Michelle N. Maurer
Frikartweg 6
3006 Bern
michelle.maurer@unibe.ch
michelle.nina@bluewin.ch

Anna Roth
Neufeldstrasse 9
3012 Bern
roth614@hotmail.com

Julian Frank
Ahornweg 7
3012 Bern
frank_julian@hotmail.com

Data Science Project

Towards a better understanding of handwriting development in the transition to school

Conceptual Design Report

30th September 2021

Abstract

Many children experience difficulties acquiring written language. Handwriting skills are meaningfully related to early academic achievement and psychosocial health. With the proposed project, we aim to identify early predictors of handwriting development. In a longitudinal study over three consecutive years, we assessed fine motor skills in kindergarten and investigated its predictive value for handwriting fluency, cognitive functions (i.e., executive functions), and visual-motor integration in first grade. Our main research question was whether fine motor skills in kindergarten predict handwriting in first grade. Data based on a sample of n = 118 children aged 4-9 were analyzed using bivariate correlations and linear regressions. First results revealed that fine motor skills in kindergarten may indeed predict handwriting in first grade. Further, we find rather weak relationships between handwriting fluency, executive skills and fine motor skills (when measured in the same year). It is likely that fine motor skills provide a playground to develop and practice cognitive skills. However, the specific underlying mechanisms are still widely unknown.

Table of Contents

1 Project Objectives

Up to 30% of school-age children suffer from handwriting difficulties [1]. This high prevalence is worrying because children with handwriting problems have a higher risk of lower academic achievement [2], social-emotional problems [3], anxiety, and depression [4], and adverse employment consequences. These factors may negatively affect the economy [5].

Studies have shown that handwriting difficulty meaningfully impacts children’s learning. For instance, a child who struggles to write digits legibly and align them neatly will probably work longer and harder on the math homework and learns less than a child with typical writing skills [6].

The contemporary literature points to the crucial role of specific motor skills (e.g., visual-motor integration, fine motor skills) and specific cognitive functions (i.e., executive functions, EF) for handwriting development. As proposed by the theoretical framework of embodied cognition, motor processes play a central role in building and strengthening cognitive functions [e.g. 7]. These theoretical assumptions are supported by evidence indicating that children’s motor skills and EF are meaningfully associated with their concurrent and prospective writing skills [e.g. 8]. However, the critical skills related to literacy development are still widely unknown.

With the proposed project, we aim to identify predictors of handwriting development. If we better understand the factors related to handwriting development, we can better recognize signs of delayed or divergent writing development. This will allow deriving hypotheses about enriching children’s learning environments, adapting instructional activities to meet their individual needs, and preventing these children from adverse life outcomes.

The aim of this project is to answer the following research questions:

1) Are there cross-sectional associations (bivariate correlations) among fine motor skills, visual-motor integration, and executive functions in first grade?

2) Do fine motor skills in kindergarten predict handwriting (fluency and time on paper) in first grade (linear regression)?

3) Do fine motor skills in kindergarten predict executive functions in first grade (even when controlling for concurrent fine motor skills in first grade?

2 Methods

The data was collected through various experiments with children in kindergarten and first grade. It contains information on fine motor skills, ability to write by hand and cognitive abilities of children observed at two time points. In order to answer the question described above, the data is read in Excel files with the help of Google Colab and Python and processed into data frames. To do so, and to be able to work together on the same files, we decided to use Google Drive. For our project, we will use several Python libraries, such as Pandas, Numpy, Statistics, Matplotlib, Math, Scipy, and Statsmodels. The following points briefly describe the most important libraries:

- Pandas is mainly to support data analysis and it works with tabular data. With this library, it is easy to import data from different file formats. Also facilitates various data manipulation operations such as merging, reshaping, selecting, as well as data cleaning features [9].
- Numpy is mainly to perform mathematical operations. This library is required to operate on the data frames created by pandas [10].
- Statistics is mainly used for calculations such as averages, measures of central location, and spread. This library also provides functions for numeric data [11].
- Matplotlib is mainly to create static, animated, and interactive visualizations [12].

In order to gain a first overview of the data, some descriptive analyses are performed (first results in chapter 3). Following this, the above hypotheses will be tested using inferential statistics. Correlations between the variables of interest will be examined using parametric or non-parametric correlation analyses, depending on the results of the prerequisite testing. After testing the preconditions for computing a linear regression, research questions are tested with regression analyses. An advantage of a regression analysis may be that it is more appropriate because it is easier to control for multiple influencing factors or to make simple predictions.

3 Data

We work with five different datasets of about 118 children. The average age is seven, the oldest child is eight years old, and the youngest is six years old. 54% are male, as can be seen in Table 1.

Table 1: Summary statistics about the children participating in the experiments

	ID	Geschlecht	Alter_Monate	Alter_Jahre
count	118.00	118.00	118.00	118.00
mean	1524.25	0.54	88.75	6.92
std	291.89	0.50	5.26	0.54
min	1104.00	0.00	80.00	6.00
25%	1216.25	0.00	85.00	7.00
50%	1608.50	1.00	88.00	7.00
75%	1805.75	1.00	92.00	7.00
max	1917.00	1.00	103.00	8.00

Each child has an ID that uniquely identifies every child in all five datasets.

First dataset named Maulwurf: Data about an experiment to test children's executive functions (EF; i.e., working memory). Children's visual-spatial working memory was assessed with this task which was adapted for children [13]. Embedded in a cover story, a mole appeared at different locations in a 4 × 4 grid. Children were asked to memorize the locations where the mole had appeared and to touch these fields in reverse order. The items appeared in a fixed pseudo-randomized order. The task started with two items, and the sequence length was increased by one if at least three out of six sequences (50%) were remembered correctly. The task was terminated at the end of a specific span length if more than three out of six sequences were incorrectly recalled. The total number of correctly recalled sequences across sequence lengths was used as a measure for visual-spatial working memory.

To answer our research questions, the following three variables are relevant:

- spanLength: This variable describes the maximum achieved working memory span a child achieved. The data type is discrete numerical and the measurement level is ordinal.
- Recognition_fields_correct: Sum of correctly recalled locations (e.g., fields where the mole appeared in the grid) in the correct sequence. The data type is discrete numerical and the measurement level is ratio.
- recognition_fieldsCorrect_irrespectivePosition: Sum of correctly recalled locations (e.g., fields where the mole appeared in the grid) while the sequence/order of fields is not considered. The data type is discrete numerical and the measurement level is ratio.

Some summary statistics about the dataset are provided in the table below.

Table 2: Summary statistics about the "Maulwurf" dataset

	spanLength	recognition_fields_correct	recognition_fieldsCorrect_irrespectivePosition
count	116.00	116.00	116.00
mean	2.79	30.35	42.54
std	0.97	15.67	16.65
min	0.00	8.00	15.00
25%	2.00	20.75	33.00
50%	3.00	27.50	38.00
75%	3.00	40.25	53.00
max	6.00	82.00	110.00

Second dataset named Hearts and Flowers: Data about an experiment to test children’s executive functions (EF; i.e., Inhibition and Shifting). The congruent block of trials of the Hearts & Flowers task [14] was used to measure children’s inhibition. Inhibition describes the ability to inhibit initial impulses or predominant responses. The mixed block of the Hearts and Flowers task was used to assess the ability to shift attention flexibly. Shifting refers to the ability to flexibly adjust to changing conditions such as to switch rules [15]. Mean reaction times of correct responses and the percentage of correct responses (accuracy) were recorded to measure inhibition (congruent block) and shifting (mixed block).

Four values are relevant to calculate children’s **inhibition**:

- correct_FGT: The number of correct trials. The data type is discrete numerical and the measurement level is ratio.
- trials_FGT: The maximal amount of trials needed. The data type is discrete numerical and the measurement level is ratio.
- RT_inhibition_FGT: The average reaction time of the correct trials. The data type is continuous numerical and the measurement level is ratio.
- accuracy_inhibition_FGT: Correct trials divided through all trials to get the success rate, what measures the accuracy. The data type is continuous numerical and the measurement level is ratio.

Four values are relevant to calculate children’s **shifting**:

- correct_MG: The number of correct trials. The data type is discrete numerical and the measurement level is ratio.
- trials_MG: The maximal amount of trials needed. The data type is discrete numerical and the measurement level is ratio.
- RT_shifting_MG: The average reaction time of the correct trials. The data type is continuous numerical and the measurement level is ratio.
- accuracy_shifting_MG: Correct trials divided through all trials to get the success rate, what measures the accuracy. The data type is continuous numerical and the measurement level is ratio.

Some summary statistics about the dataset are provided in the table below.

Table 3: Summary statistics about the "Hearts and flowers" dataset

	correct_MG	trials_MG	RT_shifting_MG	accuracy_shifting_MG	correct_FGT	trials_FGT	RT_inhibition_FGT	accuracy_inhibition_FGT
count	117.00	117.0	90.00	117.00	117.00	117.0	109.00	117.00
mean	38.79	60.0	915.95	0.65	30.02	36.0	789.89	0.83
std	21.78	0.0	445.91	0.36	9.46	0.0	217.79	0.26
min	0.00	60.0	449.46	0.00	0.00	36.0	415.25	0.00
25%	40.00	60.0	698.02	0.67	31.00	36.0	646.00	0.86
50%	49.00	60.0	818.90	0.82	34.00	36.0	766.22	0.94
75%	53.00	60.0	993.91	0.88	35.00	36.0	873.52	0.97
max	60.00	60.0	3947.90	1.00	36.00	36.0	1866.08	1.00

Third data set handwriting: In this dataset is data about an experiment on children's handwriting in first grade. The task was for each child to write four different sentences on a tablet. This allows us to measure the fluency of pen movements (i.e., degree of automaticity), how long the children need to write the sentences, how much pressure they apply, how much time the pen is on the tablet, and how much time the pen is in the air.

To answering our research questions, the following three variables are relevant:

- Timeon: Time the pen is on the tablet. The data type is continuous numerical and the measurement level is ratio.
- Timeoff: Time the pen is off the tablet. The data type is continuous numerical and the measurement level is ratio.
- NIV: Degree of automation (Average number of motion inversions) where 1 means perfectly automated. The data type is discrete numerical and the measurement level is ordinal.
- Press: Medium writing pressure. Degree of automation (Average number of motion inversions) where 1 means perfectly automated. The data type is discrete numerical and the measurement level is ordinal.

Some summary statistics about the dataset are provided in the table below.

Table 4: Summary statistics about the "handwriting" dataset

	TIMEON	TIMEOFF	NIV	PRESS
count	117	117	117	117
mean	21779	37332	5	1
std	7068	13441	2	0
min	9820	9430	1	0
25%	16091	28314	3	1
50%	20424	35239	4	1
75%	26732	44600	6	1
max	41003	79901	12	2

Fourth data set fine motor skills kindergarten: In this experiment children had to color and draw certain shapes. Based on this task, their fine motor skills were evaluated.

To answer our research questions, the following two variables are relevant:

- FineMotor_2018: Fine motor skills in 2018. The data type is discrete numerical and the measurement level is ordinal.
- FineMotor_2019: Fine motor skills in 2019 with the same children. The data type is discrete numerical and the measurement level is ordinal.

Some summary statistics about the dataset are provided in the table below.

Table 5: Summary statistics about the "motor skills in kindergarten" dataset

	ID	FineMotor_2018	FineMotor_2019
count	94.00	83.00	83.00
mean	1521.55	7.30	8.41
std	296.16	2.32	1.93
min	1104.00	2.00	2.00
25%	1217.25	6.00	8.00
50%	1605.50	7.00	9.00
75%	1807.75	9.00	10.00
max	1917.00	11.00	10.00

Fifth data set fine motor skills first grade: In this dataset is data about an experiment on children’s fine motor skills in first grade. In order to make a statement here, the children had to complete various tasks. For example, the children were given the task of tracing a curved and an angular shape, folding paper, and cutting out shapes. An average value was calculated from all these values using point values. Further, visual-motor integration was measured in several experiments. Here, too, the average value over 13 different tasks was used for further analyses.

To answering our research questions, the following three variables are relevant:

- Mean_VMI: Average measure of visual-motor integration. The data type is continuous numerical and the measurement level is ratio.
- FeinmotorikFirstGradeSumme: Sum of the point values over all fine motor skill experiments. The data type is discrete numerical and the measurement level is ordinal.

Some summary statistics about the dataset are provided in the table below.

Table 6: Summary statistics about the "motor skills in first grade" dataset

	ID	Mean_VMI	FeinmotorikFirstGradeSumme
count	48.00	48.00	48.00
mean	1419.79	1.13	30.17
std	255.74	0.19	4.81
min	1201.00	0.64	18.00
25%	1215.50	1.03	27.50
50%	1307.50	1.14	30.00
75%	1802.25	1.29	33.25
max	1816.00	1.47	40.00

4 Metadata

In this paragraph, we describe the required metadata to reproduce the analyses. To reproduce the project, we need to know detailed metadata on the sample characteristics, as findings from the analyses are specific to this sample and cannot necessarily be generalized to the population. Furthermore, information on how we collected the data is required. In addition, the specific tasks (measures) and the dependent variables included in the analyses are required to reproduce the project. Finally, metadata on the storage and accessibility of the data is necessary.

4.1 What is the metadata of the sample?

We approached school administrators and teachers of the city of Bern (Switzerland) for the project. We selected children for participation in the project if the head of the school, the teacher, and the parents, and the children themselves agreed to participate. Children only participated in the study if their German skills were sufficient to follow task instructions. Boys and girls were equally represented in the sample.

The longitudinal data of this project was collected at three different time points in children's development. Children's fine motor skills were assessed in kindergarten at two time points, in autumn 2018 and autumn 2019. We collected all remaining data (executive functions, handwriting, fine motor, and visual-motor integration skills) in spring 2021 when children were in first grade.

4.2 How did we collect the data and which variables did we use?

The data was collected in the respective schools of the children who participated in the study. We assessed children's fine motor skills, visual-motor integration, and executive functions in the class. The children were instructed to each task orally through the project team or previously recorded audio instructions. Children solved each task individually.

We used the Hearts & Flowers task and mole task ("Maulwurf") to assess children's cognitive performance on the executive function tasks. These tasks were programmed with the App Ionic and presented on a tablet computer. Children listened to standardized task instructions and feedback on their practice trials through headphones. The data on the tasks of the executive function consists of reaction times in milliseconds and accuracies, that is, the number of correct trials divided by the total number of trials.

We assessed data on children's fine motor skills in first grade with the subscale "fine motor precision" of the standardized motor test battery BOT-2 [16]. The subscale fine motor precision consists of four paper and pencil tasks and one task in which children had to fold the paper, and one in which they cut out a circle.

Kindergarten teacher measured children's fine motor skills in kindergarten with the Developmental Screening DESK 3-6 R [17]. The measure included in the study is the total score on the subscale fine motor skills. This subscale consists of 10 different fine motor tasks. We get access to this data retrospectively after having received ethical approval from all involved parties.

We assessed children's handwriting in small groups of 6 subjects or less in a separate, quiet room in the respective school. Children were presented with four sentences from the standardized writing test SLRT-II [18] in the writing task. Children were asked to copy these sentences on a white, unlined paper that was lying on top of a graphic tablet WACOM Intuos Pro. The pen was an Induction pen with a regular ballpoint which allowed to record the processes during writing. The writing tablet was connected to the computer and the software CSWin [19]. This software calculates the temporal and spatial factors of handwriting. As a measure of the degree of handwriting fluency, that is, automaticity, we used the variable "Number of Inversions in Velocity" (NIV). The NIV is defined as the number of velocity changes per unit of motion. The fewer changes in velocity during writing, the more fluid and automated the handwriting is.

4.3 Where is the metadata stored, and how can people access it?

The tasks instructions and procedure metadata is stored in text files (word documents) and pdf documents. All remaining metadata on sample characteristics, variables included in the analyses is stored in single excel files stored on our internal server. Only the project team has access to the anonymous data.

5 Data Quality

Our research questions 3 and 4 concern the longitudinal relationship between children’s prior fine motor skills (in kindergarten) and later cognitive performance and handwriting (in first grade). Consequently, we need at least two measurement points for each child to answer those questions. Specifically, data on at least one time point in kindergarten and one time point in first grade needs to be available to answer our research question. Missing values are detrimental to the quality of the data. To prevent missing data, we try to revisit schools if a child was sick or absent on the day of testing.

Another aspect that is indirectly related to data quality and impacts the conclusions drawn from the results is the sample size. Roughly speaking, the bigger the sample size, the better we can generalize the results to a more comprehensive population. Besides the size of the sample, the sample should represent the population as well as possible (e.g., gender distribution, socioeconomic background).

Furthermore, it is important for the data quality that the tasks administered have certain reliability, objectivity, and validity, which are the central psychometric properties of a test. For the sake of objectivity, all tasks were standardized in terms of their instruction, the procedure, and the interpretation of results. In addition, task validity is assumed to be good as we used widely known tasks from standardized task batteries. Concerning reliability, we calculated interrater reliability, that is, the agreement of two independent raters, to measure the reliability of the data on visual-motor integration.

It is essential to acknowledge that we can never know precisely how good the data quality is, as the data is collected in human beings, specifically in children. We cannot ensure that all children understand the task instructions equally (for example, children who speak another first language), pay attention (e.g., distractions), and are motivated to show their best performance. We can only try to create a test environment that allows children to concentrate and provide further instructions if the task is not clear. In sum, we can only try to offer ideal conditions to achieve a high test quality. Regarding the computer-based tasks, the accuracy of the data is very high, as reaction times in milliseconds can be measured, and answers are recorded automatically as correct or not correct (binary variable).

Regarding visuomotor integration, data quality and, specifically, reliability might be lower than the other data as the scoring of this task is not completely objective but is, despite set criteria, subjective. A project member scores the copied items with 0 (not accurate), 1 (somewhat accurate), or 2 (accurate) according to specific criteria, which always leaves room for interpretation and subjectivity. However, by calculating interrater reliabilities, that is, the agreement of two independent raters on the same copies allows to estimate reliability and, therefore, quality of this measure.

Scatterplots provide the first descriptive overview of the distribution. To ensure data quality, we check for outliers in each variable of interest. To test whether there are outliers in the data, we define outliers as values exceeding three standard deviations (SD) of the mean and replace them with the value of the third SD. Furthermore, we check skewness and kurtosis indicators of each variable of interest.

6 Data Flow

Our data flow is shown in figure 1 and can be explained in simplified terms as shown below:

- Data is collected in experiments, available in five excel files described above
- Converted to python data-frames and merged to one single dataframe
- After data cleaning and plausibilisation we calculated some descriptive statistics to get to know the data, see possible relationships and create plots
- After that, we apply inferential statistics (e.g. test, regressions) to test our hypotheses

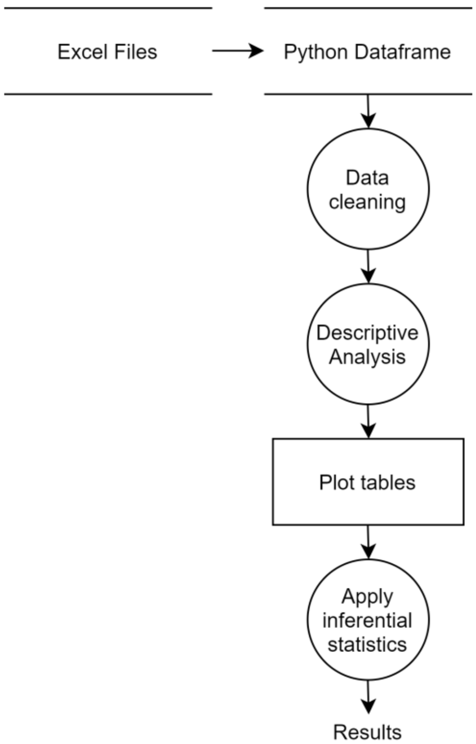

Figure 1: Data flow

7 Data Model

7.1 Conceptual data model

The goal of the conceptual model is to show the terms/entities and the relationships between them. This is in our case very simple as all the available data comes from experiments that have been conducted with children, as shown in figure 2.

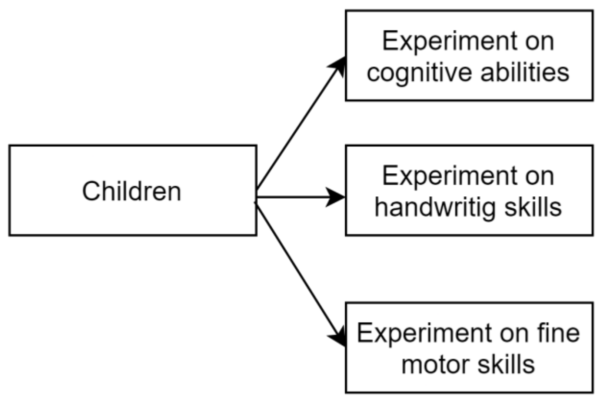

Figure 2: Conceptual data model

7.2 Logical data model

The goal of the logical data model shown in figure 3 is to show how the data sets and data files are organized. In our case, there are six different tables:

- Children: contains information about the age and gender of a child
- Executive functions:
    - Mole: Contains the results from the "mole" experiment, which tests the cognitive abilities of the children.
    - Hearts&Flowers: This contains the results from the "Hearts and flowers" experiment, which tests the cognitive abilities of the children.
- Handwriting: Contains data about processes during handwriting, such as the fluency (i.e., automaticity) of handwriting, the number of pen lifts, or the time spent with the pen on or off the paper while copying four sentences.
- Fine motor skills kindergarten: Contains data about fine motor skills based on the DESK-3-6 R Developmental Screening (Tröster et al., 2016). Data is available for 2018 and 2019, as children usually stay in kindergarten for two years.
- Fine motor skills first grade: Contains data about the same children as in the "fine motor skills kindergarten" table that shows how good their fine motor skills are one year after kindergarten (based on how good children were able to perform certain tasks like e.g. color in a star). The method to test the children's fine motor skills is not the same as in the "fine motor skills kindergarten" table.

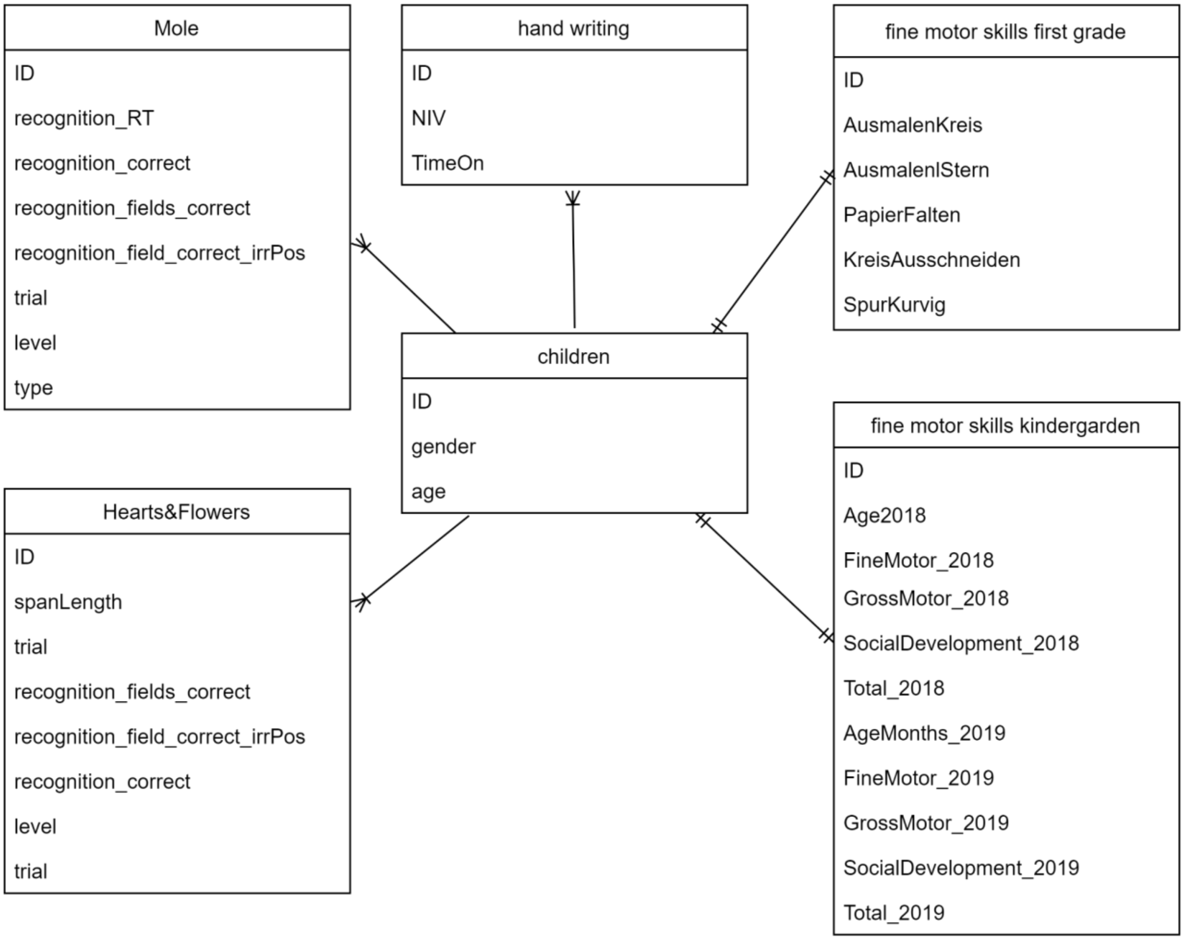

Figure 3: Logical data model

7.3 Physical data model

The data is stored in six excel sheets. One contains the data about the test persons and the other five sheets contain one dataset each. These six excel sheets are stored on a cloud server named Google Drive. To actually work with the data and to be able to combine the six sheets we used the tool Google Colab. With this approach it is possible to work with several clients on the same data, as shown in figure 4.

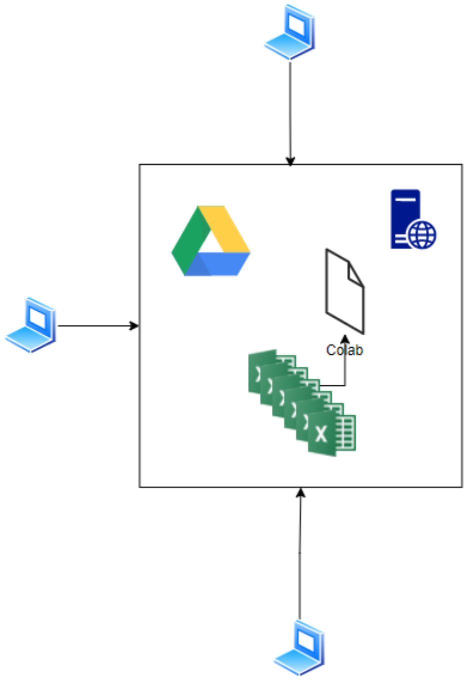

Figure 4: Physical data model

8 Risks

Each project has its risks. In this paragraph, we describe what can go wrong, what challenges we might face. In addition, we explain what countermeasures we have and how we would deal with those risks.

When collecting data in humans, we need ethical approval. In the current project, including minors, we must also have the parent’s permission to allow the child to participate in the study. Therefore, it is likely that this selective sample is biased, as children who participate and those who do not may systematically differ in certain variables. We try to inform parents about the project with a short and simply written information letter. We provide our contact details and offer to answer any questions that may come up and hinder them from giving their approval of participation. The same is true for all teachers participating in the project.

Some of the data of this project are accessed retrospectively. This implicates some risks which may impact data quality. For instance, as teachers collected the data in kindergarten, we do not know whether teachers strictly followed task instructions and procedures while collecting the data in children. Furthermore, there is the risk of missing data.

As in some schools in Biel, many children have a migration background, German is not the first language of many children. There is a risk that especially those children do not understand task instructions. We try to keep the task instructions as simple as possible and use basal German words and brief sentences. In addition, we will implement loops in the practice trials. Hence, children repeat the practice session if they do not solve the task correctly.

Another risk is missing data. The reasons for missing data can be: Children are sick or absent on the day of testing, the data is not stored correctly, or children do not want to terminate the task. We plan to revisit schools to counteract those risks if children were not present on the testing day. Such revisits imply high timely costs for the project team and might delay data collection, which is another risk. Furthermore, we train the project team to minimize errors. We have detailed descriptions and instructions on how to administer and store the data. We try to set up the data collections in a way it is fun for children. The data collection will be embedded into a cover story of a treasure hunt to keep them motivated so that each child terminates the tasks.

As in every project, many other variables which are not measured likely interfere and impact the results. Concerning the data itself, another risk is to lose data when not merging the data frames correctly, specifically when applying the inner instead of outer merge strategy.

9 Preliminary Studies

Learning to write legibly and fluently (e.g., automated) is crucial for participation in education, society, and culture [20] and academic success [21]. It is estimated that depending on the age of the children, selection criteria, and measurement instruments used, between 10 and 30% of school-age children show handwriting difficulties [1]. Fine motor and visuomotor skills are central for legible and automated handwriting [22].

Previous research has shown that specific cognitive processes are central to learning handwriting [8], especially when a task is new [23, 24]. For example, novice writers need to remember letter shapes and plan a specific movement sequence, which needs to be continuously and flexibly adapted. As handwriting becomes more practiced and automated, a child can increasingly focus on the text content, grammar, text comprehension, and planning the next word or sentence [25]. Recent study findings have shown that school-aged children with widely automated handwriting write longer texts across the school career [26] and demonstrate better reading comprehension [27].

Numerous research findings support that the accuracy with which children copy geometric shapes is not only associated with handwriting legibility [28] but is also significant for a successful transition to school and academic success [29, for a review]. For example, results from a study by Dinehart and Manfra [30] show that fine motor and visuomotor skills at age 5 make a significant explanatory contribution to later academic achievement in reading and mathematics. This was found even when controlling for the impact of gender, family socioeconomic status, and preschool language and math skills. Further study findings show that fine and visuomotor difficulties can negatively impact motivation, self-efficacy, and academic learning [31, 32, 33]. This negative impact is even more problematic when the difficulties go unrecognized [34].

However, when having a first look at the data, there is only a weak relationship between handwriting and executive functions (measured in the same year), as figures 5 shows. Regarding the findings of other studies this seems rather surprising and suggests having a closer look at the data. However, the data still shows that in our sample, children with a higher degree of automation (=lower value of NIV) of handwriting tend to have slightly better fine motor skills.

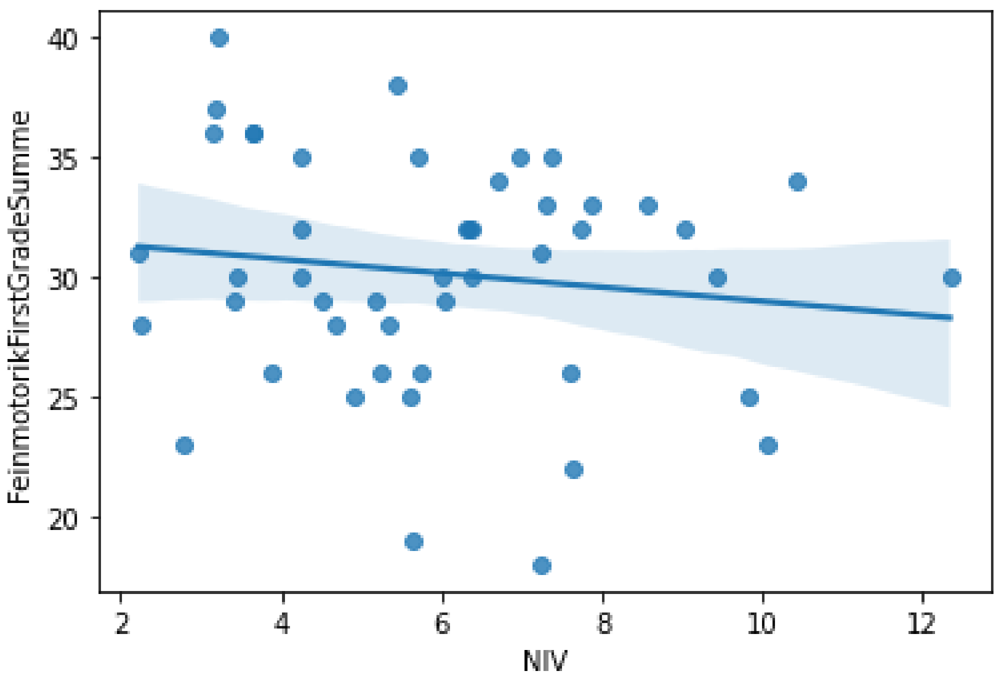

Figure 5: Relationship between degree of automation of handwriting (NIV) and fine motor skills at the same year

When analyzing the relationship between fine motor skills one year ago and automation of handwriting we find a very similar result to the above one. This suggest, that fine motor skills in kindergarten predict handwriting at least to some extent.

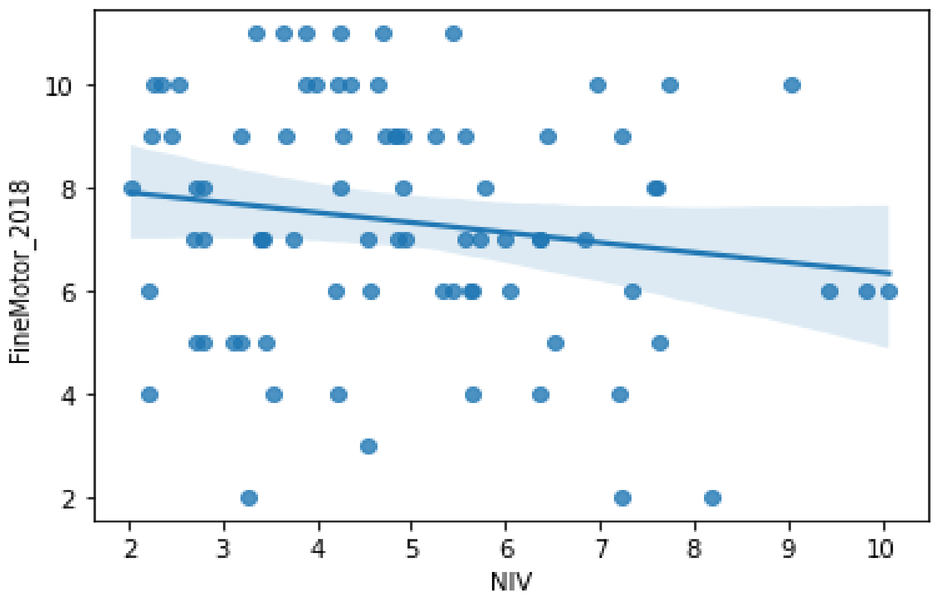

Figure 6: Relationship between degree of automation of handwriting (NIV) and fine motor skills of the year before

Further, the data doesn’t suggest that there is strong relationship between executive skills and fine motor skills (measured in the same year), as shown in figure 7 below.

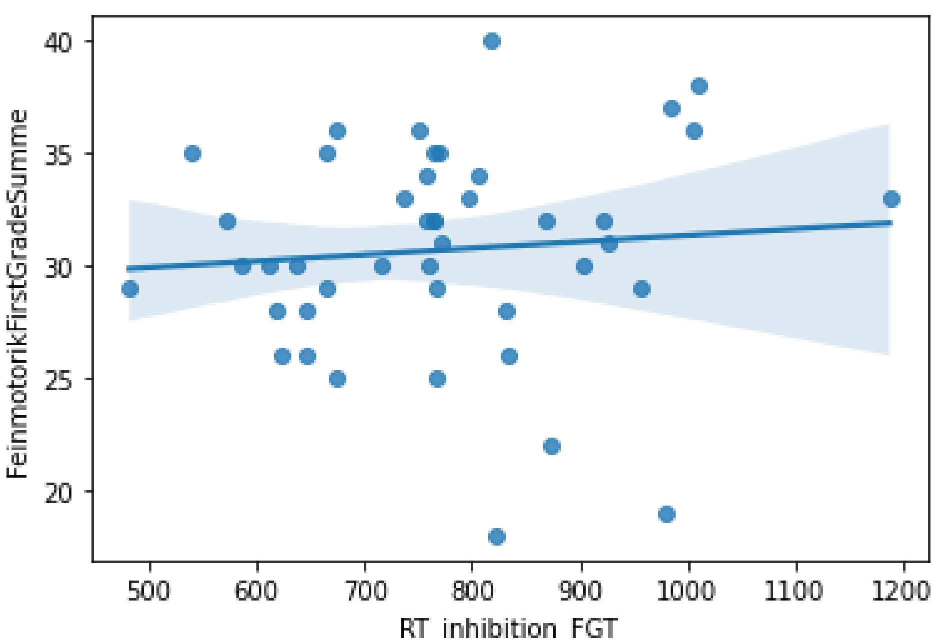

Figure 7: Relationship inhibition and fine motor skills

10 Conclusions

With the proposed project, we aim to identify early predictors of handwriting development. Further, we want to check if there are cross-sectional associations (bivariate correlations) among fine motor skills, visual-motor integration, and executive functions in first grade. To get to the bottom of those questions, we work with various python libraries and a dataset that contains results from various experiments conducted with children. These experiments provide information about handwriting skills, executive functions and fine motor skills of children.

In this report, we outlined our analysis and depicted the data flow and the data model of our analysis, identified various risks (e.g. issues with missing data and data quality) and did a few very simple preliminary studies. These preliminary results suggest that the relationships we are looking for indeed exist. However, they do not seem to be very strong. Further analysis will allow us to get further insights.

References and Bibliography

[1] Döhla, D., & Heim, S. (2016). Developmental Dyslexia and Dysgraphia : what can we learn from the one about the other ? *Frontiers in Human Neuroscience*, *6*(January), 1–12. https://doi.org/10.3389/fpsyg.2015.02045

[2] Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

[3] Fischbach, A., Schuchardt, K., Mähler, C., & Hasselhorn, M. (2010). Zeigen Kinder mit schulischen Minderleistungen sozio-emotionale Auffälligkeiten? *Zeitschrift Fur Entwicklungspsychologie Und Padagogische Psychologie*, *42*(4), 201–210. https://doi.org/10.1026/0049-8637/a000025

[4] Visser, L., Kalmar, J., Linkersdörfer, J., Görgen, R., Rothe, J., Hasselhorn, M., & Schulte-Körne, G. (2020). Comorbidities between specific learning disorders and psychopathology in elementary school children in Germany. *Frontiers in Psychiatry*, *11*, 1–12. https://doi.org/10.3389/fpsyt.2020.00292

[5] Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine and Child Neurology*, *49*(4), 312–317. https://doi.org/10.1111/j.1469-8749.2007.00312.x

[6] McCloskey, M., & Rapp, B. (2018). Developmental Dysgraphia: An Overview and Framework for Research. *Cognitive Neuropsychology*, *34*, 65–82. https://doi.org/10.1080/02643294.2017.1369016.Developmental

[7] Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*(1), 1–34. https://doi.org/10.1017/S0140525X01003910

[8] Valcan, D. S., Davis, H. L., Pino-Pasternak, D., & Malpique, A. A. (2020). Executive functioning as a predictor of children’s mathematics, reading and writing. *Journal of Applied Developmental Psychology*, *70*, 101196. https://doi.org/10.1016/j.appdev.2020.101196

[9] Wikipedia: https://en.wikipedia.org/wiki/Pandas_(software), Abruf am 16.09.21

[10] Medium: https://medium.com/fintechexplained/why-should-we-use-numpy-c14a4fb03ee9#:~:text=NumPy%20is%20an%20open%2Dsource,%2C%20statistical%2C%20and%20algebraic%20routines, Abruf am 15.09.21

[11] statistics — Mathematical statistics functions: https://docs.python.org/3/library/statistics.html, Abruf am 18.09.21

[12] matplotlib: https://matplotlib.org/, Abruf am 16.09.21
[13] Frick, A., & Möhring, W. (2016). A matter of balance: Motor control is related to children’s spatial and proportional reasoning skills. *Frontiers in Psychology*, *6*(JAN), 1–10. https://doi.org/10.3389/fpsyg.2015.02049

[14] Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. Science, 318(5855), 1387–1388.

[15] Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, *18*, 34–48. https://doi.org/10.1016/j.dcn.2015.11.005

[16] Bruininks, R. H., & Bruininks, B. D. (2005). Bruininks-Oseretsky test of motor proficiency (2nd ed.). Manual. Minneapolis: Pearson.

[17] Tröster, H., Flender, J., Reineke, D., & Wolf, S. M. (2016). *DESK 3-6 R: Dortmunder Entwicklungsscreening für den Kindergarten-Revision: Manual*. Hogrefe.

[18] Moll, K., & Landerl, K. (2014). *Lese-und Rechtschreibtest: SLRT-II: Weiterentwicklung des Salzburger Lese-und Rechtschreibtests (SLRT); Manual*. Hogrefe.

[19] Mai, N., & Marquardt, C. (2007). CSWin: Computergestützte Analyse der Bewegungsabläufe beim Schreiben. Med-Com.

[20] Vetter, M., Amft, S., Sammann, K., & Kranz, I. (2021). *G-FIPPS: Grafomotorische Förderung. Ein psychomotorisches Praxishandbuch* (third). Borgmann Media.

[21] Dinehart, L. H. (2015). Handwriting in early childhood education: Current research and future implications. *Journal of Early Childhood Literacy*, *15*(1), 97–118. https://doi.org/10.1177/1468798414522825

[22] Olinghouse, N. G., & Graham, S. (2009). The relationship between the discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology*, *101*(1), 37–50.

[23] Maurer, M. N., & Roebers, C. M. (2020). Is the fine motor-executive functions link stronger for new compared to repeated fine motor tasks? *PLoS ONE*, *15*(11), 1–9. https://doi.org/10.1371/journal.pone.0241308

[24] Maurer, M. N., & Roebers, C. M. (2021). New insights into visual-motor integration exploring process measures during copying shapes. *Psychology of Sport and Exercise*, *55*, 101954. https://doi.org/10.1016/j.psychsport.2021.101954

[25] Palmis, S., Danna, J., Velay, J. L., & Longcamp, M. (2017). Motor control of handwriting in the developing brain: A review. *Cognitive Neuropsychology*, *34*(3–4), 187–204.

[26] Salas, N., & Silvente, S. (2020). The role of executive functions and transcription skills in writing: a cross-sectional study across 7 years of schooling. *Reading and Writing*, *33*(4), 877–905.

[27] Malpique, A. A., Pino-Pasternak, D., & Roberto, M. S. (2020). Writing and reading performance in Year 1 Australian classrooms: associations with handwriting automaticity and writing instruction. *Reading and Writing*, *33*(3), 783–805.

[28] Daly, C. J., Kelley, G. T., & Krauss, A. (2003). Relationship between visual-motor integration and handwriting skills of children in kindergarten: A modified replication study. *American Journal of Occupational Therapy*, *57*(4), 459–462. https://doi.org/10.5014/ajot.57.4.459

[29] McClelland, M. M., & Cameron, C. E. (2019). Developing together: The role of executive function and motor skills in children’s early academic lives. *Early Childhood Research Quarterly*, *46*, 142–151. https://doi.org/10.1016/j.ecresq.2018.03.014

[30] Dinehart, L., & Manfra, L. (2013). Associations between low-income children’s fine motor skills in preschool and academic performance in second grade. *Early Education and Development*, *24*(2), 138–161. https://doi.org/10.1080/10409289.2011.636729

[31] Berninger, V. W., Mizokawa, D. T., & Bragg, R. (1991). Scientific practitioner: Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychologysychology*, *29*(1), 57–79.

[32] Christensen, C. A. (2005). The role of orthographic–motor integration in the production of creative and well-structured written text for students in secondary school. *Educational Psychology*, *25*(5), 441–453.

[33] Connelly, V., Dockrell, J. E., & Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educational Psychology*, *25*(1), 99–107.

[34] Sägesser, J., & Eckhart, M. (2016). *GRAFOS - Screening und Differentialdiagnostik der Grafomotorik im schulischen Kontext*. Hogrefe. https://www.hogrefe.ch