Mobile application for digital capacity assessment

Alberto Morelli[1], Alba Liso[2], Lorenzo Viviani[3], Elisa Straulino[4 5], Luisa Sartori [4 5], Laila Craighero[2], Leonardo Bocchi[1]
[1]Department of information engineering, University of Florence, Florence, Italy
[2]Department of neuroscience and rehabilitation, University of Ferrara, Ferrara, Italy
[3]Department of medical sciences, University of Ferrara, Ferrara, Italy
[4]Department of general psychology, University of Padova, Padova, Italy
[5]Padova Neuroscience Center, University of Padova, Padova, Italy

***Abstract*—Age is currently used as the sole or primary index of digital skills, categorizing smartphone users into two main groups: digital immigrants (DI), born before the 1980s, and digital natives (DN), born from the 1980s onwards. This distinction arises from the fact that DN, unlike DI, experienced the advent of smartphones during their youth, benefiting from the brain's plasticity in early developmental years, which makes acquiring new sensorimotor skills easier and more effective. However, while age plays a fundamental role in learning, it is reductive to consider it as the only determining factor. All generational cohorts who encountered smartphones learned to use them at different stages in their lives, with learning rates shaped by factors such as frequency of use, prior computer skills, and personal predispositions. To delve deeper into these dynamics, we developed an Android app and a keyboard, integrating them into an experimental pipeline consisting of four tasks (two writing tasks and two related to photo and gallery management) to collect data from smartphone users. Preliminary analyses focusing on task execution, conducted on data collected from 80 subjects, show that DN complete task with a higher average gesture speed than DI. However, some DI achieve performances comparable to those of DN, and vice versa. A regression analysis revealed a continuous relationship between age and digital skill, suggesting a more nuanced distribution of abilities across individuals. These findings indicate that further investigations, focused on gesture execution or the subject's history and lifestyle, could lead to the definition of a digital skill index with a more heterogeneous distribution among users compared to a simple binary classification based on age.**
***Keywords*—Sensorimotor skills, smartphone, machine learning, digital literacy**

I. INTRODUCTION

At the beginning of the last century, the main health concerns for the workforce were primarily related to physically demanding jobs, such as those in agriculture, mining, and the steel industry [1]. By the late 20th century and over the past two decades, thanks to the advent of information technologies and the pervasive automation of industry, a significant shift in physical capacity requirements in the workplace has emerged across much of the Western world. In countries with the highest penetration rates of computers, the Internet, cell phones and information technology in general, the skills needed for survival are increasingly shifting toward digital and communication skills, at the expense of physical skills. As a result, the physical workload in Western countries has overall become a less critical issue, while limitations in motor skills and dexterity have gained increasing importance [2]. Age plays a significant role in the development of these new skills (i.e., the age-related digital divide). The cultural changes associated with information technologies are so rapid that, at a certain point, aging becomes a form of alienation: the older people get, the more they find themselves living in a world different from the one they grew up in, requiring skills they had no opportunity to develop. In contrast, younger generations have acquired these skills during childhood and adolescence, periods when frequent use of such technologies fosters the development and refinement of specific abilities. This set of skills and attitudes, which facilitates interaction with digital technologies, is commonly referred to as digital literacy [3]. With the advent of smartphones, these generational differences have become even more evident, making it clearer whether the motor skills for using cell phones were developed at a young age or later in life.

However, all generational cohorts who experienced the advent of smartphones learned their use at a specific time in their lives, and their learning rate was affected differently by frequency of use, prior computer skills, and personal predispositions. It is therefore reasonable to expect that the range of abilities in smartphone use cannot be reduced to a simple age-based classification but rather exhibits a heterogeneous and stratified pattern, where digital skills are distributed along a continuum. The goal of this preliminary study was to develop a mobile application (SensoriMotor Digital Skill App, SMoDS-App) designed to collect mobile usage data according to a specific experimental pipeline, enabling the investigation of these concepts. The data was also used to conduct a preliminary analysis of the distribution of participants based on easily extractable features, such as the time each participant took to complete the experiment and the number of touchscreen interaction events recorded.

II. METHODS

A. Application and experimental pipeline design

The application was developed in Java using Android Studio [4], with the Google Pixel 7a as the target device and Andorid 13 as the target operating system. Alongside the application,

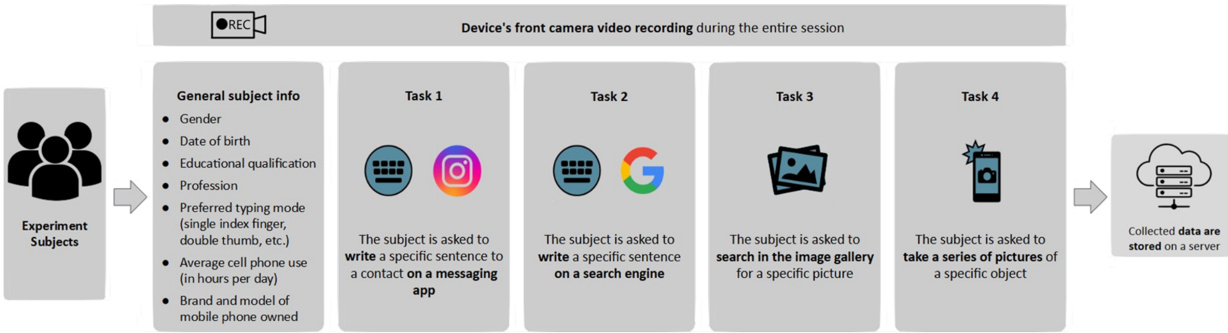


Fig. 1: Experimental pipeline scheme: for each subject, general information and a video recording were collected, along with touchscreen events recorded during a series of four specific tasks.

a custom keyboard (SMoDS-Keyboard) was developed. Since each participant used an instance of the new keyboard, this design eliminated any habitual bias that could arise if a standard keyboard had been used.

The structure of the application strictly follows a predefined experimental pipeline, outlined in Figure 1, which can be summarised in the following steps:

- Collection of personal data: gender, date of birth, educational level, profession, preferred typing method on a mobile device, average daily mobile usage, and the brand and model of the participant's phone.
- Typing task with SMoDS-Keyboard on a messaging app during which the subject is asked to type three specific sentences to a contact on a messaging app.
- Typing task with SMoDS-Keyboard on a search engine, during which the subject is asked to type the same sentences as in the previous task on a search engine
- Photo gallery search task during which the subject is asked to search for a specific image in the dedicated gallery on the application, select it and send it to a contact
- Picture-taking task during which the subject is asked to take three pictures of an object, review them via the dedicated gallery in the application and delete the two worst ones.

The tasks were designed with distinct objectives so that each interaction with the touchscreen was directed toward a specific goal. This approach aimed to encourage participants to perform gestures as naturally as possible by leveraging the well-known mechanism where the distal goal of an action influences its execution [5] [6]. In particular, the execution of gestures can be affected by whether the goal is communicative in nature [7] [8] [9]. The first and second tasks were designed to evaluate typing on the keyboard using the SMoDS-Keyboard in order to collect typing data. The phrases are always the same and are designed to be composed of letters as equally distributed between the left and right side of the keyboard as possible, while still making sense both as a possible message to a person and as a possible request to a search engine. The third and fourth tasks are designed to evaluate gestures such as swiping, double-tapping or zooming with two fingers and the ability to manage photos and the gallery. In each task, touch screen interaction events were collected individually and subdivided according to the type of event as: single tap, user interface (UI) interaction tap (such as buttons or the images thumbnail in the gallery), double tap, SMoDS-Keyboard typing, long pressure, one finger swipe and zoom with two fingers. For each of these events, both general data and event-specific data are collected. The general data include:

- **Screen coordinates:** The coordinates on the screen where the event occurred, with x and y values expressed in pixels.
- **Start/end time:** The timestamps indicating the start and end of the event, expressed in a date format (hour:minutes:seconds:milliseconds) relative to the day the experiment takes place.
- **Pressure:** The recorded pressure applied on the screen during the event, represented as a normalized value between 0 and 1, where 0 indicates no pressure and 1 indicates the maximum pressure detectable by the device.

The specific data collected based on the event type are:

- **UI interaction tap:** The name of the UI element involved in the interaction.
- **Double tap:** The time interval between the two consecutive taps.
- **Keyboard typing:** The type of character typed. In the case of special keys (e.g., Backspace or Spacebar), the corresponding key name is recorded.
- **Long pressure:** The time duration of the prolonged pressure.

Swipes, both one-finger and two-finger, have been handled differently. Specifically, the approach relied on Android's swipe management system, which considers the continuous movement of a finger on the screen as a single event but generates sub-events at regular intervals containing the current information about the finger's position and pressure. As a result, swipes are represented as groups of motion events, for which the coordinates, pressure, and timestamp of each time step are recorded, along with the total number of motion events comprising the swipe. For two-finger zoom gestures, the process simply concatenates two swipes (one for the first finger and one for the second finger). In addition, during the experiment, a video of the user's face is recorded using the front camera. The video allows for tracking the duration of the experiment, as the camera is automatically activated at the start (upon pressing the *Start Experiment* button) and deactivated at the end (upon pressing the *End Experiment and Send Data* button). Consequently, the duration of the video corresponds to the actual length of the experimental session. Furthermore, since each event on the touchscreen is timestamped, it is possible to correlate specific time instants with the corresponding moments in the video. This could enable cross-analysis of touchscreen interactions and eye movements.

The start and end of each task are recorded by pressing dedicated buttons (*start task*, *end task*), with their timestamps serving as reference points for calculating task duration. Upon completing the experiment, the video recording is saved in MP4 format on the smartphone, alongside the data collected for each task, which are stored in four separate JSON files. All files are then uploaded to a server for backup.

B. Participants

Eighty healthy participants were recruited from individuals of two distinct age groups, ensuring a clear separation with a threshold at 45 years old (corresponding to the year of birth 1980). This allowed us to adopt the classification proposed by Prensky [10], which distinguishes digital immigrants (DI), born before 1980, from digital natives (DN), born after 1980. Specifically the dataset consisted in 39 DN, 26 females, mean age = 22.2 ± 1.7 and 41 DI, 27 females, mean age = 56.8 ± 8.6. Each participant completed the experiment at their own pace under the supervision of an experimenter. All data were anonymized, and each participant provided written informed consent for data processing, ensuring full compliance with privacy regulations and the protection of individual confidentiality. The study was approved by the Ethics Committee of the University of Ferrara (Approval Code: 649/2023/Oss/UniFe).

C. Data analysis

The analysis conducted in this study is a preliminary investigation primarily focused on task execution speed. To provide a straightforward representation of performance, a Digital Skill Index (DSI) was introduced, based on the speed at which touch events occur and defined as:

$$\text{DSI} = k \times \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{d_i}{e_i}} \tag{1}$$

where:

- k is a correction factor chosen to scale the index to a meaningful range.
- n is the total number of tasks considered for the computation.
- d_i is the duration of task i (in milliseconds). This value is calculated as the difference between the timestamps of the *start task* and *end task* button presses recorded during the task execution.
- e_i is the number of events recorded during task i.
- $\frac{d_i}{e_i}$ represents the average time (in milliseconds) per event for task i.

Since the denominator represents the mean time per event across all tasks, a lower DSI indicates slower performance, characterized by a higher average time per event across the tasks. Conversely, a higher DSI reflects a faster execution, as it corresponds to a lower mean time per event. This being an initial study, it was decided to define a DSI that was as simple as possible with the intention of evaluating the analysis approach and possibly expanding its complexity in future studies.

Participants were divided into two groups (DN and DI). For each group, the mean and standard deviation of the DSI were calculated, and Welch's t-test was conducted to determine whether the differences between the groups were statistically significant. A linear regression model was also applied to explore the relationship between age and the DSI. This analysis yielded the regression slope, intercept, and the coefficient of determination (R^2), which indicates the proportion of variance in the DSI explained by age. Additionally, to evaluate the fit of the regression model, residuals (i.e., the differences between the observed and predicted DSI values) were calculated.

III. RESULTS

The analyses were conducted twice: once with the DSI calculated using only the first two tasks (typing tasks) and once using only the last two tasks (gallery management tasks). This decision was made to account for the different sensorimotor demands involved. While the typing task primarily requires keyboard-based input (mainly tapping), the gallery management task involves more gestures such as swiping and pinching for zooming, with significantly fewer taps on the screen.

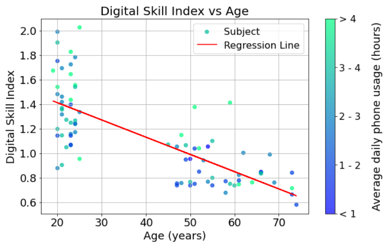
The results were visualized through several plots. A scatter plot of the DSI against age included the regression line and a gradient color scheme representing participants' average daily phone usage. Additionally, histograms were used to show the distribution of DSIs for the two age groups, highlighting overlaps and differences. Residuals were analyzed through scatter plots against both age and phone usage to identify any patterns not captured by the model. The results of these analyses are shown in Figure 2 and in Table I and II.

TABLE I: Summary of regression and statistical analysis for the typing tasks.

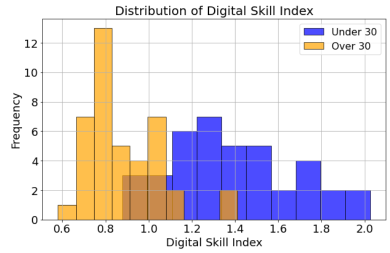
Analysis Type	DN	DI
Mean Digital Index (DSI)	1.394 ± 0.294	0.878 ± 0.180
Regression Slope	-0.014	
Intercept	1.693	
R-squared	0.538	
Welch's t-test	$t = 9.284$, $p = 0.000$	

TABLE II: Summary of regression and statistical analysis for the gallery tasks.

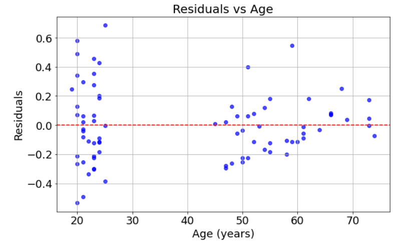
Analysis Type	DN	DI
Mean Digital Index (DSI)	0.371 ± 0.074	0.316 ± 0.058
Regression Slope	-0.001	
Intercept	0.398	
R-squared	0.125	
Welch's t-test	$t = 3.649$, $p = 0.000$	



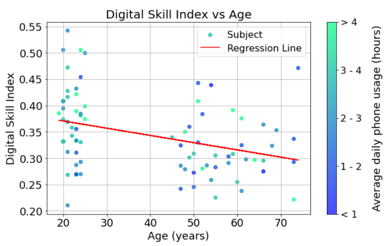
(a) Scatterplot (Typing Tasks)



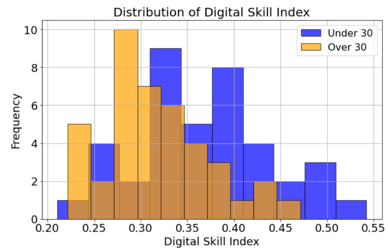
(b) Histogram (Typing Tasks)



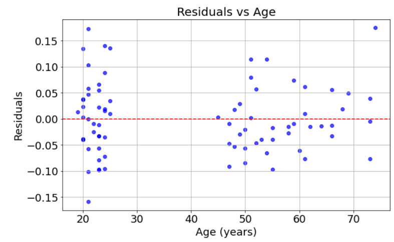
(c) Residuals (Typing Tasks)



(d) Scatterplot (Gallery Tasks)



(e) Histogram (Gallery Tasks)



(f) Residuals (Gallery Tasks)

Fig. 2: Scatterplots, histograms, and residuals for task performance analysis. The first row represents the typing tasks, while the second row corresponds to the gallery tasks.

IV. Discussion

The mean DSI values (I and II) reveal consistent differences between the two groups. In typing tasks, the DN group achieves a significantly higher mean DSI (1.394 ± 0.294) compared to the DI group (0.878 ± 0.180). This gap narrows in gallery-related tasks (0.371 ± 0.074 vs. 0.316 ± 0.058).

Welch's t-test confirms the statistical significance of these differences ($p < 0.05$), indicating a measurable performance distinction between DN and DI. However, regression analyses reveal varying levels of age influence. Typing tasks exhibit the highest R^2 value (0.538), suggesting that while age plays a stronger role in these tasks, it is insufficient to fully explain performance, highlighting the potential contribution of other factors. In gallery management tasks, the R^2 value drops to 0.125, indicating a minimal influence of age on performance.

The scatterplots show that younger participants generally achieve a higher DSI, but the overlap between the two age groups challenges a strict binary classification. This overlap is particularly evident in gallery tasks, where the skill gap is less pronounced. Histograms (Figures 2b and 2e) further illustrate this, with the DN group showing higher and broader DSI distributions in typing tasks, while gallery tasks display closer distributions between age groups.

Residual plots (Figures 2c and 2f) provide deeper insights into variability. In typing tasks, residuals are widely dispersed, especially among younger participants, suggesting individual differences significantly affect performance. Gallery tasks, however, show tighter residual clustering, particularly for the DN group, indicating more uniform performance influenced by experience rather than age. Notably, outliers in both age groups deviate from expected trends, emphasizing that performance cannot be solely predicted by age.

These findings underscore the complexity of digital skill assessment. The DSI overlap between age groups, particularly in gallery tasks, suggests that binary classifications based solely on age (e.g., DN vs. DI) oversimplify digital skill distribution. Instead, a more nuanced approach considering lifestyle, experience, and specific skill domains is needed.

Future research should refine the DSI by focusing on individual gesture execution rather than overall task performance, exploring non-age-related factors such as education, profession, and other factors that may shape specific smartphone usage patterns.

References

[1] H. Heuer, "Motor behavior and work: from physical load to motor skills," *Sport Kinetics*, vol. 97, pp. 243–250, 1999.

[2] ——, "Technologies shape sensorimotor skills and abilities," *Trends in neuroscience and education*, vol. 5, no. 3, pp. 121–129, 2016.

[3] M. Appel, "Are heavy users of computer games and social media more computer literate?" *Computers & Education*, vol. 59, no. 4, pp. 1339–1349, 2012.

[4] Google, "Android studio," https://developer.android.com/studio.

[5] R. G. Marteniuk, C. L. Mackenzie, M. Jeannerod, S. Athenes, and C. Dugas, "Constraints on human arm movement trajectories." *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 41, no. 3, p. 365, 1987.

[6] C. Ansuini, L. Giosa, L. Turella, G. Altoè, and U. Castiello, "An object for an action, the same object for other actions: effects on hand shaping," *Experimental Brain Research*, vol. 185, pp. 111–119, 2008.

[7] L. Sartori, C. Becchio, B. G. Bara, and U. Castiello, "Does the intention to communicate affect action kinematics?" *Consciousness and cognition*, vol. 18, no. 3, pp. 766–772, 2009.

[8] L. Sartori, C. Becchio, M. Bulgheroni, and U. Castiello, "Modulation of the action control system by social intention: unexpected social requests override preplanned action." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, no. 5, p. 1490, 2009.

[9] L. Craighero, U. Granziol, and L. Sartori, "Digital intentions in the fingers: I know what you are doing with your smartphone," *Brain Sciences*, vol. 13, no. 10, p. 1418, 2023.

[10] M. Prensky, "Digital natives, digital immigrants part 1," *On the horizon*, vol. 9, no. 5, pp. 1–6, 2001.