- ORIGINAL ARTICLE -

A Dataset and Post-Processing Method for Pointing Device Human-Machine Interface Evaluation

Un Conjunto de Datos y Método de Post-Procesamiento para Evaluación de Interfaz Humano-Máquina de Dispositivos de Tipo Apuntador

Rocío Madou[1], Federico N. Guerrero[1], and Enrique M. Spinelli[1]

[1]*Institute of Research in Electronics, Control and Signal Processing LEICI (UNLP-CONICET-CIC), La Plata, Argentina*
rocio.madou@ing.unlp.edu.ar

Abstract

The evaluation of human-machine interfaces (HMI) requires quantitative metrics to define the ability of a person to effectively achieve their goals using the HMI. In particular, for pointing-device type HMIs such as the computer mouse, an experiment quantifying movement by performing repetitive target selections allows defining a useful metric known as throughput (TP) using the Fitts' Law test. In this work, a dataset obtained from an automated protocol application is presented, which is made publicly available through an on-line platform. A post-processing method to obtain performance parameters from the dataset is also presented, and its output is used to validate the data against similar experiments in the literature.

Keywords: Dataset, Fitts' Law, Human-Machine Interface

Resumen

La evaluación de interfaces humano-máquina (HMI) requiere métricas cuantitativas para definir la capacidad de una persona para lograr eficazmente sus objetivos utilizando la HMI. En particular, para las HMIs del tipo dispotivo apuntador como el ratón de la computadora, un experimento que cuatifica el movimiento al realizar selecciones repetitivas de objetivos permite definir una métrica útil conocida como rendimiento (TP) utilizando la prueba de la Ley de Fitts. En este trabajo, se presenta un conjunto de datos obtenido a partir de la aplicación de un protocolo automatizado, el cual está disponible públicamente a través de una plataforma en línea. También se presenta un método de post-procesamiento para obtener parámetros de rendimiento a partir del conjunto de datos, y sus resultados se utilizan para validar los datos en comparación con experimentos similares en la literatura.

Palabras claves: Conjunto de datos, Interfaz Humano-Máquina, Ley de Fitts

1 Introduction

Human-Machine Interfaces (HMIs) provide a path of communication between a person and a device [1], a paradigmatic example being a computer and its input peripherals. In particular, there is a class of HMIs that seek alternative paths of communication aimed at people with motor disabilities or, in certain work or entertainment environments, achieve a different way of interacting with a desired device while still performing the most common function in computers such as pointer devices do [2, 3]. In order to evaluate the performance of HMIs that control a pointer on the screen a useful metric is Fitts' Law test [4, 5].

Fitts' Law test consists of a series of movement tasks where a person is asked to repeatedly reach and select a target using a specific tool, and it is mostly useful in the evaluation of continuous movement [5]. A version of this test has been selected in ISO standard 9241 for computer mouse evaluations [6]. The output of the test is an indicator of the average achieved information rate named Throughput (TP). The TP has been used as a metric related to the Information Transfer Rate (ITR) in brain-computer interfaces [7] and as a valuable tool to compare different pointing-device technologies and the abilities of different users when using these devices [8, 9, 10].

Based on its usefulness for HMI evaluation, an application that performs an automated Fitts' Law test was designed and tested by the authors in controlled experiments [11]. The data produced in these experimental sessions provides a complete account of the users' performance which contains further useful information compared with the usually reported statistical parameters. This data can be used to research the characteristics of the quantification algorithms and equations and as a benchmark in the design stages of similar experiments and general HMI evaluations. Studies of features of the test itself can be conducted when the full data is available. For example, processing strategies can be tested using weighted versions of the traditional algorithms [12], trade-offs that are not evident from the output parameters can be researched

[13], as well as alternative metrics [14]. In particular, the motivation for this work arose from the need to account for precise target hit rate which is important in effective PC control; this processing for example could not be performed over available sets of results. Therefore, in this work, first an open dataset called BMEP is presented; next, a method to analyze the dataset in post-processing is described so as to simplify its use; finally, the obtained samples are validated by comparing their characteristics with results form similar experiments in the literature.

2 Related work

In order to provide a useful dataset, well-determined and repeatable experimental conditions must be established. In several studies, ad hoc software applications were developed to perform Fitts' Law tests, either in standardized form or following variations for specific evaluations [15, 2, 16, 17]. These type of software implementations allow for a repeatable way of executing the test, more so when the experimental protocol is included as part of the preset execution.

Pointing device HMIs are aimed at controlling specific applications implemented to simplify the use of the HMI, and generic everyday applications where the HMI fully replaces a standard mouse. The Fitt's Law test has been used to test novel devices [18, 19, 20, 21] by employing the pointing device in standardized tests where the performance for generic applications can be derived, and also adapting non-standardized tasks [15] to better reflect the performance when using specific applications. Moreover, other types of experiments are designed when the data of interest differs form that produced by the test [22].

The availability of a full dataset of a test run can be processed so as to provide metrics for various of the aforementioned scenarios, accounting for specific characteristics such as considering the effect of accuracy by alternative algorithms, normalizing movement time against traveled distance, or as a benchmark otherwise unavailable, for example. Previous efforts to build a public dataset of the Fitts' Law test have been successfully undertaken on a large scale [23] however the results seem currently unavailable to the best of the authors' knowledge, thus, this work is a first step toward providing such presently nonexistent data, albeit initially at a smaller scale.

In experiments with fewer participants seeking to characterize different HMIs through Fitts' Law tests a number of participants between 12 and 19 have been recruited [24, 17, 16, 15], showing that relevant conclusions can be obtained with dataset sizes in this range.

3 Method

3.1 Fitts' Law test implementation

The Fitts' Law test is performed by placing two or more reference areas on the screen and asking a subject to repeatedly touch one or the other. The goal is to select the target "as quickly and accurately as possible" [16]. A complete session for one subject consists of a set of serial tapping tests where the subject taps back and forth between two rectangular targets as represented in Fig. 1. Each test of the session probes one difficulty level, repeating the task at that level several times to increase the statistical significance. The different tests try to cover a representative gamut of difficulty levels. Therefore, at the end of one session a subject will have completed N_t tasks (with different difficulty index) with N_r repetitions each.

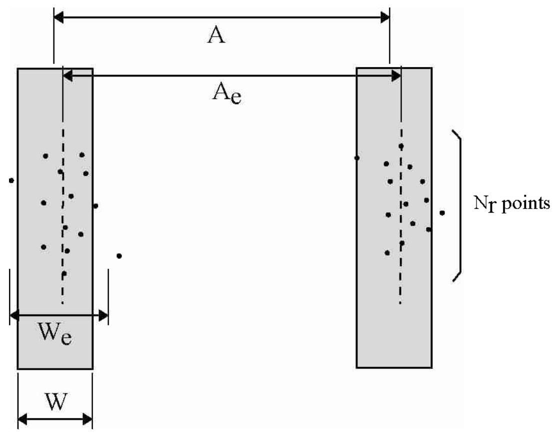

Figure 1: Schematic representation of a test for a fixed ID value.

When tests are performed on a PC, the targets are shown on the screen and, for the case of the 1D test, they consist in 2 rectangles of width W separated a distance A (see Fig. 1). Touching or tapping the targets is achieved by driving the pointing device to the target's delimited area and performing a "click", referring to the selection action that in the standard mouse peripheral consists in pressing the left button. Parameters W and A allow quantifying the difficulty level with a value defined as the index of difficulty (ID) that is calculated as

$$ID = \log_2(\frac{A}{W} + 1). \tag{1}$$

The width of the targets and distance between them are selected to cover a wide range of ID values, for example 1 to 8 according to Soukoreff & MacKenzie [4].

The average time that a subject takes in completing one repetition of a test is the movement time (MT). An averaged MT value is obtained, computed as

$$MT = \frac{1}{Nr} \sum_{k=1}^{k=Nr} MT_k, \tag{2}$$

where MT_k is the time between two taps of that test. From these metrics, a value is obtained defined as throughput (TP) that is calculated for each test (i.e., for each ID):

$$TP = \frac{ID}{MT}. \quad (3)$$

The throughput can be calculated for one test, for one subject, or a device. The result of a full session for each subject is the average TP across the Nt tests

$$TP = \frac{1}{Nt} \sum_{k=1}^{k=Nt} TP_k. \quad (4)$$

Further, the performance of a device is defined as the mean value of the subjects' performance indicators, obtaining an average TP for the device, and a coefficient of variation (CV) for this TP.

An alternative set of parameters use additional information regarding the actual position of the tap values instead of depending only on the movement time. After Nr repetitions of a test, a point cloud is generated (see Fig. 1). The position of the points along the horizontal axis x is then processed. The points are divided into groups belonging to each target. The mean x position of each group is obtained to define the two mean locations of each clicking area μ_{Xtl} and μ_{Xtr}. Once the mean locations are found, a single array that contains all the relative positions of the clicks with respect to the center of their respective targets is formed and the standard deviation (σ) is calculated.

Using μ_{Xtl}, μ_{Xtr}, the effective distance or amplitude (Ae) can be found, which represents the actual distance moved by the user:

$$Ae = |\mu_{Xtl} - \mu_{Xtr}|, \quad (5)$$

Next, the effective target width (W_e) can be obtained, which refers to the size of the target as perceived by the user. A method to obtain W_e is available [16] by which a normal distribution is assumed for the point positions, with a 96% hit rate and 4% of selections missing the target which corresponds to a width of

$$We = 4.133 \times \sigma. \quad (6)$$

As a result, a new effective index of difficulty (ID_e) is calculated using the effective parameters We and Ae according to

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \quad (7)$$

From this, an effective throughput (TP_e) is obtained.

$$TP_e = \frac{IDe}{MT}. \quad (8)$$

Finally, Fitts' Law proposes there is a linear relationship between MT and ID for each person and device given by

$$MT = b * ID + a, \quad (9)$$

where a and b can be found using a least-squares linear regression analysis using the ID and MT values across all tests of a subject. The r^2 value resulting from the regression analysis is also stored.

For tap values that have been effectuated very far from the target location, an "outlier" detection is put in place which discards distances higher than $5 \times W$ from the target's center. The MT to an outlier is thus not considered.

3.2 Dataset

Fifteen (15) people (12 males, 3 females) were recruited for the previously reported experimental tests [11]. Their ages ranged from 22 to 56 years, either with no visual impairments or using adequate vision correction. They were computer literate subjects and used their right hand as the preferred hand to complete the tests, although subjects 2 and 9 were left-handed. After post-processing, the file from one participant was left out from the datased since it contained unusable data due to a lack of attention to the consigned task [11]. All participants have signed informed consents to publish the experiment data and the records were anonymized.

The BMEP dataset has been organized into directories which hold the files produced during a complete experiment, in which several participants are asked to perform the same test session. The directories are sequentially numbered, with the addition of the range of dates when the experiment took place (with the format YYYYMMDD-YYYYMMDD). Each directory contains a markdown description file *readme.md* where the test is explained including a description of the experiment, the hardware, and the participant data. Each directory also contains the output data of the experiments which are a set of comma separated value (CSV) tables in one file for each subject. The presently reported results correspond to directory `Dataset001 [20230301-20230308]`.

The format of the file for each subject includes a header before the data. The header has 4 lines. It describes the name of the file (the timestamp of the moment the experiment is initiate), the configuration set in the test interface and the number of repetitions of each task Ir. The final line of the header describes the table indicating, in order, the information represented by each column, and the units, identified by square brackets ("[]"'). Below the header, the sampled data is listed. A typical file is schematized in Table 1, and the structure for the data is further detailed in Table 2.

3.3 Post-processing

Every result file corresponding to a complete session with one subject is treated individually by an Octave 8.2.0 program (compatible with MATLAB) to obtain the test output parameters for that subject. Within each file, the Nt tests must be processed and then the

Table 1: Description of files in the dataset.

Line	Description		Example file content
1	Header	Timestamp	20230302-142723
2		Software configuration	[1D - Original - ...] (Full HD (1080p)) †
3		Total repetitions	15
4		Column description	It,Ir,A(pixels),W(pixels),H(pixels),... ‡
5	Data	Test 1 Repetition 1	1,1,1024,32,810,811
6		Test 1 Repetition 2	1,2,1024,32,810,729
...		...	...
$Nt * Nr + 4$		Test Nt Repetition Nr	16,15,128,16,810,645

† Further description in the text (sec. 3.1).
‡ Detailed in Table 2.

Table 2: Description of data columns.

Acronym	Description	Unit
It	Number of task	[1 to Nt]
Ir	Number of repetition	[1 to Nr]
A	Movement amplitude	pixels
W	Target width	pixels
H	Target height	pixels
MT	Movement time	milliseconds
X	X position of the click	pixels
Y	Y position of the click	pixels
Xt	X position of the center of the target	pixels
Yt	Y position of the center of the target	pixels
HIT	Target hit achieved	[1 or 0]

average subject's metrics are calculated. Finally, the results across all subjects are processed to obtain the device's performance parameters.

The program that processes each file uses an array of data structures shown schematically in Fig. 2 to capture the information of each test, organize it, and calculate the outputs. The algorithm used to process each file is presented in Algorithm 1.

As detailed in Fig. 2, there are two structures, `pos` and `post`, that have been defined to store the information about the pointing device and target positions. In `post`, the targets are differentiated and the clicks are segregated according to the target they belong to.

After data from the *Nt* tests of a file has been processed, *a* and *b* coefficients defining the linear relationship between ID and MT are found, together with the r^2 coefficient. Next, once the values are calculated and stored, rows of the table are arranged in ascendant order by A, W and H for visualization purposes. The results of TP, MT and error rate of each stage of the general test are plotted, and a fourth figure with the plot of the correlation between MT and ID values, with their linear regression, is also produced.

Algorithm 1 Pseudocode of the processing algorithm for each file. The data and results are stored in an array of structures as seen in Fig. 2 named `test_data`

```
Import file header and data to environment
Nr ← Read from header
Nt ← data size / Nr
for t from 1 to Nt do
    test_data[t] ← test information: It, A, W, H
    test_data[t] ← ID calculated from A, W
    for r from 1 to Nr do
        if outlayer then
            continue
        end if
        test_data[t].MTk[r] ← MTk
        test_data[t].pos[r] ← target and mouse positions: X, Y, Xt, Yt
        test_data[t].post ← target and mouse position processing
        test_data[t] ← error count update
    end for
    test_data[t].MT ← average of MTi values
    test_data[t].TP ← TP calculated from ID, MT
    test_data[t] ← calculate effective values - Ae, We, He, IDe, TPe
end for
calculate linear regression coefficients a, b, r^2
```

4 Results and discussion

A dataset of 14 instances of Fitts' Law test was obtained and made publicly available through a repository in the platform Github [25]. The repository is divided into directories identified by a numeric code and the range of dates in which the samples were taken. The directories contain text files with data in CSV format and a document written markdown language that describes the tests that produced the data in the directory, the configuration and conditions in which the experiment was performed, and the necessary information about the participants.

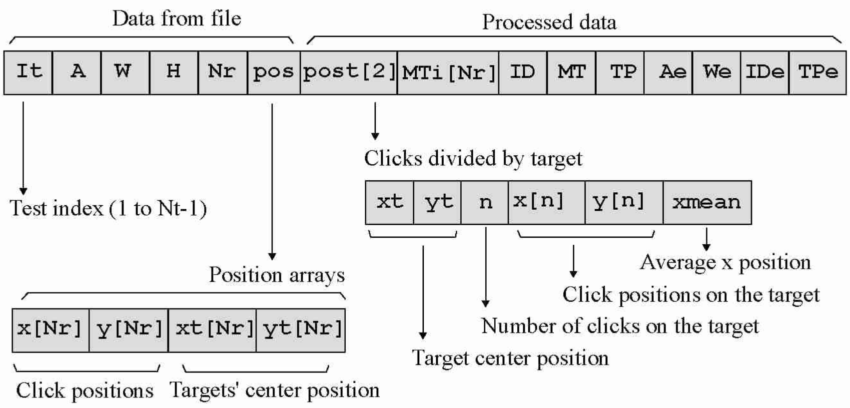

Figure 2: Data structure used to organize information for each test of a file.

The experimental data from Dataset001 [20230301-20230308] was obtained as described in sec. 3.2 and processed with the method presented in sec. 3.3. Useful data was extracted from 14 of 17 files, since in 3 files a different experimental protocol was followed [11], which was detected by examining the file header. This excluded protocol is a variation of Fitts' test introduced in 1964 [26], where the order of selection of the target was determined by a stimulus light. These files were added to the dataset due to their potential relevance in future work, however, they do not constitute part of the results presented in this study.

In order to validate the dataset, the post-processing results are presented in detail. The performance parameters for each participant are shown in Table 3. Accounting for inter-subject variability, the results are consistent. Some participants placed more relevance in accuracy than rapidness, and so they achieve higher MTs and higher hit rate, e.g. the 3 most accurate ($99.5 \pm 0.2\%$) have an average MT of 813 ± 82ms while the 3 least accurate ($91.2 \pm 1.5\%$) achieve a MT of 683 ± 61ms.

A sample plot from one participant is shown in Fig. 3. The error rate was very low, signaling that a relatively higher MT was needed to achieve the goals, producing a lower TP. However, when the effective values are considered, the improved accuracy is accounted for, and the TP_e reflects this achieving higher values as expected. The saw-tooth pattern of the MT vs A,W curve is also characteristic of the test [17], and an examination of the curves for all participants show a general conformance with this pattern.

An additional metric involving the normalized MT with respect to Ae is presented in Fig. 4. It can be seen that the amount of time per traveled pixel that each repetition takes decreases when the distance between targets increases. This is consistent with users slowing the speed of movement near the target area to achieve a higher degree of accuracy and increasing the speed over long stretches where accuracy is less important. This behavior is shared by all subjects as can be seen in Fig. 5.

Finally, the performance indicators calculated for

Table 3: Summary of individual participant results.

ID	Hits [%]	MT [ms]	TP [b/s]	b	a	(r^2)	CV	TP_e [b/s]	b_e	a_e	r_e^2	CV_e
1	99.2	897.9	4.32	335.5	140.6	0.83	0.24	5.11	390.4	106.8	0.75	0.31
2	93.3	676.4	5.70	245.1	107.8	0.93	0.21	5.82	242.5	106.8	0.89	0.23
3	98.3	688.2	5.61	204.6	120.9	0.93	0.17	6.85	154.4	112.4	0.86	0.13
4	92.9	613.7	6.31	229.3	96.1	0.89	0.22	7.15	230.3	85.2	0.84	0.21
5	99.6	807.8	4.78	242.5	141.3	0.95	0.18	6.09	74.5	148.5	0.92	0.11
6	95.0	818.2	4.72	298.9	129.9	0.91	0.22	4.93	292.5	127.5	0.83	0.23
7	90.0	682.6	5.63	305.7	94.2	0.92	0.23	5.81	281.7	98.7	0.83	0.22
8	99.6	734.7	5.28	278.8	114.0	0.86	0.23	6.29	242.2	105.1	0.76	0.22
9	95.4	756.2	5.12	294.5	115.4	0.85	0.23	6.00	313.6	95.7	0.70	0.23
10	90.8	755.3	5.09	352.4	100.7	0.92	0.25	5.49	320.3	102.4	0.83	0.20
11	98.8	682.1	5.66	220.6	115.4	0.93	0.20	6.38	238.0	98.7	0.89	0.23
12	97.5	882.2	4.38	290.2	148.0	0.90	0.21	5.03	316.3	124.0	0.82	0.21
13	95.4	674.4	5.78	176.5	124.5	0.87	0.19	6.87	108.5	122.3	0.83	0.17
14	95.4	717.1	5.41	240.2	119.2	0.84	0.22	6.05	179.4	122.9	0.75	0.17

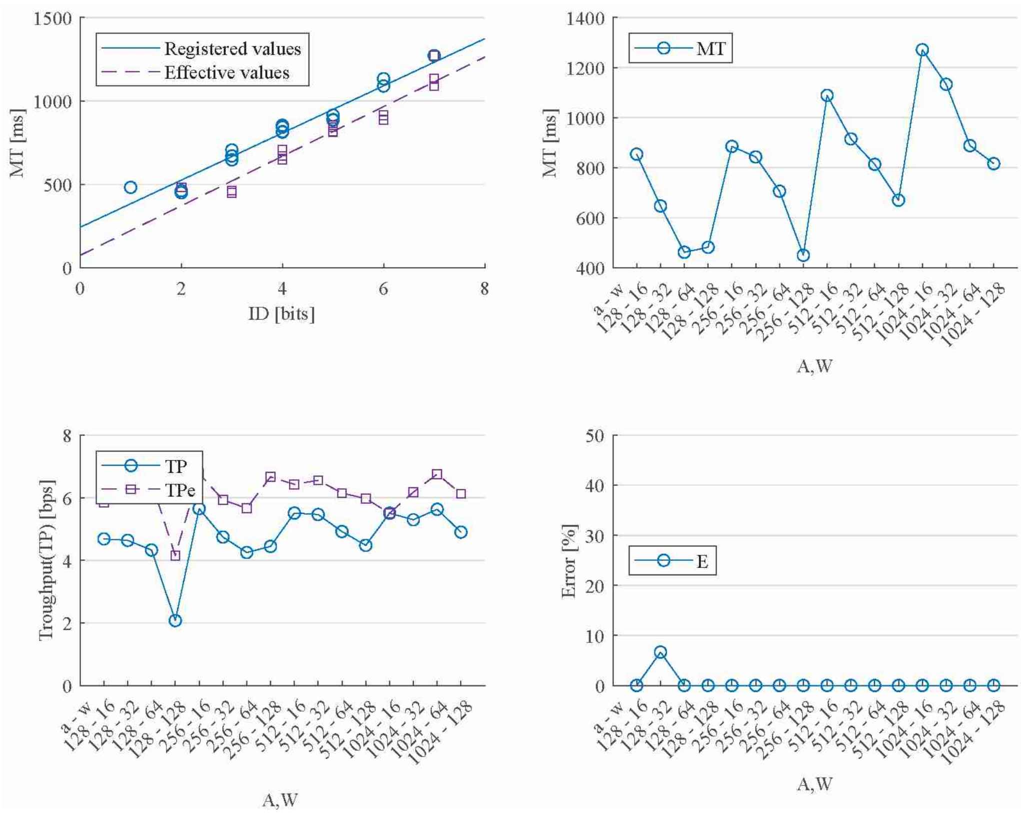

Figure 3: Example performance plot from file 20230303-155824 in the dataset (index It=8 from Table 3).

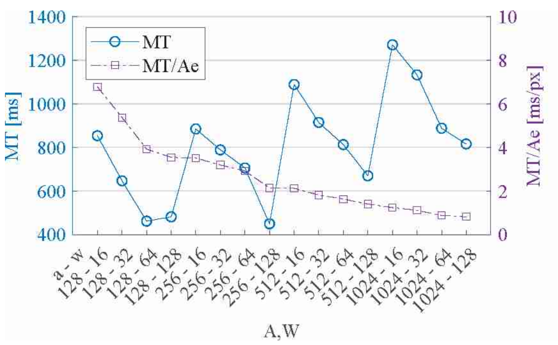

Figure 4: Comparison of curves of MT and normalized MT with respect to Ae plot from file 20230303-155824 in the dataset (index It=8 from Table 3.)

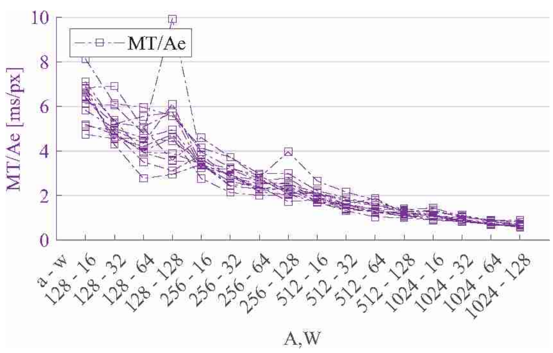

Figure 5: Normalized MT curves of each subject.

the computer mouse used in the experiments are presented in Table 4. The values are within the range of results found in the literature. A review by Soukoreff & MacKenzie [4] presented a TP range of 3.7 to 4.9 bps using older mechanical mice. A 2D comparative performance test by Sambrooks, Lawrence & Wilkinson [27, Fig. 2] resulted in a higher average TP of $7.5 \pm 1.4\%$ for 15 participants. A 2009 1D test by Sasangohar, MacKenzie & Scott [17] showed higher accuracy and higher MT (97.9 % and 965 ms) thus the TP is lower (3.83 bps vs 5.27 bps), however the effective values are not available and they could show a better approximation for the previously discussed trade-off between accuracy, MT and TP. Nevertheless, this trade-off has also been demonstrated through two different designs of Fitts 2D evaluation [10, 28] which resulted in the effective indicators detailed in Table 4.

Table 4 illustrates that the hit rate is comparable to those reported in other studies. While in line with what was discussed about the throughput, the MT was 1.74% lower than Sanchez's results and 4.2% lower than Nappenfeld's findings, which impacted consistently in the regression parameters. Sanchez et al con-

Table 4: Performance indicators for the pointing device.

Averaged effective performance indicator	This work	Sanchez et. al, 2021	Nappenfeld et. al, 2018
Hits	95.8±3.2 %	97.5±4.1 %	100 %
Movement time (MT)	741.9±83.4 ms	973.9±165.5 ms	810 ms
Throughput (TP)	5.98±0.69 bps	3.09±0.48 bps	2.16 bps
Regression line slope (b)	241.7 ms/bit	226.4 ms/bit	150 ms/bit
Regression line intercept (a)	111.2 ms	364.4 ms	550 ms
Coefficient of correlation (r^2)	0.8233	0.9926	0.23
Coefficient of variation of TP (CV)	0.2046	-	0.27

ducted the tests with a population aged between 6 and 8 that could have influenced the difference in MT. Furthermore, the coefficient of correlation is lower than the value reported by Sanchez et. al, likely product of the higher number of A&W combinations in this work. In contrast, r^2 was higher than Nappenfeld et. al where there was more dispersion in the results.

The processing method presents a potential weakness since there are clicks that result in error because they are not made within the target area, and are neither considered outlier values because the distance to the center is less than 3×W. This occurs mainly when the target is wide for low ID values, and any point in the screen is valid. In these cases the measured MT does not reflect the real time that would take a user to select the desired element. In the presented dataset, however, there are no instances of this issue.

5 Conclusions

A dataset containing results of a Fitts' Law test for pointing device Human-Machine Interfaces was presented. The information was collected with a previously reported experimental protocol [11] and has been made publicly accessible through the BMEP public repository on Github [25]. This dataset has potential to serve as a valuable resource for future investigations, enabling the incorporation of data from diverse experiments. Additionally, a post-processing method was introduced to obtain useful metrics per Fitts' Law from the dataset. The results from this processing algorithm were presented as a validation of the dataset, demonstrating a favorable agreement with similar studies documented the literature. Future work will include the expansion of the dataset by increasing number of participants involved in the same experiment, along with additional experiments using different classes of HMIs and tests configurations.

Competing interests

The authors have declared that no competing interests exist.

Funding

This work was supported by Research Project PIP-0323 from the National Scientific and Technical Research Council (CONICET), Argentina; and Research Projects PPID/I0016 from La Plata National University (UNLP), Argentina.

Authors' contribution

R.M. wrote the program, conducted the experiments, contributed to the analysis of the results, and wrote the manuscript; F.N.G. contributed to the analysis of the results and the writing of the manuscript; E.M.S. conceived the idea, supervised the work, and revised the manuscript. All authors read and approved the final manuscript.

References

[1] H. R. Hartson and D. Hix, "Human-computer interface development: concepts and systems for its management," *ACM Computing Surveys (CSUR)*, vol. 21, no. 1, pp. 5–92, 1989.

[2] A. J. Molina-Cantero, J. A. Castro-García, F. Gómez-Bravo, R. López-Ahumada, R. Jiménez-Naharro, and S. Berrazueta-Alvarado, "Controlling a mouse pointer with a single-channel eeg sensor," *Sensors*, vol. 21, no. 16, p. 5481, 2021.

[3] P. R. Thomas, "Performance, characteristics, and error rates of cursor control devices for aircraft cockpit interaction," *International Journal of Human-Computer Studies*, vol. 109, pp. 41–53, 2018.

[4] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci," *International journal of human-computer studies*, vol. 61, no. 6, pp. 751–789, 2004.

[5] D. E. Thompson, L. R. Quitadamo, L. Mainardi, S. Gao, P.-J. Kindermans, J. D. Simeral, R. Fazel-Rezai, M. Matteucci, T. H. Falk, L. Bianchi, *et al.*, "Performance measurement for brain–computer or brain–machine interfaces: a tutorial," *Journal of neural engineering*, vol. 11, no. 3, p. 035001, 2014.

[6] A. Ham, J. Lim, and S. Kim, "Do we need a faster mouse? empirical evaluation of asynchronicity-induced jitter," in *The 34th Annual ACM Symposium on User Interface Software and Technology*, pp. 743–753, 2021.

[7] X. Hu, A. Song, Z. Wei, and H. Zeng, "Stereopilot: A wearable target location system for blind and visually impaired using spatial audio rendering," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1621–1630, 2022.

[8] R. B. Widodo, R. M. Quita, R. Setiawan, and C. Wada, "A study of hand-movement gestures to substitute for mouse-cursor placement using an inertial sensor," *Journal of Sensors and Sensor Systems*, vol. 8, no. 1, pp. 95–104, 2019.

[9] K. S. Jones, T. J. McIntyre, and D. J. Harris, "Leap motion- and mouse-based target selection: Productivity, perceived comfort and fatigue, user preference, and perceived usability," *International Journal of Human–Computer Interaction*, vol. 36, no. 7, pp. 621–630, 2020.

[10] C. Sanchez, V. Costa, R. Garcia-Carmona, E. Urendes, J. Tejedor, and R. Raya, "Evaluation of child–computer interaction using fitts' law: A comparison between a standard computer mouse and a head mouse," *Sensors*, vol. 21, no. 11, p. 3826, 2021.

[11] R. Madou, F. N. Guerrero, and E. M. Spinelli, "Protocolo experimental para análisis de usabilidad de periféricos de tipo apuntador," in *VII Jornadas de Investigación, Transferencia y Extensión de la Facultad de Ingeniería (La Plata, 2023)*, pp. 170–178, Facultad de Ingeniería UNLP, 2023.

[12] E. Al-Imam and E. Lank, "Biasing response in fitts' law tasks," in *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, pp. 460–465, 2006.

[13] Y. Guiard, H. B. Olafsdottir, and S. T. Perrault, "Fitt's law as an explicit time/error trade-off," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1619–1628, 2011.

[14] J. Gori, O. Rioul, Y. Guiard, and M. Beaudouin-Lafon, "One fitts' law, two metrics," in *Human-Computer Interaction–INTERACT 2017: 16th IFIP TC 13 International Conference, Mumbai, India, September 25–29, 2017, Proceedings, Part III 16*, pp. 525–533, Springer, 2017.

[15] M. F. Roig-Maimó, I. S. MacKenzie, C. Manresa-Yee, and J. Varona, "Evaluating fitts' law performance with a non-iso task," in *Proceedings of the XVIII International Conference on Human Computer Interaction*, pp. 1–8, 2017.

[16] I. S. MacKenzie, "Fitts' law," *The wiley handbook of human computer interaction*, vol. 1, pp. 347–370, 2018.

[17] F. Sasangohar, I. S. MacKenzie, and S. D. Scott, "Evaluation of mouse and touch input for a tabletop display using fitts' reciprocal tapping task," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, pp. 839–843, SAGE Publications Sage CA: Los Angeles, CA, 2009.

[18] X. Guo, "A Fitts' law evaluation and comparison for human and manipulator on touch task," *Cognitive Computation and Systems*, vol. 4, pp. 265–272, Sept. 2022.

[19] D. J. L. L. Pinheiro, J. Faber, S. Micera, and S. Shokur, "Human-machine interface for two-dimensional steering control with the auricular muscles," *Frontiers in Neurorobotics*, vol. 17, p. 1154427, June 2023.

[20] L. Pandey and A. S. Arif, "Design and Evaluation of a Silent Speech-Based Selection Method for Eye-Gaze Pointing," *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, pp. 328–353, Nov. 2022.

[21] Yubin Liu, C. B. Sivaparthipan, and A. Shankar, "Human–computer interaction based visual feedback system for augmentative and alternative communication," *International Journal of Speech Technology*, vol. 25, pp. 305–314, June 2022.

[22] M. Zapata, K. Valencia-Aragón, and C. Ramos-Galarza, "Experimental Evaluation of EMKEY: An Assistive Technology for People with Upper Limb Disabilities," *Sensors*, vol. 23, p. 4049, Apr. 2023.

[23] K. Goldberg, S. Faridani, and R. Alterovitz, "Two large open-access datasets for fitts' law of human motion and a succinct derivation of the square-root variant," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 1, pp. 62–73, 2014.

[24] I. S. MacKenzie and W. Buxton, "Extending fitts' law to two-dimensional tasks," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 219–226, 1992.

[25] GIBIC-LEICI, "Github public repository of bmep datasets." `https://github.com/gibic-leici/bmep-datasets`. (Accessed on 04/05/2023).

[26] P. M. Fitts and J. R. Peterson, "Information capacity of discrete motor responses.," *Journal of experimental psychology*, vol. 67, no. 2, p. 103, 1964.

[27] L. Sambrooks and B. Wilkinson, "Comparison of gestural, touch, and mouse interaction with fitts' law," in *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, pp. 119–122, 2013.

[28] N. Nappenfeld and G.-J. Giefing, "Applying fitts' law to a brain-computer interface controlling a 2d pointing device," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 90–95, IEEE, 2018.

Citation: R. Madou, F. N. Guerrero and E. M. Spinelli. *A dataset and post-processing method for pointing device human-machine interface evaluation.* Journal of Computer Science & Technology, vol. 23, no. 2, pp. 127–134, 2023.
DOI: 10.24215/16666038.23.e11
Received: April 14, 2023 **Accepted:** August 15, 2023.