation. Other things being equal, higher sample sizes allow detecting smaller effect sizes in terms of their statistical significance and with a sufficiently large sample, any effect size can be detected. Statistical significance, however, needs to be distinguished for practical significance that is of primary relevance for stakeholders. Assuming that most stakeholders would be reluctant to support a program that only produces small effect sizes ($d = .20$), programs with medium effect sizes ($d = .50$) are, of course, more desirable. This clearly is a rule of thumb (Bloom et al., 2008) and the effect sizes of typical educational interventions range between small and medium (Hill et al., 2008). McCartney and Rosenthal (2000) even discuss situations in which much smaller effect sizes might be acceptable. Note that in order to detect effect sizes in the medium order of magnitude, the sample size does not need to be particularly large.

Certainly, literature suggests that school-based intervention programs targeting socio-emotional skills have effect sizes in the range of $.20 < d < .30$ and only for some outcome variables medium effect sizes of $d = .50$ are reached (Corcoran et al., 2018; Durlak et al., 2011; Luo et al., 2022). In these meta-analyses and reviews, however, all

types of programs are included ranging from single intervention trials to more sustainable programs regularly implemented in schools. When the goal is to achieve sustainable funding for the interventions program, it is not helpful to run a large (and therefore expensive) evaluation study that at the end will demonstrate the statistical significance of an effect that has only little practical relevance.

We therefore recommend taking a return-on-investment perspective (see, for example, Barrett et al., 2020) when calculating the effect size that would convince potential stakeholders and funders. If an intervention is very cheap, then smaller effect sizes are more acceptable as compared to a situation where an intervention is expensive. For this reason, no break-even number can be recommended here without knowing the costs and (long-term) benefits of a program. However, one needs to consider that small proximal effect sizes will translate into even smaller distal effect sizes (even if one assumes that small effect sizes can accumulate over time).

When calculating the optimal sample size, however, one needs to consider that participants are clustered within school classes and classes are clustered within schools which reduces the information available as compared to a non-clustered sample that is typically assumed for the calculation of test power. Also, it might be of interest to conduct effectiveness analyses in some theoretically interesting subpopulations such as single countries or age groups. Here, in each and every subpopulation the optimal sample size needs to be achieved, which can dramatically increase the overall sample size when many subgroups are considered.

Challenges of Designing a RCT for Young Children

The report on existing literature (Pesando, 2024) suggests there have been at least a few attempts to test the effectiveness of educational programs aimed at enhancing children's financial literacy skills by means of RCTs. These studies have shown mixed results:

- Some RCTs have documented significant effects on certain outcomes but not others.
- Programs targeting early childhood are generally effective at improving outcomes like self-regulation but are less successful at enhancing financial literacy. This suggests that the "social" components of these programs often overshadow the "financial" aspects.
- There is inconclusive evidence regarding the value of including explicit financial literacy curriculum in early childhood education, potentially due to the complexity

of measuring financial knowledge at such a young age. The literature suggests that exposure to experiential and hands-on methods (even briefly) have proven more effective for young children compared to traditional lecture-based or video lessons. Yet, findings in the existing literature do not properly treat issues such as unobserved heterogeneity, potential biases resulting from enumerators' training, and pre-existing differences among children.

Framing a RCT Design for Aflatot: Considerations and Recommendations

As observed in the preceding sections, there are many factors to be taken into consideration in developing an RCT for the Aflatot program. Deviating from a rigorous RCT approach can significantly obscure the interpretation of results and reduce the

Flexibility in Control Group Design

While the integrity of the RCT design must be maintained, some flexibility can be considered in the control group setting, sample size, geographic distribution of participating schools, or simplicity of data collection methods. However, the reliability and validity of the measures used should not be compromised.

impact and credibility of the findings. However, it is not always possible to influence all the conditions required for a successful RCT. Following are some best practices for designing a RCT for Aflatot:

Methodological Considerations

Control Group Design

Proper selection of control groups is crucial.

While no-treatment control group designs are the easiest to implement, they present the downside of not being able to control for non-specific effects. Placebo or active control groups are recommended over no-treatment controls to enhance the validity of the findings.

Moderators of Effectiveness

Identifying **potential moderators** of program effectiveness and considering them in summative evaluation questions enhance understanding and allow for tailored interventions.

Data Collection and Analysis

Employ a variety of data collection methods that align with the evaluation questions and adhere to high psychometric standards. Ensure to **partner with child development**

experts for assessing outcomes among young children, to maintain methodological rigor. Additionally, rigorous data analysis and reporting procedures, including **pre-registration of analysis plans**, are vital for maintaining credibility.

Resource Considerations

When resources are limited (as they often are) it is possible to achieve high reliability and validity by the following means:

- First, before creating your own data collection methods, research whether existing methods could meet your needs. Although they may not perfectly align with your specific domain, developing a new method can be labor-intensive and offers no assurance of success. Using established methods can save time and potentially yield more reliable data.
- Second, when assessment time is limited, it is preferable to measure fewer constructs with sufficient reliability (e.g., by using longer scales or more observations), rather than more constructs with low reliability and hence low validity (e.g., by using short scale versions or observing behavior only once).
- Third, piloting all data collections methods first in the target sample is usually a good idea, even when the data collection method is well-established. This will allow testing for whether these methods work as intended in the specific context and to make adjustments before the actual evaluation study. Piloting is particularly beneficial when choosing between two methods where it's uncertain which will be more psychometrically rigorous. In general, pilot studies are less expensive than the actual evaluation studies so that making mistakes in the pilot study is less consequential.

Ethical Considerations

In conducting the RCT for Aflatot, ethical considerations must be observed. Before getting started, it is essential to secure approval from relevant ethical review boards. In addition, all participating schools, parents, and children must provide informed consent, fully understanding the scope of the study and their rights.

Informed consent involves providing participants or their representatives with comprehensive information about the study's purpose, procedures, risks, benefits, and their rights as participants. Participants should have the opportunity to ask questions

and freely decide whether to participate or withdraw from the study without repercussions.

Informed consent procedures, however, may inadvertently contribute to selective attrition if parents who are more supportive of the intervention or children who experience more progress or simply have more fun in the intervention are more likely to consent to their child's participation. This can bias the sample and undermine the representativeness of the findings and poses a risk to the interpretability of evaluation findings, particularly in school-based intervention programs involving young children. Researchers must carefully consider strategies to minimize selective attrition, such as providing incentives for participation, maintaining open communication with participants and their families, and ensuring that withdrawal from the study is voluntary and without consequences.

Protecting participants' confidentiality and privacy is also essential, particularly in sensitive research involving children. Researchers must take measures to safeguard the confidentiality of participants' personal information and ensure that data are stored securely to prevent unauthorized access or disclosure.

Furthermore, respecting participants' autonomy involves recognizing their right to make informed decisions about their participation in the evaluation. Researchers should ensure that participants, particularly children, understand the purpose and implications of the evaluation and have the opportunity to express their views and preferences.

Finally, researchers have a responsibility to minimize potential harm or risks to participants throughout the evaluation process. This includes implementing appropriate safeguards to mitigate any adverse effects of the intervention or evaluation procedures on participants' well-being, physical or psychological health, or academic performance.

Recommendations

Ethical review boards or experts in ethical research practice should be consulted even if there is no legal requirement to do so. They might help to identify problematic aspects that eventually might jeopardize the reputation of the evaluation study or even the entire programme. Above that, it could pay off to invest some resources in a small communication campaign (e.g., a professionally designed website or a brochure) in order to build trust with parents and minimize selective drop out.

Sample RCT Design

Overview

This RCT aims to determine the efficacy of the Aflatot program in improving financial literacy and socio-emotional outcomes among school children. By utilizing a controlled study design, this evaluation will provide data on the program’s effectiveness.

Sample Primary Research Questions

- What is the impact of the Aflatot program on enhancing children's financial literacy?
- How does the Aflatot program affect children's socio-emotional skills?
- Are there differences in program impacts across various subgroups such as age, socioeconomic status, and gender?

Potential Exploratory Questions

- What additional child outcomes (e.g., self-regulation, initiative-taking) are influenced by the Aflatoun program?
- How do different learning components within the Aflatot program moderate its effects?

Steps to Conducting the RCT / Timeline (12-15 months)

The implementation of the RCT could take anywhere from 12 - 15 months, depending on a variety of factors, including the availability and level of cooperation of schools, access to resources, capacity and competence of evaluators, and others. The specific steps might follow the progression outlined below (rough timeline estimates suggested).

Step	Description	Est. Duration
Sampling, Recruitment and Consent	Sample schools from master list, engage schools, and obtain consent from parents and assent from children.	2 month
Baseline Assessment	Administer pre-intervention assessments to measure financial literacy and socio-emotional skills.	1 month

Randomization	Randomly assign participants to control and treatment groups ensuring stratified sampling based on key demographics.	1 week
Intervention Implementation	Execute the Aflatoun curriculum across treatment groups.	6 months
Endline Assessment	Collect post-intervention data on the same measures as baseline to assess changes.	1 month
Data Analysis	Analyze the collected data to determine the effectiveness of the intervention.	2 months
Reporting	Compile findings into a comprehensive report detailing the outcomes and implications of the study.	1 month

Ethical Considerations

- Ensure all study procedures are approved by relevant institutional review boards.
- Clearly explain the study to participants and obtain written consent from parents and assent from children.

Sample

Population

The sample will include children aged 3 - 6 years from schools participating in the Aflatot program.

Sampling Frame

Optimally, there would be a master list of all schools participating in the program in a given period of time. For sampling schools, explicit and/or implicit stratifiers can be defined (e.g., country, language, socio-economic status). In order to allow valid conclusions on the population of students, the probability of selecting a school should be proportional to its size. Within school, sample the school classes. A school class is the natural smallest sampling unit as the program is conducted in school classes. From a sampling theory perspective, the question comes up how many school classes should be included from a single school. Having more school classes per school can make data collection more economic but due to clustering the design effect increases and the

sampling design becomes less efficient. Statistically, it might be more efficient sampling 5 school classes from 5 schools as compared to sampling 10 school classes from 2 schools. There is always an optimum depending on the size of the sampling units, the variance of their size, and the intra-class coefficient. Calculate the design effects for different types of sampling strategies and find this optimum. Also consider that with more classes per school there is more risk of “cross-group contamination” in the sense that children in the intervention group might influence children in the control group who are in the same school.

Statistical Power and Sample Size

For calculating the optimal sample size, we are assuming a design effect $deff = 3$ in order to account for the clustering of children in schools and school classes. This is quite a conservative estimate that allows a comfortable safety margin. As the intraclass correlation coefficient, we assumed $ICC = .10$, which corresponds to values typically found in school-based educational (Bosker & Witziers, 1995) or health-related research (Shackleton et al., 2016). This number is already “included” in the assumed design effect. Also, a sufficiently high power of 95% is chosen which means that if there is a group difference in reality, it will be detected in 95 out of 100 evaluations. Type I error is set to 5% which means that if there is no group difference in reality, it would still be detected in 5 out of 100 evaluations. This scenario leads to the following sample sizes per group that are necessary to detect a given effect size with a two-sided statistical test. It is obvious that for detecting small effect sizes of $d = .20$ or smaller, the necessary sample size per group increases dramatically.

Effect size d	Sample size per group
.10	7797
.20	1950
.30	867
.40	489
.50	312

Treatment Conditions

- Full Aflatot Curriculum: Participants receive the complete financial and social education program.
- Full Aflatot Curriculum plus Parental Involvement.

- Control Group: Participants are given a placebo (placebo group) or continue with their usual curriculum without additional interventions (no-treatment group). Specific guidance is provided in the ‘recommendations’ section below.

Defining Outcomes

A proper RCT should focus on short-term outcomes first, measuring children’s understanding of the program, children’s “knowledge,” and children’s socioemotional skills that are related to financial literacy, such as self-regulation, emotion management, initiative taking, and attachment behavior.

- Knowledge outcomes are usually measured through proxies for numeracy, by means of picture books, comic books, or in-class games. One example could be counting the number of correct tasks children get when conducting a specific game.
- “Social” outcomes can be tested through standard tools common in developmental psychology (one example is the IDELA score, which also includes self-regulation and executive function) - the current Child Observation Tool seems quite effective for measuring these.
- In addition to these core outcomes, it would be worthwhile to include a measure of domain-specific *self-concept*, even though obtaining such self-reports from younger children is challenging. As compared to behavioral measures, self-concept measures are typically more sensitive to change. Furthermore, it can be argued that a positive self-concept in the respective domain is a precursor for actual performance.

Generally, it is strongly recommended to assess more proximal outcomes in the first stages of an evaluation process. Sticking close to what was actually conveyed in the program does not only make it more likely to find pre-post-improvements but will also help identify those parts of the program that did not work out as intended. When children cannot remember the basic contents of the program, they will not be able to apply it, and expecting any distal outcomes in that situation is futile.

Data Collection Methods

Utilize both direct assessments (e.g., numeracy games) and observational tools (e.g., Child Observation Tool) to measure financial and socio-emotional outcomes at baseline and endline.

The following data collection methods are used in the literature and should be considered to be included in the evaluation study:

Children's outcomes	Option 1	Option 2	Option 3
Financial outcomes			
Financial planning	In-Class Assessment Tool (Aflatot)		Money and Pensions Service (several examples with pictures and comic books and questions to test) - the review by Jay et al. (2022) has some examples with graphics, alongside scoring templates associated to each answer
Budgeting			
Financial awareness	In-Class Assessment Tool (Aflatot)		The few interventions that test actual knowledge tend to develop their own tests, adapted to the tools used
Attitudes towards saving	In-Class Assessment Tool (Aflatot)	IDELA captures emergent literacy and early numeracy, hence it captures many of these aspects in "basic" form	
Financial behavior			
Financial knowledge	40-item test for knowledge developed as part of specific program (Curugan et al. 2019; Kwartalino, Philippines), reading questions out loud to children		
Financial interest	Preschool Learning Behavior Scale (PLBS, McDermott et al. 2002; Ramli et al. 2022)		
Social outcomes			
Ability to identify emotions/empathy	In-Class Assessment Tool (Aflatot)	IDELA	Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Self-efficacy	In-Class Assessment Tool (Aflatot)	IDELA	Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Future orientation	In-Class Assessment Tool (Aflatot)		Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Initiative-taking	In-Class Assessment Tool (Aflatot) Child Observation Tool		Devereux Early Childhood Assessment for Preschoolers

	(Aflatot)		(Lebuffe & Naglieri 2013)
Locus of control	In-Class Assessment Tool (Aflatot) Child Observation Tool (Aflatot)	IDELA	Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Delayed gratification	In-Class Assessment Tool (Aflatot) Child Observation Tool (Aflatot)		Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Self-regulation	In-Class Assessment Tool (Aflatot) Child Observation Tool (Aflatot)	IDELA	Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Attachment and relationships	In-Class Assessment Tool (Aflatot)	IDELA	Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Respect towards diversity	In-Class Assessment Tool (Aflatot) Child Observation Tool (Aflatot)		Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Sensitivity towards diversity	In-Class Assessment Tool (Aflatot)		Devereux Early Childhood Assessment for Preschoolers (Lebuffe & Naglieri 2013)
Prosocial behavior	Preschool and Kindergarten Behavior Scale (Ramli et al. 2022; Smart Money Kit, Malaysia)	Child Prosocial Behavior Questionnaire (10 items)	
Program understanding and/or satisfaction	Series of basic direct questions specific to each intervention, or external observation		

Moderators and Stratification

- Stratify randomization by variables such as age, gender, and school characteristics to control for potential confounding effects and to explore

- subgroup analyses.
- Moderators might include:
 - rural/urban location of school
 - parental involvement
 - time spent with children on different activities out of school
 - type of school (public/private)
 - type of school curriculum
 - time spent at school
 - financial literacy program as part of school curriculum or not
 - financial literacy program provided by school staff vs external staff
 - financial literacy program provided at school or somewhere outside of school
 - frequency of attending Aflatot's session by children
 - frequency of attending Aflatot's session by parents
 - affective measures (e.g. using the Self-Assessment Manikin by Bradley and Lang, 1994)

Recommendations for Implementation

Testing and Validation of Measures

Test the psychometric properties of the outcome measures in pilot studies before using them in the actual evaluation. Observation tools require trained coders who are using well-developed coding manuals as every disagreement will lower the reliability and hence the validity of the measurement. If resources are limited, assess fewer outcome variables but do not compromise on their psychometric properties.

Assessing Outcomes

Children's outcomes should be measured both through trained "external" observers, i.e., progress of children tracked by school staff, but also through children's progress through hands-on tools.

A holistic approach also tests outcomes of school staff (teachers/trainers) and parents. These outcomes should go beyond just stakeholders' understanding of the program - it should test stakeholders on baseline understanding and post-intervention understanding, as well as baseline knowledge and post-intervention knowledge.

- Parents' understanding of programs should be assessed both through external observers (i.e., school staff), as well as through survey questions that parents answer themselves.

- Testing financial literacy skills of parents may be useful and novel - this may provide a sense of how financially literate they are and, as a proxy, how much financial socialization occurs within the household. Post-intervention financial knowledge or savviness of parents may also tell us something about spillover effects from children to parents.
- Another important measure to test both at baseline and post-intervention is parental involvement with children - this could be another good proxy of how much financial socialization occurs in the household, as well as on the extent to which parents can be agents of change.
 - This could be measured through a simple **time-use module** asking parents to report the primary activities they conduct with children every 30-min of a day, as well as through standard indicators of parental involvement.

Longitudinal and Follow-Up Studies

To the extent possible, it is essential to assess persistence of effects on school staff, children, and parents. We recommend a follow-up after 2 years and, potentially, after 5 years. Follow-ups should also include observational tools as well as hands-on tasks (obviously adapted for higher age). This would also allow detecting “sleeper effects”, i.e. effects that only become apparent after some time.

Treatment Group Strategies

In an RCT setting, we recommend two treatment groups, one that targets children only, and one that targets parents too (i.e., children + parents). As such, the design may have one control group and two treatment arms. We recommend that this additional treatment arm targets parents explicitly, i.e., includes content that parents can leverage to discuss financial matters with kids at home. This could be something like a financial socialization seminar. This is also a novel aspect we did not come across in the literature. Potentially, Aflatoun could also think about more treatment arms, including one where parents are only tested in terms of engagement (e.g., through the current Family Engagement Tool), one where parents are included “indirectly” (such as giving children a book to bring home to read with parents, or sending parents an invitation to visit a bank), and one where parents are targeted on their own with a specific intervention (e.g., the above financial education seminar). Overall, the key point is that parents need to be involved much more than they currently are. The younger the children, the more important parental involvement is.

The methodologically most appropriate (but probably also most challenging) control group design would be a placebo control group design in which children would receive some training unrelated to the financial knowledge and social skills conveyed in the actual program. Given the content of the program, finding an appropriate placebo control intervention is not easily accomplished. Knowledge conveyed in this group

should be unrelated to numeracy and the social intervention should not promote self-regulation skills. Examples could be placebo programs related to life-style topics such as health or nutrition, programs focusing on physical activity, art classes, or interventions targeting mindfulness. A selection of school-based programs targeting socio-emotional skills can be found in Siddiqui and Ventista (2018).

A no-treatment or waiting group design would be the second best option that would, however, not allow controlling for non-specific effects. No compromise whatsoever should be made with regard to the random assignment to the intervention with regard to control groups as this will immediately limit the validity of the evaluation study.

Some evidence suggests that an alternative effective strategy could be to invite parents to school and give them teacher-made homework activities to be done jointly by children and guardians.

Concluding Remarks

This report provides a blueprint for assessing the Aflatot program, and may be adapted to guide research on educational interventions for children. Specific to Aflatoun, the analyses provided in both reports could potentially inform Aflatoun's strategic decisions on several fronts. The gaps identified in the literature review point to a few immediate revisions and enhancements that could be made to the Aflatot program. The propositions in Deliverable 2, the research feasibility report, could inform the design of pilot studies to test a variety of outcomes. These pilots will serve as preliminary evaluations before a full-scale RCT, allowing for adjustments based on real-world feedback, resource availability, and operational logistics. Additionally, by highlighting methodological and resource considerations, the feasibility report assists the Project Host in strategic planning and resource allocation, ensuring that immediate resources are wisely invested.

It is also our hope that access to these reports will promote collaborative learning, as these resources circulate and are utilized by various actors in the educational and developmental sectors. The literature review synthesizes existing studies on the effectiveness of financial literacy and socio-emotional programs for young children, providing educators and curriculum developers with evidence-based insights. The feasibility report on conducting RCTs adds a methodological toolset to the public domain, guiding implementers of similar childhood programs on how to rigorously assess educational interventions.

Appendix A

Evaluation Research: An Overview

Evaluation research is a systematic approach to assessing the effectiveness, efficiency, and impact of programs, policies, interventions, or initiatives. It involves gathering, analyzing, and interpreting data to determine whether the objectives of the program or intervention have been met and to inform decision-making processes. By considering these different aspects of evaluation research, researchers can conduct rigorous and meaningful evaluations that contribute to evidence-based decision-making and program improvement. In this feasibility report, we are focussing on the following aspects that are most relevant to consider in evaluation research.

- **Purpose and objectives.** Evaluation research begins with clearly defined purposes and objectives. These objectives outline what the evaluation aims to achieve, such as assessing program effectiveness, identifying areas for improvement, or informing decision-making.
- **Evaluation questions.** Formulating evaluation questions is crucial for guiding the evaluation process. These questions help define what information needs to be collected and analyzed to assess the program's performance and impact.
- **Evaluation design.** The evaluation design outlines the overall approach, methods, and procedures for conducting the evaluation. Common evaluation designs include experimental, quasi-experimental, and non-experimental designs, each with its own strengths and limitations.
- **Data collection methods.** Evaluation research employs various data collection methods, such as surveys, interviews, focus groups, observations, and document reviews. The choice of data collection methods depends on the evaluation questions, the nature of the program or intervention, and the available resources.
- **Data analysis.** Once data is collected, it needs to be analyzed to draw meaningful conclusions. Data analysis techniques may include descriptive statistics, inferential statistics, content analysis, thematic analysis, and qualitative coding.
- **Validity and reliability.** Ensuring the validity and reliability of the evaluation findings is essential for establishing the credibility and trustworthiness of the results. Validity refers to the accuracy and soundness of the evaluation findings, while reliability pertains to the consistency and stability of the results over time and across different contexts.
- **Ethical considerations.** Evaluation researchers must adhere to ethical principles and guidelines throughout the evaluation process. This includes obtaining informed consent from participants, protecting their confidentiality and

privacy, and ensuring that the evaluation does not cause harm or undue stress to participants.

- **Reporting and dissemination.** The findings of the evaluation should be clearly documented and communicated to relevant stakeholders. This may involve preparing evaluation reports, presentations, and other dissemination materials tailored to the needs and preferences of different audiences.
- **Utilization of findings.** The ultimate goal of evaluation research is to inform decision-making and improve programs and policies. Therefore, it is essential to consider how the evaluation findings will be used and to actively engage stakeholders in the interpretation and application of the results.
- **Continuous improvement.** Evaluation is an iterative process, and feedback from stakeholders should be used to refine and improve both the program being evaluated and the evaluation methods themselves.

Formative Evaluation

Formative evaluation focuses on assessing the implementation of a program or intervention while it is still in progress. Its primary aim is to provide feedback and insights that can be used to improve the program's design, implementation, and delivery. Formative evaluation seeks to answer questions such as:

- Are the program activities being implemented as intended?
- What are the strengths and weaknesses of the program design?
- How are participants responding to the program?
- What adjustments or modifications are needed to enhance program effectiveness?

Formative evaluation is particularly valuable during the early stages of program development and implementation when there is still flexibility to make changes based on feedback. It helps identify challenges and opportunities for improvement, ultimately contributing to the program's success.

Summative Evaluation

In contrast, summative evaluation focuses on assessing the outcomes and impacts of a program or intervention after it has been implemented. Its primary aim is to determine the overall effectiveness, efficiency, and impact of the program in achieving its intended goals and objectives. Summative evaluation seeks to answer questions such as:

- What are the outcomes and impacts of the program?

- To what extent have the program objectives been achieved?
- What are the costs and benefits of the program?
- What lessons can be learned for future programming?

Summative evaluation is typically conducted towards the end of the program or intervention period to provide a comprehensive assessment of its success and inform decisions about its continuation, expansion, or termination.

References

- Amagir, A., Groot, W., Maassen van den Brink, H., & Wilschut, A. (2018). A review of financial-literacy education programs for children and adolescents. *Citizenship, Social and Economics Education, 17*(1), 56–80.
- Arrondel, L., Debbich, M., & Savignac, F. (2013). Financial Literacy and Financial Planning in France. *Numeracy, 6*(2), 1–19.
- Avci, F. (2022). Financial Literacy Skills in Preschool Period and Evaluation of Financial Literacy Skills in 2013 Preschool Education Program. *Journal of Inclusive Educational Research, 2*(1), 1–6.
- Barrett, C. A., Gadke, D. L., & VanDerHeyden, A. M. (2020). At What Cost?: Introduction to the Special Issue “Return on Investment for Academic and Behavioral Assessment and Intervention.” *School Psychology Review, 49*(4), 347–358. <https://doi.org/10.1080/2372966X.2020.1817718>.
- Batty, M., Collins, J. M., & Odders-White, E. (2015). Experimental evidence on the effects of financial education on elementary school students’ knowledge, behavior, and attitudes. *Journal of Consumer Affairs, 49*(1), 69–96.
- Batty, M., Collins, J. M., O’Rourke, C., & Odders-White, E. (2020). Experiential financial education: A field study of my classroom economy in elementary schools. *Economics of Education Review, 78*(June), 102014.
- Becchetti, L., Caiazza, S., & Coviello, D. (2013). Financial Education and Investment Attitudes in High Schools: Evidence from a Randomized Experiment. *Applied Financial Economics, 23*(10), 817–836.
- Berry, J., Karlan, D., & Pradhan, M. (2018). The Impact of Financial Education for Youth in Ghana. *World Development, 102*, 71–89.
- Birbili, M., & Kontopoulou, M. (2015). Financial Education for Preschoolers: Preparing Young Children for the 21st Century. *Childhood Education, 91*(1), 46–53.
- Bosker, R. J., & Witziers, B. (1995). *A meta-analytical approach regarding school effectiveness: The true size of school effects and the effect size of educational leadership* (ED392147). ERIC. <https://files.eric.ed.gov/fulltext/ED392147.pdf>
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., & Zafar, B. (2016). Financial Education and the Debt Behavior of the Young. *Review of Financial Studies, 29*(9), 2490–2522.
- Bruhn, M., Leão, L. de S., Legovini, A., Marchetti, R., & Zia, B. (2016). The impact of high school financial education: Evidence from a large- scale evaluation in Brazil. *American Economic Journal: Applied Economics, 8*(4), 256–295.
- Cedeño, D., Lannin, D. G., Russell, L., Yazedjian, A., Kanter, J. B., & Mimnaugh, S. (2021). The effectiveness of a financial literacy and job-readiness curriculum for

- youth from low-income households. *Citizenship, Social and Economics Education*, 20(3), 197–215. <https://doi.org/10.1177/20471734211051770>
- Chen, H., & Volpe, R. P. (1998). An Analysis of Personal Financial Literacy Among College Students. *Financial Services Review*, 7(2), 107–128.
- Chu, Y. L. L., & Toh, T. L. (2020). A framework for designing mathematics instruction using comics at the primary school level. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, 5(3), 218–230. <https://doi.org/10.23917/jramathedu.v5i3.11373>
- Cronqvist, H., & Siegel, S. (2015). *The Origins of Savings Behavior*.
- Curugan, A. A. M., Masnan, A. H., & Norwani, N. binti M. (2019). Effect of Kwartalino Financial Education Program on Kindergarten Learners' Financial Knowledge. *International Journal of Social Science and Economic Research*, 4(4), 3144–3164.
- Danes, S. M., Haberman, H. R., & Candidate, M. A. (2007). *Teen Financial Knowledge, Self-Efficacy, and Behavior: A Gendered View*. <http://ssrn.com/abstract=222840648>
- Drever, A. I., & Else-Quest, N. M. (2021). Financial literacy among children. In *The Routledge Handbook of Financial Literacy*. Routledge.
- Drever, A. I., Odders-White, E., Kalish, C. W., Else-Quest, N. M., Hoagland, E. M., & Nelms, E. N. (2015). Foundations of financial well-being: Insights into the role of executive function, financial socialization, and experience-based learning in childhood and youth. *Journal of Consumer Affairs*, 49(1), 13–38. <https://doi.org/10.1111/joca.12068>
- Ďuríková, E., & Vaněk, B. (2016). Pre-post evaluation in Slovakia. Open Society Foundation, Slovakia & Aflatoun International, Netherlands.
- Erner, C., Goedde-Menke, M., & Oberste, M. (2016). Financial Literacy of High School Students: Evidence from Germany. *The Journal of Economic Education*, 47(2), 95–105.
- Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8), 1861–1883.
- Fornero, E., & Monticone, C. (2011). Financial Literacy and Pension Plan Participation in Italy. *Journal of Pension Economics and Finance*, 10(10), 547–564.
- Gnatuk, C. A., & Granovsky, N. (2014). *Building Your Preschooler's Financial Capabilities: Information for Building Your Preschooler's Financial Capabilities: Information for Parents and Caregivers Parents and Caregivers*. https://uknowledge.uky.edu/fcs_reports/117
- Guiso, L., & Jappelli, T. (2009). Financial literacy and Portfolio Diversification. *CSEF Working Papers 212, Centre for Studies in Economics and Finance*.
- Harter, C., & Harter, J. F. R. (2010). Is Financial Literacy Improved By Participating In A Stock Market Game? *Journal for Economics Educators*, 10(1), 21–32.
- Hill, C.J., Bloom, H.S., Black, A.R. and Lipsey, M.W. (2008), Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2: 172-177. <https://doi.org/10.1111/j.1750-8606.2008.00061.x>

- Holden, K., Kalish, C., Scheinholz, L., Dietrich, D., & Novak, B. (2009). *Financial Literacy Programs Targeted on Pre-School Children: Development and Evaluation*. 1–52.
- Jay, T., Rashid, S., Dervou, I. X., & Moeller, K. (2022). *Measuring Financial Literacy of Children Aged 4 to 6 years: Design and small-scale testing*.
- Johnson, E., & Sherraden, M. S. (2007). From Financial Literacy to Financial Capability among Youth. *Journal of Sociology & Social Welfare*, XXXIV(3), 119–147.
- Jorgensen, B. L., & Savla, J. (2010). Financial Literacy of Young Adults: The Importance of Parental Socialization. *Family Relations*, 59(4), 465–478.
- Karimli, L., Shephard, D. D., McKay, M. M. K., Batista, T., & Allmang, S. (2020). Effect of Non-formal Experiential Education on Personal Agency of Adolescent Girls in Tajikistan: Findings from a Randomized Experimental Study. *Global Social Welfare*, 7(2), 141–154.
- Lusardi, A. (2015). Financial Literacy: Do People Know the ABCs of Finance? *Public Understanding of Science*, 24(3), 260–271.
- Lusardi, A., & Mitchell, O. S. (2007). Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education. *Business Economics*, 42(1), 35–44.
- Lusardi, A., & Mitchell, O. S. (2011a). Financial Literacy and Retirement Planning in the United States. *Journal of Pension Economics and Finance*, 10(04), 509–525.
- Lusardi, A., & Mitchell, O. S. (2011b). Financial Literacy Around the World: an Overview. *Journal of Pension Economics and Finance*, 10(4), 497–508.
- Lusardi, A., Mitchell, O. S., & Curto, V. (2010). Financial Literacy among the Young. *The Journal of Consumer Affairs*, 44(2), 358–380.
- Lusardi, A., & Tufano, P. (2015). Debt Literacy, Financial Experiences and Overindebtedness. *Journal of Pension Economics and Finance*, 14(4), 329–365.
- Mandell, L. (2008). Financial Literacy of High School Students. In J. J. Xiao (Ed.), *Handbook of Consumer Finance Research*. Springer.
- McCartney, K. and Rosenthal, R. (2000), Effect Size, Practical Importance, and Social Policy for Children. *Child Development*, 71: 173-180.
<https://doi.org/10.1111/1467-8624.00131>
- McCormick, M. H. (2009). The effectiveness of youth financial education: A review of the literature. *Journal of Financial Counseling and Planning*, 20(1), 70–83.
- McCormick, M. H., & Godsted, D. (2006). *Learning Your Monetary ABCs: The Link between Emergent Literacy and Early Childhood Financial Literacy Learning*. 2006-NFI-03, 2–14.
- Moreno-Herrero, D., Salas-Velasco, M., & Sánchez-Campillo, J. (2018). Factors that influence the level of financial literacy among young people: The role of parental engagement and students' experiences with money matters. *Children and Youth Services Review*, 95, 334–351. <https://doi.org/10.1016/j.childyouth.2018.10.042>
- O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis. *Educational Psychologist*, 41(3), 181–206.
https://doi.org/10.1207/s15326985ep4103_4

- O'Neil-Haight, M. (2010). Educator Teams up to Teach Finance to Young Children. *Journal of Family and Consumer Sciences*, 102(2), 43–47.
- Pesando, L. M. (2018). Does financial literacy increase students' perceived value of schooling? *Education Economics*, 26(5), 488–515.
<https://doi.org/10.1080/09645292.2018.1468872>
- Ramli, N. N., Roslan, S., Sulaiman, T., Kadir, S. A., & Zaremohzzabieh, Z. (2022). The Effect of a Smart Money Kit on the Financial Interest, Financial Management Behavior, and Prosocial Level of Preschoolers. *Pertanika Journal of Social Sciences and Humanities*, 30(3), 1283–1297.
- Scheinoltz, L., Holden, K., & Kalish, C. (2011). Cognitive Development and Children's Understanding of Personal Finance. In *International Series on Consumer Science* (pp. 29–47).
- Schug, M. C., & Hagedorn, E. A. (2005). The Money Savvy Pig™ Goes to the Big City: Testing the Effectiveness of an Economics Curriculum for Young Children. *The Social Studies*, 96(2), 68–71.
- Shackleton, N., Hale, D., Bonell, C., & Viner, R. M. (2016). Intra-class correlation values for adolescent health outcomes in secondary schools in 21 European countries. *SSM - Population Health*, 2, 217–225. <https://doi.org/10.1016/j.ssmph.2016.03.005>
- Shephard, D. D., Kaneza, Y. V., & Moclair, P. (2017). What curriculum? Which methods? A cluster randomized controlled trial of social and financial education in Rwanda. *Children and Youth Services Review*, 82(March), 310–320.
- Shim, S., Xiao, J. J., Barber, B. L., & Lyons, A. C. (2009). Pathways to life success: A conceptual model of financial well-being for young adults. *Journal of Applied Developmental Psychology*, 30(6), 708–723.
- Skagerlund, K., Lind, T., Strömbäck, C., Tinghög, G., & Västfjäll, D. (2018). Financial literacy and the role of numeracy—How individuals' attitude and affinity with numbers influence financial literacy. *Journal of Behavioral and Experimental Economics*, 74, 18–25. <https://doi.org/10.1016/j.socec.2018.03.004>
- Siddiqui, N., & Ventista, O. M. (2018). A review of school-based interventions for the improvement of social emotional skills and wider outcomes of education. *International Journal of Educational Research*, 90, 117–132.
<https://doi.org/10.1016/j.ijer.2018.06.003>
- Smith, R. C., Sharp, E. H., & Campbell, R. (2011). Evaluation of financial fitness for life program and future outlook in the Mississippi Delta. *Journal of Economics and Economic Education Research*, 12(2), 25–39.
- Soto, C. J., Napolitano, C. M., Sewell, M. N., Yoon, H. J., & Roberts, B. W. (2024). Going Beyond Traits: Social, Emotional, and Behavioral Skills Matter for Adolescents' Success. *Social Psychological and Personality Science*, 15(1), 33–45.
<https://doi.org/10.1177/19485506221127483>
- Supanantarok, S., Lensink, R., & Hansen, N. (2017). The Impact of Social and Financial Education on Savings Attitudes and Behavior Among Primary School Children in Uganda. *Evaluation Review*, 41(6), 511–541.
- Te'eni-Harari, T. (2016). Financial literacy among children: the role of involvement in saving money. *Young Consumers*, 17(2), 197–208.

- Totenhagen, C. J., Casper, D. M., Faber, K. M., Bosch, L. A., Wiggs, C. B., & Borden, L. M. (2015). Youth Financial Literacy: A Review of Key Considerations and Promising Delivery Methods. *Journal of Family and Economic Issues*, 36(2), 167–191.
- Van Campenhout, G. (2015). Revaluing the Role of Parents as Financial Socialization Agents in Youth Financial Literacy Programs. *The Journal of Consumer Affairs*, 49(1), 186–222.
- van den Heuvel-Panhuizen, M., van den Boogard, S., & Doig, B. (2009). Picture books stimulate the learning of mathematics. *Australasian Journal of Early Childhood*, 34(2).
- van Rooij, M. C. J., Lusardi, A., & Alessie, R. J. M. (2012). Financial Literacy, Retirement Planning, and Household Wealth. *The Economic Journal*, 122, 449–478.
- Wahyuni, S., Liza, L. O., Syahdan, Rusandi, M. A., & Situmorang, D. D. B. (2023). ‘Treasure hunt’: Using loose parts media to develop social financial education model for early children. *Heliyon*, 9(6), e17188.
- Xiao, J. J., Ahn, S. Y., Serido, J., & Shim, S. (2014). Earlier Financial Literacy and Later Financial Behaviour of College Students. *International Journal of Consumer Studies*, 38(6), 593–601.
- Xu, Y., Briley, D. A., Brown, J. R., & Roberts, B. W. (2017). Genetic and environmental influences on household financial distress. *Journal of Economic Behavior and Organization*, 142, 404–424. <https://doi.org/10.1016/j.jebo.2017.08.001>
- Zhou, J., Feng, S., Wu, L., & Wang, S. (2024). The Impact of Financial Education for Children: Evidence from an Experiment in China. *Asia-Pacific Education Researcher*, 33(1), 157–169.