

## Why STEM Inquiry Programs must begin in Early Childhood Education

Po Chu NG<sup>1</sup>

### Abstract

STEM education, focusing on cultivating a "scientific spirit" and "learning by doing" through inquiry, observation, and problem-solving, has gained global attention. However, evidence supporting its efficacy for Kindergarten children (K1-K3) is limited. This paper reports findings from a Science Inquiry Program (SIP) intervention study in 18 early childhood settings (6 experimental; 12 control) in Hong Kong, aligned with the Curriculum Development Council (2017) "Nature and Life" guidelines. Surveys at three timepoints (T0, T1, T2) addressed: Do children in the experimental group improve more than the control group in multiple outcomes? Results show that with teacher professional development and emphasis on agency, a light-dose intervention can significantly enhance skills, attitudes, confidence, interest, and reduce negative emotions. The essence of STEM inquiry should be to cultivate learners' curiosity and problem-solving skills through securing educator-child bonds to ease inquiry anxiety, reinforcing early STEM's value and implementation success in Hong Kong's preschool context.

Keywords: STEM; early childhood education; pre-school; inquiry; inquiry-based curriculum

Corresponding author: Dr. Po Chu NG, Email: [pcng.creative@gmail.com](mailto:pcng.creative@gmail.com)

---

<sup>1</sup> Dr. Ng Educational Studio Limited, Hong Kong

## **Introduction**

Engaging students in self-initiated scientific inquiry processes is central to the Science Inquiry Program (SIP), which cultivates a "scientific spirit" and "learning by doing" to address the abstract nature of STEM education (Chou, 2020; Curriculum Development Council, 2017). This approach mitigates barriers like teachers' limited training, children's lack of opportunities, and parental disengagement (Fuller et al., 2021), exacerbated in Hong Kong by resource constraints and rote learning traditions. Drawing from attachment theory (Bowlby, 1969), secure educator-child relationships in the SIP reduce anxiety, enhancing engagement in inquiry tasks. The global trend of increasing STEM focus (Fuller et al., 2021; Yüceliyigit & Toker, 2021) underscores the purpose of this research which includes: (1) designing a STEM program for early childhood to foster curiosity, problem-solving, and environmental awareness, and (2) evaluating its effects on students' STEM learning.

## **Aims**

The overarching aim is to examine the effects of the Science Inquiry Program (SIP) on student outcomes related to STEM learning.

## **A program utilising Inquiry-based learning in STEM to improve student engagement**

The research attempted to answer one research question (RQ): Do Kindergarteners in the experimental group improve more than the control in multiple outcomes after the SIP?

The significance of this research lies with its potential in providing an insight to optimise the benefits of applying an inquiry-based approach to STEM education for young children. The research findings could guide early childhood educators to design, plan, and implement a STEM model (see Aldemir & Kermani, 2017) and achieve the desired goals of STEM inquiry at a very young age by capitalising on their curiosity to develop essential skills such as critical thinking. The findings may also guide us to improve current early childhood STEM practices to obtain the best effects.

This research will make a significant contribution to the discourse on STEM education and inquiry-based learning. Previous STEM education research with preschool samples is limited, and studied either teachers or students, used small samples, and may not use a longitudinal design, making it hard to claim the effects of intervention. The present research uses a longitudinal design and a larger sample than previous studies (e.g. student samples less than 150 in each of five studies in Yüceliyigit and Toker's (2021) meta-analysis). Hence, the validity and reliability of the results contribute to a clearer understanding of STEM learning.

STEM is increasingly popular because STEM process skills developed in early childhood are believed to have lifelong positive effects on learners' development (Yang et al., 2024). That is, the skills children learn from a STEM program, the attitude they develop through the learning process, and the competence and confidence they build up in the explorations and sense-making will cascade and transfer to other areas of learning (Stephenson Reaves et al., 2022).

## **Focus and contents of the SIP**

STEM education at the Kindergarten stage should not be advanced scientific theories or creating amazing products. The essence of STEM should be to cultivate students' "scientific spirit" and "scientific learning steps", that is, to guide children to observe, be curious, ask questions, find possible answers, and then organize the experience gained in the process. STEM also emphasizes the concept of "learning by doing", which emphasizes whether children can use "integrated application" skills to solve problems. Accordingly, the "Nature

and Life” area in the *Hong Kong Kindergarten Education Curriculum Guide* (Curriculum Development Council, 2017) emphasizes that children should be able to:

- develop curiosity about the environment and phenomena around and to enjoy exploring the surroundings and nature;
- cultivate an objective and open-minded attitude; to acquire basic exploratory skills through observation, questioning and making assumptions; and to develop abilities to solve problems;
- appreciate, respect and care for nature and live an environmentally-friendly life.

The emphasis on inquiry, scientific spirit, problem-solving skills, and positive learning attitude, all require children to use their comprehensive abilities, which all contribute to the learning concept of STEM. In essence, for young children, STEM activities are designed to help develop the ability to “explore, solve problems, conduct engineering activities, and integrate curriculum” (Chou, 2020). Even though all children are naturally endowed with curiosity, Curriculum Development Council (2017) emphasises the importance of developing both curiosity and skills through learning activities to facilitate students to fully achieve their potentials in life. Achieving the goals of fostering curiosity, scientific thinking, problem-solving, and a positive learning attitude requires children to actively use their full range of existing abilities. These educational goals cannot be achieved passively; children need to actively draw upon their full set of abilities to succeed.

The Science Inquiry Program (SIP) illustrated here was developed by Ng (2020, 2024), integrating the above concepts. Through the SIP, children were expected to

"be curious about things and be willing to explore the things around them and natural phenomena; have an objective and open attitude towards things, learn basic methods of exploring things such as observation, questioning and hypothesis, and develop problem-solving skills" (Curriculum Development Council, 2017).

Essentially, the curriculum includes three elements – "scientific and mathematical cognition", "constructive learning" and "exploratory learning" – echoing the learning objectives and learning expectations of the "Nature and Life" category in the *Hong Kong Kindergarten Education Curriculum Guide* (Curriculum Development Council, 2017). These three elements are implemented in the SIP as scientific exploration activities, free play activities, and exploratory learning activities.

**Focusing on scientific exploration.** The study reported here focused on scientific exploration activities for children at all levels of Kindergartens (K1, K2, K3) in Hong Kong (Figure 1). The activities were designed based on the principle of "scientific learning spirit" which guides children to explore, observe, ask questions, predict, and organize their learning experiences (Curriculum Development Council, 2017).

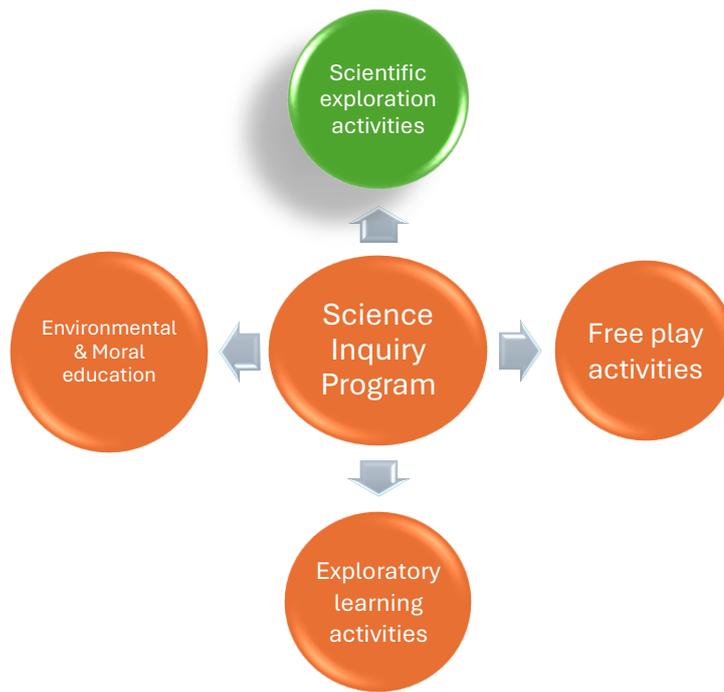


Figure 1. Focusing on scientific exploration activities (Ng, 2020, 2024)

Importantly, the "science exploration activities" conducted in this study followed the principles of being age-appropriate, scope-appropriate, and level-appropriate, and selected relevant activities involving environmental education, moral education, etc. for K1-K3, suited to their learning abilities. Emphasis was placed on learning by doing and taking scientific steps as the spirit of exploration. During an activity, the teacher would lead the children to discover the problem to be explored, and then the children would discuss ways to solve the problem (forming hypotheses) and what tools and materials to prepare before moving on to the next step, and then find a solution "hands-on". Through the hands-on activities, children would verify or refute their earlier assumptions and hypotheses. Finally, they would record the procedures and results of their hands-on activities on worksheets and make a summary to share with others in class. In the process of such activities, success and merit were assessed in terms of how children carried out the activities according to "scientific steps" (including "exploring problems", "predicting and inferring", "verifying methods", and "discovering and interpreting results"). Obtaining consistent experimental results or attempting to find a definite solution was not particularly celebrated.

### Expected outcomes of SIP

This study evaluates the Science Inquiry Program (SIP) on Hong Kong Kindergarteners' outcomes, addressing gaps and limitations in prior research (Yücelyiğit & Toker, 2021). It examines: (1) skills development, crucial for STEM learning; (2) attitudes toward STEM, understudied in preschoolers despite positive effects in older learners (Saraç, 2018); (3) self-concept enhancement, vital for sustainable learning (Kadir et al., 2017; Marsh et al., 2023); and (4) reduction of negative emotions (e.g., fear, anxiety), which may hinder inquiry activities. By integrating attachment theory (Bowlby, 1969), SIP fosters secure educator-child bonds to mitigate anxiety, aiming to bridge the research gap on preschool STEM impacts.

**Skills.** In this research, focus was placed on nurturing skills more than knowledge. Although skills and knowledge are intertwined, and good skills are often built upon adequate

knowledge such that skills and knowledge go hand in hand in education (Choy & Yeung, 2024), preschool children have very limited knowledge to base on for building skills. However, if facilitated, they are able to apply their innate ability of exploration to acquire and exercise skills to make sense of the environment they live in. Hence, an emphasis on learning and practising skills to explore and to make sense of the surrounding phenomena would better encourage students to engage in hands-on inquiry than an emphasis on knowledge. For optimal acquisition of skills, children need facilitation by an adult through scaffolding (Vygotsky, 1978) to build advanced skills from the basics, and this is the essence of the SIP.

**Attitude.** Attitude is an influential factor that facilitates performance and personal satisfaction. A learner's behaviours are essentially influenced by attitude (Ajzen & Fishbein, 1977). A learner who has a positive attitude towards learning will display favourable learning behaviours. Lee et al. (2019) have demonstrated that students holding a positive attitude towards technology and a belief in the ability to use technology tend to choose to become more familiar with technology, and to make the best use of it. However, even though a positive attitude at school or at work is highly valued, unlike skills, a desirable attitude is not always teachable (Josiam et al., 2010). In our program, it is unclear whether the students hold a positive attitude, and whether our program results in improvement of their attitude. It would not be surprising if no change was detected in attitude towards STEM through the program, given the stable nature of attitudes once established (Petty & Cacioppo, 1986).

**Confidence (cognitive self-concept).** Self-concept is a game changer for academic success (Marsh et al., 2023). Two components of academic self-concept have been examined in recent research (Kadir et al., 2017). Cognitive self-concept refers to a student's sense of competence (or confidence) whereas affective self-concept refers to interest and enjoyment. Cognitive self-concept is related to academic achievement and shorter-term performance outcomes (Kadir et al., 2020). Hence, we may expect that enhancing cognitive STEM self-concept may enable young learners to perform well in STEM. We use the term 'confidence' hereafter for ease of interpretation and understanding.

**Interest (affective self-concept).** Affective self-concept refers to a student's interest and positive affect to study (Kadri et al., 2017). A positive affective self-concept tends to have more long-term effects on motivation and engagement in future learning (Kadir et al., 2017). Students with a strong affective self-concept are likely to attend to a topic of interest and subsequently achieve (Choy & Yeung, 2022). Hence, we may anticipate that through enhancing affective self-concept, educators may enable young learners to engage and achieve in STEM in the long term. We use the term 'interest' hereafter for ease of interpretation and understanding.

**Emotion (negative).** Negative emotions (e.g., fear, anxiety, worry) may have negative influences on an individual's behaviour and performance (Pekrun, 1992). For preschool children, as they may not feel comfortable in inquiry activities with minimal adult assistance, negative emotions may also arise. Our concern in this research is whether the program can bring down the level of negative emotion for the students attending SIP.

### **Addressing a Significant Gap in Research**

While STEM education has attracted much attention and STEM programs are advocated by scholars and educational institutions, with many arguing for early starts in SIP activities, there is unfortunately inadequate compelling evidence to support why STEM inquiry must begin in early childhood. To support an early start of an SIP, strong evidence is

required to show that (1) an SIP can be implemented with an effective program design, (2) preschool learners are able to benefit from the SIP, (3) success of learning is measurable, and (4) gains from SIP are noteworthy even with a small dosage of intervention (such as teacher professional development for a few hours followed by teacher implementation of SIP in a few lessons in Kindergarten). The current research attempted to illustrate these possibilities in support of an early STEM education for young children.

### **Methods**

The research used a longitudinal design, collecting data at multiple timepoints: Time zero (T0) before the program starts, T1 after 3 months, and T2 after another 3 months. The findings presented in this report include T0, T1, and T2 students' self-report survey data about STEM education.

### **The Sample**

The sample included students from classes in preschools in the Hong Kong Special Administrative Region of China. Data were collected at three time points (T0, T1, and T2, 3 months apart). As participants may drop out, or fail to provide data at a certain point, while new participants may join in at later stages, the sample sizes may vary in the analyses presented in this report.

At T0, the sample was comprised of 18 Kindergartens, with a total of 149 classes and 277 teachers. The total numbers of participating students at T0 were 3457. Students' minimum age=2 years 8 months, maximum age=6 years 6 months, mean age = 4 years 10 months (51.0% male). As some students attending the sampled Kindergartens were above the maximum age and might affect the interpretation of the findings of this research for preschoolers, the analysis used the data from children at or below 6 years and 6 months old to explicate the intervention effect of the program specifically for Kindergarten learners ( $N=3400$ ).

All the participating schools provide a 3-year Kindergarten education. At T0, the composition of students by class level reported by the teachers was: K1:  $N=1019$  (30.0%); K2:  $N=1210$  (35.6%); K3:  $N=1171$  (34.4%). Valid samples for the analyses presented here are  $N=3400$ . (Data for children over 6 years 6 months old are excluded from this analysis.)

### **The Instrument**

Two quantitative surveys on learning in the Science Inquiry Program were administered. In the student survey. Variables included students' self-ratings on Skills, Confidence, Interest and Negative Emotion, each comprising multiple response items. Ratings of each item ranged from 1 = low to 5 = high. In the teacher survey, teachers were asked to rate each student's attitude towards the SIP on the same 5-point scale. The scores from both surveys were used in the analysis. Sample items are given in Table 1.

### **Procedure**

Ethics approval was sought and granted by the institution conducting the research before the schools and participants were recruited. Parents were contacted by the participating schools and were provided with information about the research. Signed consents were collected from parents for their children to participate. Teachers in the experimental group briefed their students about the research, and all students were encouraged to ask any question, especially when being surveyed by their teacher orally. All students from 18 kindergarten were surveyed as least 1 week before the start of the program (T0). All participants (and their parents) could choose to, or not to, participate in the program, at any time during the study. All responses to the instruments were obtained online, and

confidentiality and anonymity were maintained for all participants in the research.

## Analysis

**Preliminary analysis.** Descriptive statistics are reported. Reliability analyses by Cronbach's Alpha were conducted where appropriate.

**Construct validation.** To validate the hypothesized constructs, a series of confirmatory factor analysis (CFA) was conducted using the lavaan package in R (R4.3.1; Rosseel, 2023). Model fit was assessed by Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the comparative fit index (CFI). The chi-square test statistics are also reported. In general, the values of TLI and CFI equal to or larger than .90 are considered an acceptable fit while .95 indicates an excellent fit to the data (Byrne, 2012). The value of RMSEA ranging between .05 and .08 would indicate a close fit to a fair fit (Bowen & Guo, 2012). Factor loadings and latent factor correlations would be examined. Factor loadings show the relations of each underlying construct (each hypothesized construct) with each of the observed variables (the survey items). The higher the factor loading is, the more reliable the item is to measure the targeted factor or construct. The latent factor correlations show the associations of the latent constructs, which should be clearly smaller than 1 to be differentiable from each other; otherwise, they should be treated as a single factor.

For those measurements which would be compared across groups, invariance tests were also conducted using structural equation modelling (Byrne, 2012) to ascertain measurement equivalence across groups (i.e., experimental vs. control groups). The baseline was a configural invariant model in which the structure of the measurement model would fit the multiple groups. Then tested was the metric invariant model which would constrain the factor loadings across groups to be equal. Scalar invariance would further constrain the intercept to be the same across groups. Intercept equivalence ensures that the group scores can be statistically compared. The differences in fit statistics were used to evaluate the invariance tests, and the change in CFI should be less than .01 and the change in RMSEA should be within .015 for evidence of invariance (Cheung & Rensvold, 2002). In essence, these measurement invariance tests examine whether the respondents in different groups (e.g., experimental vs. control groups) interpret the constructs in a conceptually similar way. In other words, an invariant instrument across groups indicates that it measures the same construct across multiple groups of teachers.

**Mean comparison.** When the instrument is found invariant across groups, we can confidently compare the scores of the groups and determine whether they are statically similar or different. Group comparisons may be conducted with traditional analysis of variance (ANOVA) procedures. Where longitudinal data are available, a within-subject design may be used such that each individual is matched across time points for a more solid scrutiny of change over time.

## Results

### Preliminary Analysis

Descriptive statistics for the constructs examined in this research are reported in Table 1. Cronbach's Alpha reliability estimates for constructs with multiple measured items are also reported in Table 1.

**Table 1. Constructs – Descriptive Statistics and Alpha Reliability**

| Construct         | Time          | Items  | N    | Mean | SD   | Alpha |
|-------------------|---------------|--|------|------|------|-------|
| Attitude          | 0             | 4  | 3361 | 4.02 | 0.85 | .83   |
| Confidence        | 0             | 4  | 3361 | 3.90 | 0.89 | .85   |
| Interest          | 0             | 3  | 3361 | 4.08 | 0.83 | .78   |
| Emotion           | 0             | 4  | 3361 | 2.88 | 1.07 | .84   |
| Attitude          | 1             | 4  | 3185 | 4.24 | 0.76 | .86   |
| Confidence        | 1             | 4  | 3185 | 4.14 | 0.77 | .87   |
| Interest          | 1             | 3  | 3185 | 4.33 | 0.70 | .81   |
| Emotion           | 1             | 4  | 3185 | 2.80 | 1.09 | .86   |
| Attitude          | 2             | 4  | 2630 | 4.36 | 0.68 | .87   |
| Confidence        | 2             | 4  | 2630 | 4.24 | 0.76 | .90   |
| Interest          | 2             | 3  | 2630 | 4.49 | 0.64 | .87   |
| Emotion           | 2             | 4  | 2630 | 2.73 | 1.17 | .91   |
|                   | <b>Source</b> | <b>Sample item</b>   |      |      |      |       |
| <b>Skills</b>     | Teacher       | The child can use different ways to overcome difficulties in a science inquiry |      |      |      |       |
| <b>Attitude</b>   | Teacher       | The child likes to find the answer to a science problem by themselves          |      |      |      |       |
| <b>Confidence</b> | Teacher       | The child is full of confidence, learning Science Inquiry very fast            |      |      |      |       |
| <b>Interest</b>   | Teacher       | The child likes Science Inquiry  |      |      |      |       |
| <b>Emotion</b>    | Student       | I feel anxious during science inquiry  |      |      |      |       |

**Instrument Validation**

To validate the hypothesized constructs, a series of confirmatory factor analysis (CFA) was conducted (Tables 2 and 3). Model 1 testing a 5-factor model (Skills, Attitude, Confidence, Interest, and Emotion) resulted in a good fit (CFI=.977, TLI=.972, RMSEA=.042)(Table 2). The solution of Model 1 is reported in Table 3. The Alpha reliability estimates are good (all>.70). The students are positive in their self-assessed skills ( $M=3.29$  on a 1-5 scale), attitude assessed by their teacher ( $M=4.02$ , also above the mid-point of the scale), self-reported confidence ( $M=3.90$ ), and interest ( $M=4.08$ ), both above the mid-point of the scale. However, they are low in negative emotion ( $M=2.88$ ) which is encouraging as Emotion items referred to negative feelings such as fear, anxiety, etc.

The factor loadings are reasonable, all being well above .50. Positively worded constructs (Skills, Attitude, Confidence, and Interest) are positively correlated with one another, but they are either negatively correlated with negative Emotion ( $rs=-.051$  and  $-.008$  respectively for Skills and Confidence) or show very weak (close to zero) correlations ( $rs=.070$  and  $.094$  respectively for Attitude and Interest). This pattern of latent correlations supports the discriminant validity of the instrument.

**Testing equivalence of T0 measures between conditions of students.** It is important to first establish equivalence of the two conditions (Experimental vs. Control) at the baseline (T0) such that any change at later time-points (e.g., T1, T2) can be interpreted in terms of intervention effects. Subsequent analyses can then be conducted to test whether the experimental and control groups were similar or different in the mean scores of constructs at other time points.

A series of invariance tests between two conditions (experimental vs. control) was conducted. Model S2 testing configural invariance had a good fit (CFI=.968, TLI=.962, RMSEA=.049). Model S3 testing metric invariance also had a good fit (CFI=.967, TLI=.963, RMSEA=.049). Between Models S2 and S3, the change in CFI (.968-.967=.001) was less

than .01, and in RMSEA (.049-.049=.000) was less than .015. Further, Model S4 testing scalar invariance also had a good fit (CFI=.966, TLI=.963, RMSEA=.049). Comparing Model S4 (scalar) with Model S3 (metric), the change in CFI (.967-.966=.001) was less than .01, and there was no change in RMSEA, both=.049 (change was less than .015). These results satisfied Cheung and Rensvold's (2002) requirements for condition (experimental vs. control) invariance. This result provides a strong basis for conducting and interpreting group comparisons with confidence.

**Table 2. CFA Models**

| <i>Construct Validation</i>   |          |           |      |      |       |
|---|----------|-----------|------|------|-------|
|   | $\chi^2$ | <i>Df</i> | CFI  | TLI  | RMSEA |
| Model 1. 5-factor model (Skills, Attitude, Confidence, Interest, Emotion; <i>N</i> =3287)       | 972.41   | 142       | .977 | .972 | .042  |
| Model 2. Condition invariance: Configural (Experimental <i>N</i> =1206; Control <i>N</i> =2081) | 1424.26  | 284       | .968 | .962 | .049  |
| Model 3. Condition invariance: Metric (Experimental <i>N</i> =1206; Control <i>N</i> =2081)     | 1476.08  | 298       | .967 | .963 | .049  |
| Model 4. Condition invariance: Scalar (Experimental <i>N</i> =1206; Control <i>N</i> =2081)     | 1533.92  | 312       | .966 | .963 | .049  |

**Table 3. CFA Solution for T0**

| <i>Construct Validation of Student Model 1</i> |        |          |            |          |                  |
|--|--------|----------|------------|----------|------------------|
| Factor loadings                                | Skills | Attitude | Confidence | Interest | Negative Emotion |
| Alpha  | .92    | .83      | .85        | .78      | .84              |
| Mean   | 3.29   | 4.02     | 3.90       | 4.08     | 2.88             |
| <i>SD</i>                                      | (0.92) | (0.85)   | (0.89)     | (0.83)   | (0.27)           |
| Item 1   | .801** | .719**   | .749**     | .652**   | .706**           |
| Item 2   | .926** | .740**   | .783**     | .802**   | .793**           |
| Item 3   | .919** | .696**   | .744**     | .778**   | .792**           |
| Item 4   | .813** | .783**   | .815**     | --       | .715**           |
| Correlations                                   |        |          |            |          |                  |
| Skills   | --     |          |            |          |                  |

|                |        |        |        |        |    |
|----------------|--------|--------|--------|--------|----|
| Attitude       | .503** | --     |        |        |    |
| Cognitive Self | .470** | .871** | --     |        |    |
| Affective Self | .462** | .876** | .871** | --     |    |
| Emotion        | -.051* | .070*  | -.008  | .094** | -- |

Note: \* $p < .05$ . \*\* $p < .001$ .

### Group Comparisons

Scale scores were calculated by averaging the item scores for each of the validated constructs. Group comparisons were then conducted with the scale scores applying analysis of variance (ANOVA).

### Comparing T0, T1, T2 Across Groups

To examine intervention effects, a repeated-measures ANOVA for five constructs (Skills, Attitude, Confidence, Interest, and Negative Emotion) with longitudinal data collected at three time points (T0, T1, and T2) was conducted. The focus of the 2 (group) x 3 (time; within-subjects) ANOVA is the potential of any interaction effect. Any noteworthy interaction effect detected would indicate some changes that may be attributed to the intervention. The results are reported in Table 4.

For Skills, the main effect of condition is statistically significant,  $p < .001$ ; the main effect of time is also statistically significant,  $p < .001$ . More importantly, the condition x time interaction effect is significant,  $p < .001$ . As shown in Table 4, while both conditions improved over time, the experimental group's skills had a greater increase from T0 ( $M=3.26$ ) to T1 ( $M=3.83$ ), a 17.48% increase, than the control group ( $M$ s from 3.32 to 3.64) respectively, 9.64%. Figure 2 shows that the difference in improvement is pronounced at T1, after which there is no obvious additional improvement for the experimental group. Despite the evidence of a positive intervention effect on student skills, the effect is small ( $\eta^2 = .01$ ).

For Attitude, the main effect of condition is not statistically significant,  $p > .05$ . The main effect of time is statistically significant,  $p < .001$ , indicating improvement for both conditions over time. The condition x time interaction effect is significant,  $p < .001$ . As shown in Table 4, while both conditions improved over time, the level of attitude that the experimental group reached at T1 ( $M=4.32$ ) required an additional 3 months for the control to reach at T2 ( $M=4.33$ ). Figure 3 shows that the difference in improvement is pronounced at T1, but the line flattens thereafter. Again, despite evidence of intervention effect, the effect size is small ( $\eta^2 = .01$ ).

For Confidence, the main effect of condition is not statistically significant,  $p > .05$ . The main effect of time is statistically significant,  $p < .001$ , indicating improvement for both conditions over time. The condition x time interaction effect is significant,  $p < .001$ . As shown in Table 4 and Figure 4, while both conditions improved over time, the two groups did not differ at T1 or T2, and the difference between groups were observed with the experimental group being relatively lower at T0 but improved at T1) Again, despite evidence of intervention effect, the effect size is small ( $\eta^2 = .01$ ; Table 4).

For Interest, the main effect of condition is not statistically significant,  $p > .05$ . The main effect of time is statistically significant,  $p < .001$ , indicating improvement for both conditions over time. The condition x time interaction effect is significant,  $p < .001$ . As shown in Table 4, the level of affective self-concept took another 3 months for the control group at T2 ( $M=4.43$ ) to match that of the experimental group at T1 ( $M=4.44$ ). Figure 5 shows that even though students in the experimental group started significantly lower ( $M=4.01$ ) than the

control group ( $M=4.16$ ), the improvement in the experimental group is noteworthy ( $p<.001$ ,  $\eta^2=.03$ ). Furthermore, although it was noted earlier that caution may be needed in interpreting findings for Interest, this result indicates that whereas the experimental students' low interest at T0 may be puzzling, the positive effect of the intervention for interest is undoubtedly impressive.

For Emotion (noting that this refers to negative feelings), the main effect of condition is not statistically significant,  $p>.05$ . The main effect of time and the condition x time interaction effects are both significant,  $p<.001$ . As shown in Table 4 and Figure 6, while the control condition remains similar across time points ( $Ms=2.88, 2.87, 2.89$  respectively), the experimental group shows continual decreases ( $Ms=2.88, 2.67, 2.50$  respectively), which are statistically significant and noteworthy ( $p<.001$ ,  $\eta^2=.04$ ).

**Table 4. Students – Condition (C) x Time (T) Repeated-Measures ANOVA Results**

| Condition (C) |           | Experimental |        |        | Control |        |        | ANOVA               |                     |                     |              |
|---------------|-----------|--------------|--------|--------|---------|--------|--------|---------------------|---------------------|---------------------|--------------|
| Time (T)      |           | T0           | T1     | T2     | T0      | T1     | T2     | Condition           | Time                | CxT                 | CxT $\eta^2$ |
|               | <i>N</i>  | 935          | 935    | 935    | 1531    | 1531   | 1531   | <i>F</i>            | <i>F</i>            | <i>F</i>            |              |
| Skills        | <i>M</i>  | 3.26         | 3.83   | 4.11   | 3.32    | 3.64   | 3.91   | 12.73**             | 1078.62**           | 44.28**             | .01          |
|               | <i>SD</i> | (0.91)       | (0.76) | (0.73) | (0.95)  | (0.88) | (0.86) | ( <i>MSE</i> =0.55) | ( <i>MSE</i> =0.31) | ( <i>MSE</i> =0.31) |              |
|               | <i>N</i>  | 1003         | 1003   | 1003   | 1527    | 1527   | 1527   |                     |                     |                     |              |
| Attitude      | <i>M</i>  | 3.97         | 4.32   | 4.42   | 4.09    | 4.24   | 4.33   | 0.52                | 292.65**            | 31.83**             | .01          |
|               | <i>SD</i> | (0.87)       | (0.68) | (0.64) | (0.83)  | (0.78) | (0.70) | ( <i>MSE</i> =0.43) | ( <i>MSE</i> =0.28) | ( <i>MSE</i> =0.28) |              |
| Confidence    | <i>M</i>  | 3.82         | 4.13   | 4.27   | 3.97    | 4.16   | 4.22   | 1.91                | 284.79**            | 21.93**             | .01          |
|               | <i>SD</i> | (0.88)       | (0.70) | (0.70) | (0.89)  | (0.81) | (0.79) | ( <i>MSE</i> =0.46) | ( <i>MSE</i> =0.30) | ( <i>MSE</i> =0.30) |              |
| Interest      | <i>M</i>  | 4.01         | 4.44   | 4.57   | 4.16    | 4.29   | 4.43   | 3.03                | 414.08**            | 67.73**             | .03          |
|               | <i>SD</i> | (0.81)       | (0.58) | (0.57) | (0.84)  | (0.74) | (0.67) | ( <i>MSE</i> =0.34) | ( <i>MSE</i> =0.28) | ( <i>MSE</i> =0.28) |              |
| Emotion       | <i>M</i>  | 2.88         | 2.67   | 2.50   | 2.88    | 2.87   | 2.89   | 1.81                | 28.70**             | 91.09**             | .04          |
|               | <i>SD</i> | (1.05)       | (1.09) | (1.11) | (1.05)  | (1.04) | (1.18) | ( <i>MSE</i> =0.80) | ( <i>MSE</i> =0.65) | ( <i>MSE</i> =0.34) |              |

Note: \* $p < .05$ . \*\* $p < .001$ .

Figure 2. Student skills in 2 conditions at 3 time points

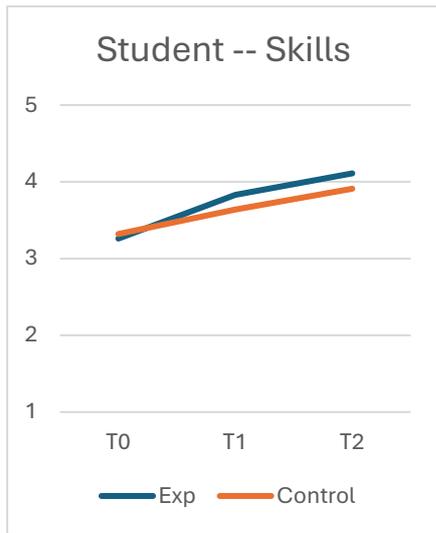


Figure 3. Student attitude in 2 conditions at 3 time points

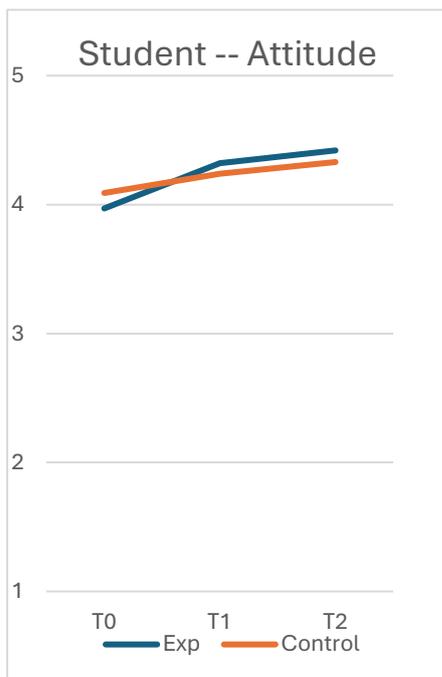


Figure 4. Student confidence in 2 conditions at 3 time points

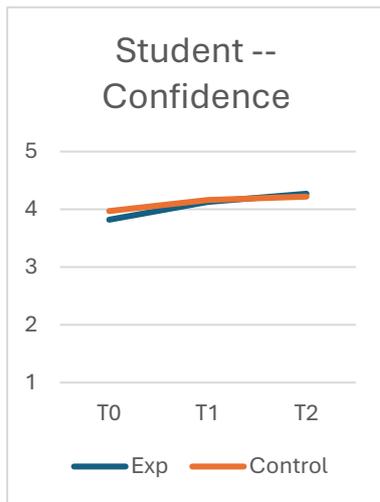


Figure 5. Student interest in 2 conditions at 3 time points

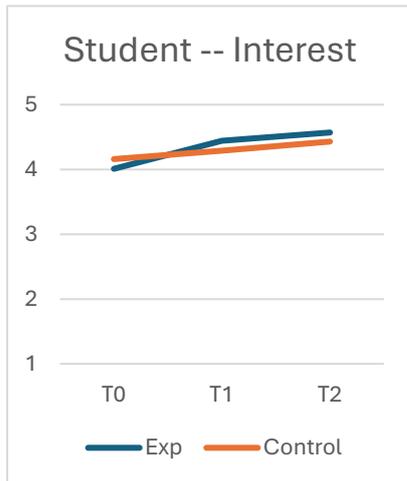
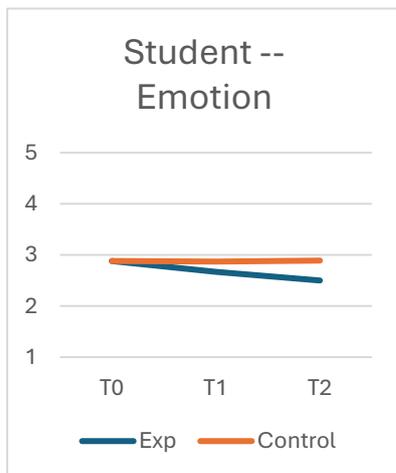


Figure 6. Student negative emotion in 2 conditions at 3 time points



## Discussion

The research attempted to answer one research question: Do Kindergarteners in the experimental group improve more than kindergarteners in the control group in multiple outcomes after the SIP? The answer to the RQ is clearly positive, indicating that the program is successful, evidenced by the intervention effects for all student outcomes measured. Although the effects do not seem astounding, the statistical analyses have provided clear evidence of success, given the light dose of intervention: only 1 session of 180 minutes for teacher professional development, and 2 sessions of 30 minutes (total 60 minutes) for program implementation in class. Teachers benefited from the professional development in terms of STEM teaching skills, and they witnessed their students' gains in STEM skills attributed to the program they conducted in class.

The findings from this research have contributed to the discourse on STEM in early childhood education, providing an insight to optimising the benefits of STEM education by first guiding early childhood educators to design, plan, and implement an SIP (see Aldemir & Kermani, 2017). The results highlight the positive effects of Science Inquiry Program for young children which has not been clearly illustrated in previous studies (Stephenson Reaves et al., 2022; Yang et al., 2024; Yüceliyigit & Toker, 2021). The research design collecting multivariate data at multiple timepoints has also added rigour and credibility to the demonstration of how young children benefit from an SIP and why STEM inquiry should start early in early childhood. The focus on skills and the cultivation of students' scientific spirit and scientific learning steps, and promotion of curiosity, open-mindedness, and free exploration is the essence of the current SIP (Chou, 2020; Curriculum Development Council, 2017) and this focus has proved to work for the young learners as well as their teachers implementing the program. Hence, evidence of success points to the merit of this focus.

Despite the evidenced success, the program may be further improved by strengthening the dosage and including self-concept enhancement to yield stronger and sustainable long-term intervention effects (Kadir et al., 2020). As the SIP was designed to ignite a lifelong passion for STEM in Kindergarteners by fostering curiosity, hands-on problem-solving, and a scientific mindset, we can envisage positive effects of our SIP approach would lead to measurable improvements in exploratory skills, critical thinking, and environmental awareness. This dynamic program empowers young learners to confidently engage with the world around them, building a strong foundation for future STEM success. To achieve the best effects, STEM inquiry programs must begin in early childhood education to cultivate the

essential skills, attitude, confidence, and interest and minimise negative emotions from the earliest stages of learning. In essence, STEM inquiry must begin as early as possible.

### **Conclusion and Recommendations**

The findings of this research indicate that scientific inquiry programs can and should start in early childhood education. SIP activities can benefit all learners and evidence shows that a well-designed SIP can benefit young learners in preschool education. Based on our findings, the following recommendations are made:

1. The “Exploratory (STEM) curriculum” (Ng, 2020, 2024) can be implemented in the preschool stage and should be implemented as early as possible. Research has shown that children’s inquiry skills, inquiry attitudes, inquiry confidence, and interest in inquiry are significantly improved after implementing the curriculum, while negative learning emotions are significantly reduced. Therefore, young learners are ready and capable of constructing learning from an appropriately structured STEM inquiry. Schools should adopt this curriculum across all Kindergarten levels to ensure consistent exposure and skills development.
2. This study only implemented one of the exploration activities of the "Scientific Exploration Activities " in the Science Inquiry Program. It is recommended to expand the research to observe the continual development of the experimental group while allowing the waitlist control group to fully participate in the implementation of the curriculum, and then collect research data for analysis to fully consolidate the intervention effects irrespective of intervention sequence.
3. Further research may focus on the other two untested areas of “Inquiry-based (STEM) Curriculum” (free construction and free play, exploratory learning activities) in school, and compare the intervention effects with those in the present investigation to provide further insight into various modes and approaches to STEM education for young children. This extension of research may provide us with a more comprehensive understanding of the best way to obtain the best intervention effects.
4. Follow the principle of “training first, then practice” so that teachers can have a full understanding and confidence in the content and skills of leading exploration activities (STEM courses) before engaging in other related programs. This is a win-win situation for both teachers and children in terms of learning and teaching.
5. SIP activities should be designed based on the principles of age-appropriate, scope-appropriate, and level-appropriate, such as the "inquiry-based (STEM) curriculum"

(Ng, 2020, 2024), in order to effectively promote children's knowledge, skills, attitudes, self-confidence, and interest.

6. It is recommended to involve different stakeholders in school-based exploratory activities (STEM courses) so that schools can better understand the ideas of different stakeholders, complement each other's ideas, and make the development of learning and teaching more comprehensive.

8. Researchers should collect both quantitative data for objective analysis and qualitative data for complementary explanations for why and how the SIP was successful. The research should also adopt a longitudinal design to substantiate the benefits of SIP over time compared to a control group that did not receive the intervention.

## References

- 伍寶珠(2024)。STEM教育的迷思。《博士童行——經濟日報》，7/10/2024。
- 伍寶珠(2020)。適齡適量適度的學習。《博士童行——經濟日報》，4/11/2020。
- 周淑惠(2020)。《幼兒STEM教育：課程與教學指引》。台北：心理。
- 課程發展議會(2017)。《學前教育課程指引》。香港：教育局。
- Aldemir, J., & Kermani, H. (2017). Integrated STEM curriculum: Improving educational outcomes for head start children. *Early Child Development & Care*, 187(11), 1694-1706. <https://doi.org/10.1080/03004430.2016.1185102>
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888–918. <https://doi.org/10.1037/0033-2909.84.5.888>
- Bowen, N. K., & Guo, S. (2012). *Structural equation modeling*. Oxford University Press.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus*. Taylor and Francis.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Choy, M. W. C., & Yeung, A. S. (2022). Cognitive and affective academic self-concepts: Which predicts vocational education students' career choice? *International Journal of Education Research Open*. <https://doi.org/10.1016/j.ijedro.2022.100123>
- Choy, M. W., & Yeung, A. S. (2024). Tourism vocational education: Relations of input and output. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 34, 100479. <https://www.sciencedirect.com/science/article/pii/S1473837624000029>
- Fuller, J. A., Luckey, S., Odean, R., & Lang, S. N. (2021). Creating a diverse, inclusive, and equitable learning environment to support children of color's early introductions to STEM. *Translational Issues in Psychological Science*, 7(4), 473-486.
- Josiam, B. M., Devine, F. G., Baum, T., Crutsinger, C., & Reynolds, J. S. (2010). Attitudes to Work of Generation Y Students in Hospitality Management: a comparative analysis of students in England, Scotland and Northern Ireland. *Journal of Hospitality & Tourism Education*, 22(1), 44–53. <https://doi.org/10.1080/10963758.2010.10696968>
- Kadir, M. S., Yeung, A. S., & Diallo, T. M. O. (2017). Simultaneous testing of four decades of academic self-concept models. *Contemporary Educational Psychology*, 51, 429-446. <https://doi.org/10.1016/j.cedpsych.2017.09.008>
- Kadir, M. S., Yeung, A. S., Ryan, R. M., Forbes, A., & Diallo, T. M. O. (2020). Effects of a dual-approach instruction on students' science achievement and motivation.

*Educational Psychology Review*, 32, 571-602. <https://doi.org/10.1007/s10648-018-9449-3>

- Lee, C., Yeung, A. S., & Cheung, K. W. (2019). Learner perceptions versus technology usage: A study of adolescent English learners in Hong Kong secondary schools. *Computers & Education*, 133, 13-26. <https://doi.org/10.1016/j.compedu.2019.01.005>
- Marsh, H. W., Craven, R. G., Yeung, A. S., Mooney, J., Franklin, A., Dillon, A., Barclay, L., vanWestenbrugge, A., Vasconcellos, D., See, S-M., Roy, D., Kadir, M., & Durmush, G. (2023). Self-concept a game changer for academic success for high-achieving Australian Indigenous and NonIndigenous students: Reciprocal effects between self-concept and achievement. *Contemporary Educational Psychology*, 72, 102135. <https://doi.org/10.1016/j.cedpsych.2022.102135>
- Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology: An International Review*, 41(4), 359–376. <https://doi.org/10.1111/j.1464-0597.1992.tb00712.x>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19(1), 123-205.
- Rosseel, Y. (2023). The lavaan tutorial. <https://lavaan.ugent.be/tutorial/tutorial.pdf>
- Stephenson Reaves, J., Likely, R. & Arias, A. M.. (2022). Design principles for considering the participatory relationship of students, teachers, curriculum, and place in project-based STEM units. *Education Sciences*. 12(11), 760. <https://doi.org/10.3390/educsci12110760>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Massachusetts: Harvard University Press.
- Yang, W., Du, Y., Wu, R., & Xiang, S. (2024). Development and validation of the children's STEM Habits of Mind Questionnaire. *Early Childhood Education Journal*, 52(3), 493-501.
- Yeung, A. S., Lau, S., & Nie, Y. (2011). Primary and secondary students' motivation in learning English: Grade and gender differences. *Contemporary Educational Psychology*, 36, 246-256. <https://doi.org/10.1016/j.cedpsych.2011.03.001>