EMPHASIZING HUMANS IN INDUSTRY 5.0: A CROSS-AGE ANALYSIS OF BEHAVIORAL ENTROPY AND COGNITIVE WORKLOAD IN VR-BASED TELEROBOTICS

Federica Nenna[1], Davide Zanardi[2] and Luciano Gamberini[1]
[1]*Department of General Psychology, University of Padova, Padova, Italy*
[2]*Department of Developmental Psychology and Socialization, University of Padova, Padova, Italy*

ABSTRACT

Human motion patterns have been shown to reflect cognitive processes in psychology and applied Human-Computer Interaction (HCI) research. Specifically, behavioral entropy of users' movement trajectories has been proposed to consistently reflect workload and fatigue in various scenarios, including Virtual Reality (VR). In VR, this measure is typically calculated based on the movements of the VR controllers, known as Entropy of Controller Movements (ECM). While having great potential as a highly sensitive and unobtrusive measure of human workload, the effectiveness of ECM in more applied and realistic instances of VR-based applications (e.g., teleoperation platforms) has not been proven yet. Furthermore, the existing literature is based on young experimental samples, living open the question of whether age-related changes in motor performance might prevent using ECM as a measure of workload. We here covered these aspects by assessing relations between workload and ECM in 15 young and 15 senior participants who physically maneuvered an industrial robot in VR. Specifically, they were asked to guide the robot through a pick-and-place task by using their own physical movements in VR. Our findings evidence some criticalities of ECM as a measure of workload in our VR-based industrial contexts, opening new questions on its applicability and effectiveness.

KEYWORDS

Behavioral Entropy, Virtual Reality, Workload, Industrial Robotics, Entropy of Controller Movements

1. INTRODUCTION

1.1 Virtual Reality in Manufacturing and Industrial Robotics

Industry 5.0 is enabling greater use of robotic and autonomous systems in manufacturing while bringing peculiar attention to humans. Always more considerations are done regarding human workload in the workplace (Coronado et al., 2022; Panagou et al, 2023) and a great deal of effort is spent on better designing human-machine interactions, to lower users' fatigue and provide workers with the best robotic support.

Lately, physical industrial settings are being transferred into digital platforms allowing users to operate or inspect robots and machines also remotely. Such platforms, also known as digital twins (Van der Valk et al., 2020), can exploit the most various interfaces, including Virtual Reality (VR) (Havard et al., 2019). VR is particularly attractive in this sector for the following reasons. First, VR can faithfully reproduce real scenarios, even the most complex ones, with high levels of immersion. This feature is significantly useful in the robotics and manufacturing industry, where it is crucial for usability test settings to be as close as possible to the real ones. Second, implementing novel features in physical robots for usability tests is time-consuming and requires a great deal of work (Dautenhahn, 2018). Also, interacting with autonomous machines can be particularly dangerous in some situations, or even impossible in others (e.g., Guo et al., 2021). By using simulation software in VR, it is possible to overcome issues related to the hardware, generate usability test settings, or even get immersed in physically unreachable environments in a much faster, more efficient, safer, and cheaper way

(Dianatfar et al., 2021; Duguleana et al., 2011). Third, VR allows interactions with any virtual object exploiting human physical and embodied mechanisms. This calls to action human spatiality, which is the innate ability to act in physical space and thus facilitate any interaction with the virtual replica of physical objects (Villani et al., 2018).

1.2 VR-Based User Performance and Workload Metrics

Besides the mere advantage of enabling physical interactions, VR interactivity allows for collecting vast amounts of data related to user behavior and performance. This data can be used to improve the user experience, as well as inform research in fields such as cognitive science, psychology, ergonomics and Human-Computer Interaction (HCI). A broad body of literature in the work sector has demonstrated how workload and performance are strictly related, and how high workload levels are associated with mental fatigue, frustration, errors, and distraction (Galy et al., 2012; Young et al., 2015). Such factors can have important consequences in industrial robotics, and therefore, are particularly worth measuring, also in VR.

In this respect, all VR hardware (i.e., headset, controllers, body trackers) continuously generates time series data about their position and rotation. Such data can be leveraged to compute, for example, the start time and duration of any interaction with a virtual object, the users' movement velocity within a work setting, their position with respect to the digital replica of a robot or a machine, their task efficiency, or their work peace (Nenna et al., 2022; Nenna et al., 2023). All these metrics can report on the users' behaviors within a VR-based industrial or work context, enhancing our understanding of how people interact with virtual environments.

Remarkably, time series data on the VR controller position were further leveraged to gain insights into users' workload (Reinhardt et al., 2019; 2020). More precisely, the authors computed the Entropy of Controller Movement (ECM), which was demonstrated to be significantly modulated by the users' mental workload. While entropy measures have most likely been computed on gaze data and in desktop and mouse- or joystick-based settings within the human research areas (Chatzithanos et al., 2021; Diaz-Piedra et al., 2019; Goodrich et al., 2004; Reinhardt and Hurtienne, 2018; Stillman et al., 2018; Wu et al., 2020), the evidence of the effectiveness of such a measure within virtual environments opens new possibilities for continuous and indirect workload monitoring in VR-based industrial robotics.

1.3 Behavioral Entropy as a Measure of Workload

The concept of entropy refers to the degree of irregularity, randomness, and disorder in a system. It is typically used to quantify the complexity of different structures or processes or to learn about the randomicity of data or component variations (Wehrl, 1978). While it was initially developed to describe physical phenomena, the concept of entropy can actually extend to different kinds of data, including time series data, thus applying to various phenomena and application areas.

In the HCI and ergonomics fields, entropy has been employed to infer human workload, fatigue, or decisional processes by analyzing the unpredictability of specific movement trajectories. Nakayama, Boer, and colleagues (1999; 2000) first employed measures of entropy in the steering wheel of a vehicle to estimate drivers' workload (i.e., steering entropy). Subsequently, this concept was generalized to human activity at large, which is currently known as *behavioral entropy*. Some studies computed behavioral entropy on users' movements via mouse (e.g., McKinstry et al., 2008; Reinhardt and Hurtienne, 2018; Stillman et al., 2018), or VR controllers (Reinhardt et al., 2019; 2020); some others also computed the entropy of eye movements as a measure of workload (e.g., Wu et al., 2020).

In all such cases, when individuals interact with technology, they perform task-related movements whose complexity and accuracy can be influenced by the task demand (e.g., complexity, precision). As the task complexity or precision increases, users may experience greater cognitive and physical demands, resulting in more random and unpredictable movement trajectories. More specifically, Goodrich et al. (2004) proposed that, when operators face high workload or other factors causing degraded performance, they might select less efficient behaviors, anticipate less and react more, therefore resulting in more fragmented or exaggerated actions (i.e., reactive behaviors). Under lower workloads, instead, operators are likely to perform anticipatory behaviors, which are smoother with lower magnitudes and less frequent changes. As a result, examining the entropy of human movements can provide insights into the levels of users' workloads.

Goodrich et al. (2004) tested these hypotheses in seven users teleoperating a robot through direct or shared control, while additionally performing an arithmetic task. They concluded that behavioral entropy allowed to identify the most complex conditions of the teleoperation task. More recently, Chatzithanos et al. (2021) used behavioral entropy in a remote inspection scenario, whereby a sample of three operators teleoperated a robot to navigate an arena through a joypad. Different workload levels were created via dual-tasking and were directly related to the entropy values.

On the measurement of behavioral entropy within immersive virtual environments, Reinhardt et al. (2019) measured the Entropy of Controller Movements (ECM) in twenty-two participants executing a simple rhythm game in VR. They found a clear relation with their workload, indicating ECM as a promising mental workload measurement even in VR. In a subsequent experiment, the same authors tested twenty students performing the e-crossing task in VR, and demonstrated positive relations between the task difficulty, mental workload reported at the NASA-TLX and ECM (Reinhardt et al., 2020).

While the literature on behavioral entropy in immersive VR is not that extensive, ECM seems to have great potential as a highly sensitive and unobtrusive measure of workload in various VR scenarios. Nonetheless, to the best of our knowledge, literature assessing the effectiveness of ECM in more applied and realistic instances of VR-based environments (e.g., teleoperation platforms) is missing. Furthermore, the experimental samples underpinning the current literature are consistently young. On this point, it is known how older individuals demonstrate age-related changes in motor performance and multi-joint trajectories, with decreasing smoothness and accuracy of the movements (Seider et al., 2002). Therefore, it is also crucial to further assess the possible impact of age on the effectiveness of ECM as a measure of workload in VR.

2. THE PRESENT STUDY

With this study, we aim at covering the identified gaps and learn more about the effectiveness of ECM as a measure of workload in VR. We thus addressed a VR-based industrial robotics framework, and compared results obtained by young (<30 years old) and senior users (>50 years old). Specifically, a total of thirty participants - divided into two age groups - guided an industrial robotic arm through a pick-and-place task in VR by physically moving their arms, under low (single-task) and high (dual-task) mental demands. We also administered the NASA-TLX questionnaire to reference the participants' perceived workload after both task conditions. Based on the literature on behavioral entropy (Reinhardt et al., 2019; 2020) and considering the possible age-related motor performance degradation (Seider et al., 2002), we expect: H1) higher ECM and higher scores on the NASA-TLX questionnaire in the dual-task compared to the single-task, independently from the participants' age; H2) a generally higher ECM in the senior compared to the young group.

3. METHODOLOGY

3.1 Participants

15 individuals (9 females) composed the senior group, who reported being more than 50 years old (M_{age}= 57.1, SD_{age}= 6.2), and 15 individuals (6 females) the young group, who reported being less than 30 years old (M_{age}= 27.8, SD_{age}= 6.4). All participants signed informed consent. The inclusion criteria were the following: absence of past or present neurological/psychiatric disorders, being right-handers, possessing normal or corrected-to-normal vision with contact lenses, and normal color vision. The local ethics committee approved the research methodology, and the study was conducted following the principles of the Declaration of Helsinki. One senior participant was excluded due to limited proficiency with technological devices, which caused extremely long training and a severe inability to perform the task. All participants reported to be inexperienced with VR and telerobotics, particularly the senior once.

3.2 Technical Setup and Experimental Procedure

Participants were provided with HTC VIVE Pro Eye and both its controllers. The virtual environment was programmed in Unity (version 2020.2.1f1). All data was automatically saved on the internal storage of the local laboratory computer at the end of each experimental session.

Before starting the experiment, participants conducted a training session based on the same tasks used in the experimental phase to familiarize themselves with the virtual environment and minimize individual differences related to the ability to use the virtual system. Afterward, during the experimental session, all subjects controlled a robotic arm in VR to execute a pick-and-place task under different demands (single- and dual-task) presented in random order. The young group completed 40 trials for each experimental condition. Differently, for the senior group, the number of trials was lowered to 20, as they showed longer training duration and greater difficulties in familiarizing with the tasks. This may be due to their lack of technological literacy (Wildenbos et al., 2018), which may pose a barrier to the repeated execution of VR-based tasks for older adults.

In the single-task, participants were called to physically drive the robot to pick a bolt from the workstation and place it into a blue box. Figure 1 depicts one complete task trial, which was split into two task phases: the Pick phase, which required higher movement precision to accurately align the robot effector with the bolt, and the Place phase, which required less movement precision as the box where to release the bolt is provided with a larger area. Given that the two task phases demanded different levels of precision of the movement trajectories, we computed the ECM within each of them independently. Oppositely to the single-task, in the dual-task participants additionally performed an arithmetic task to create a higher level of task demand. Specifically, they were presented with a series of digits randomly ranging between 1 and 10 every 2sec, with a jitter of 0.3sec. They were thus asked to sum the numbers all the way through the trial, and then report the result of the mental calculations on a virtual numeric keyboard once the trial was completed.

The robot control modality followed the same mechanisms developed by Nenna et al. (2022; 2023). Specifically, participants approached their right hand to the robot effector and then grasped it by pressing the grip button on the controller. Therefore, they dragged the robot to the desired position by physically moving their own arm within the virtual space, producing a movement trajectory. To enable the picking or placing operations, they then pressed the pad button on the left controller and the robot automatically went down on the workstation to either pick or place the bolt. Once the bolt entered the box, a new bolt and box randomly appeared on the workstation, and a new trial started.

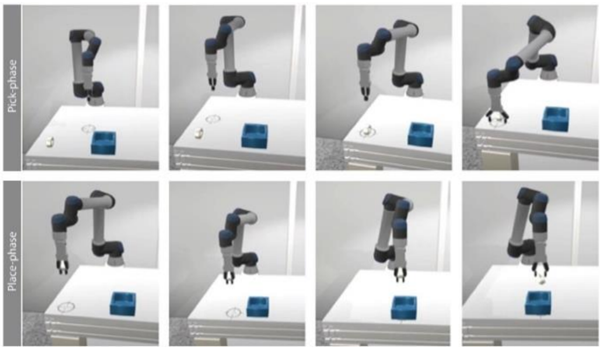

Figure 1. Depiction of one trial of the pick-and-place task, divided into two task phases (i.e., Pick and Place)

3.3 Behavioral and Workload Measures

The levels of perceived mental workload were measured through the NASA-TLX questionnaire (Hart and Staveland, 1988) after the single- and the dual-task. The global workload score was calculated by averaging the scores in its six sub-scales.

For the computation of behavioral entropy, we isolated the movement trajectories performed during the pick and the place action and only included trials exempt from movement disruptions (e.g., the participant lost the robot grip and returned to grasp the robot again, causing fragmented trajectories); therefore, all actions under examination were continuous and uninterrupted. We then choose to determine sample entropy using the method of Richman and Moorman (2000). This method best fits the randomness intrinsic to systems behaving in real-world or complex environments and has been demonstrated to be the preferred method in applied research for mental workload assessments, also in VR scenarios (Reinhardt et al., 2019). Sample entropy can be defined as the negative logarithm of the probability that if two sets of data points of length m are similar, they will remain similar at $m+1$ (Richman and Moorman, 2000). Therefore, if areas in a trajectory that appear similar at one length are no longer similar at a greater length, greater dispersion and complexity are observed in a trajectory, which increases the sample entropy (Hehman et al., 2015). For its computation, we converted the data to normalized time and shifted the absolute controller positions on the three axes (x, y, z) in order to always start from zero (0, 0, 0) in each trial. Based on the overview of statistical tools for calculating sample entropy of Chen and colleagues (2019), we used the function *sample_entropy* in Rstudio (Team, 2020) from the package *pracma* (Borchers, 2022), which allows setting specific parameters like the embedding dimension m (length of sequences to be compared for similarity) and the tolerance r (the threshold for determining similarity between windows). By following the approaches of Hehman et al. (2015) and Reinhardt et al. (2019), we set m=2 and r=0.2; thus, we compared windows with a length of 2, and each sequence was determined to be similar if it was within a tolerance of 0.2 multiplied by the standard deviation of the data. We finally calculated the ECM on the three axes (ECM-X, ECM-Y, ECM-Z) and the ECM-total by averaging the three individual ECMs (Reinhardt et al., 2019).

3.4 Statistical Analysis

All data were analyzed through Generalized Linear Models (GLMs from lme4 package, Bates et al., 2014). We computed a model for each ECM calculation (ECM-X, ECM-Y, ECM-Z, ECM-total) over the factors Task (single-task, dual-task) and Age (young, senior). For the analysis of the NASA-TLX scores, we additionally included the factor Item (mental demand, temporal demand, physical demand, performance, effort, frustration) to further explore possible differences within each of the questionnaire sub-scales. Each model was chosen based on data distribution (Delignette-Muller and Dutang, 2015). Post hoc were performed on each significant interaction with the application of the Bonferroni correction for multiple comparisons (Bonferroni et al., 1936).

4. RESULTS

Table 1. Descriptive statistics of the NASA-TLX questionnaire scores

Dimension	Young Single-task M ± SD	Young Dual-task M ± SD	Senior Single-task M ± SD	Senior Dual-task M ± SD
Global score	5.79 ± 4.85	10.30 ± 5.28	5.54 ± 4.87	11.5 ± 5.89
Mental demand	2.67 ± 1.35	13.80 ± 4.25	4.14 ± 4.44	15.6 ± 2.71
Physical demand	8.40 ± 5.47	11.9 ± 4.48	5.21 ± 4.02	8.14 ± 5.63
Temporal demand	7.27 ± 3.47	7.00 ± 3.40	7.14 ± 4.75	11.2 ± 5.58
Performance	4.33 ± 4.75	7.93 ± 5.64	4.43 ± 3.39	9.36 ± 5.49
Effort	8.27 ± 5.18	13.9 ± 3.83	9.71 ± 6.70	16.8 ± 2.94
Frustration	2.73 ± 2.76	7.47 ± 5.01	2.57 ± 1.40	8.14 ± 5.95

Table 2. Descriptive statistics of the ECMs

Task phase	Axis	Young Single-task M ± SD	Young Dual-task M ± SD	Senior Single-task M ± SD	Senior Dual task M ± SD
Pick	ECM-x	0.026 ± 0.013	0.017 ± 0.015	0.021 ± 0.026	0.010 ± 0.008
	ECM-y	0.199 ± 0.099	0.137 ± 0.086	0.112 ± 0.089	0.104 ± 0.085
	ECM-z	0.093 ± 0.084	0.080 ± 0.041	0.052 ± 0.060	0.033 ± 0.037
	ECM-tot	0.106 ± 0.048	0.078 ± 0.028	0.062 ± 0.040	0.049 ± 0.034
Place	ECM-x	0.044 ± 0.038	0.045 ± 0.042	0.023 ± 0.029	0.019 ± 0.025
	ECM-y	0.216 ± 0.144	0.205 ± 0.145	0.145 ± 0.100	0.111 ± 0.090
	ECM-z	0.064 ± 0.065	0.068 ± 0.072	0.037 ± 0.050	0.028 ± 0.048
	ECM-tot	0.108 ± 0.059	0.106 ± 0.058	0.068 ± 0.042	0.053 ± 0.038

The analysis of the NASA-TLX revealed a main effect for Task (X^2 = 121.58, $p < 0.001$) and Item X^2 = 81.26, $p < 0.001$). Significant interactions were observed between Task and Item (X^2 = 41.78, $p < 0.001$) and Item and Age (X^2 = 16.57, $p < 0.01$). No significant main effect was found for Age ($p = 0.19$). Post-hoc tests on the Task-Item interaction showed significantly higher levels of mental demand ($p<.0001$), effort ($p < 0.0001$), and frustration ($p < 0.001$) in the dual- compared to the single-task. Differently, no significant contrasts were found for the Age-Item interaction. Descriptive statistics are resumed in Table 1. For the analysis of behavioral entropy, we resumed the descriptive statistics in Table 2, the results of the GLMs in Table 3, and all the post-hoc contrasts are depicted in Figure 2.

Table 3. Results of the GLMs performed in the pick and place phases on each ECM measure

Axis	Task phase	Task	Age	Task * Age
ECM-x	Pick	X^2 = 66.66 ***	X^2 = 1.62 ns	X^2 = 4.03 *
	Place	X^2 = 0.04 ns	X^2 = 28.53 ***	X^2 = 3.07 ns
ECM-y	Pick	X^2 = 1.47 ns	X^2 = 7.79 **	X^2 = 1.33 ns
	Place	X^2 = 7.61 **	X^2 = 13.22 ***	X^2 = 10.53 **
ECM-z	Pick	X^2 = 7.27 **	X^2 = 15.44 ***	X^2 = 0.28 ns
	Place	X^2 = .008 ns	X^2 = 36.49 ***	X^2 = 8.76 **
ECM-tot	Pick	X^2 = 15.73 ***	X^2 = 17.38 ***	X^2 = 0.05 ns
	Place	X^2 = 5.52 *	X^2 = 22.85 ***	X^2 = 18.58 ***

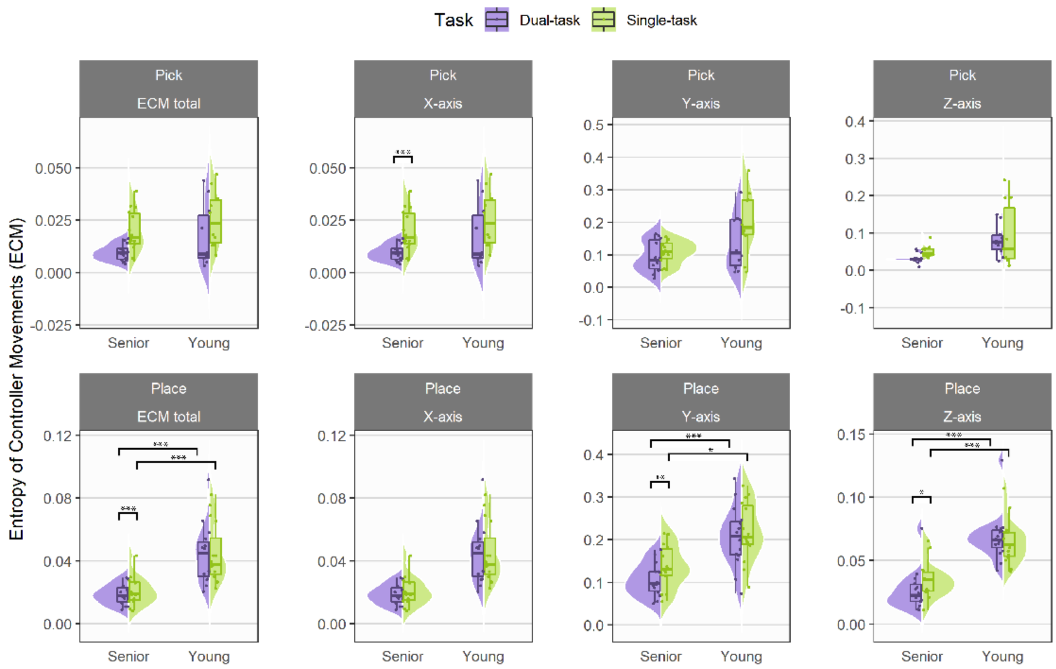

Figure 2. Averaged behavioral entropy in each task phase (pick, place) and task condition (dual-task, single-task), divided by age (senior, young). The dots on the boxplots indicate each participant's averaged entropy. The half violins show the data distribution. The stars indicate the significance levels of the post hoc tests (*$p \leq .05$; **$p \leq .01$; ***$p \leq .001$)

5. DISCUSSION AND CONCLUSIONS

In this study, we applied the concept of behavioral entropy to the trajectory of VR controller movements, exploring its effectiveness in measuring human workload within a VR-based robotic teleoperation scenario. We further explored possible differences between young and senior users, who are prone to decreasing smoothness of movements (Seider et al., 2002), which might affect the effectiveness of ECM as a measure of workload. We thus defined different task demands through the dual-tasking: participants physically drove a robotic arm once as a single task, and once concurrently with an arithmetic task (dual-task). As indicated by the NASA-TLX, regardless of participants' age, the dual-task elicited significantly higher mental demand, effort, and frustration than the single-task. Additionally, young and senior participants self-reported similar workload levels, suggesting no age-dependent differences in perceived workload.

However, unexpectedly, our findings on the ECM deviate from what previously observed (i.e., Chatzithanos et al., 2021; Goodrich et al., 2004; Reinhardt et al., 2019). Specifically, we first hypothesized (H1) higher ECM in the most demanding task condition (i.e., dual-task). Oppositely, we observed systematically higher ECM values in the single-task compared to the dual-task as specifically suggested by ECM-x and ECM-z in the pick phase, ECM-y in the place phase, and ECM-tot both in the pick and place phases, indicating an inverse relationship between the entropy of movement trajectories and the task difficulty. Only senior participants demonstrated slightly higher ECM in the dual- compared to the single-task in the place phase, which aligns with prior research (Chatzithanos et al., 2021; Goodrich et al., 2004; Reinhardt et al., 2019). On this matter, we argue that the secondary task may have intrusively interfered with the primary pick-and-place particularly in the young group, influencing users' motor trajectories. Specifically, in the dual-task, participants summed a series of numbers throughout each pick-and-place action and were always instructed to be as fast and accurately as possible in both tasks. However, the faster they placed the bolt into the box, the fewer mental calculations they had to compute. Differently, in the single-task, they were free to act impulsively while following the instruction of being both fast and accurate. This may have led to more conscious, regular, and smooth motor trajectories in the dual-task to correctly pick and then place the bolt as fast as possible, and more impulsive and dispersive trajectories in the single-task.

An additional note on this regard: previous research, which demonstrated the efficacy of behavioral entropy in detecting varying levels of workload resulting from dual-tasking, was conducted in desktop environments using either a mouse or joypad (Chatzithanos et al., 2021; Goodrich et al., 2004). However, to the best of our knowledge, no study investigated the effectiveness of behavioral entropy in VR contexts leveraging dual-tasking. In their VR-based experiments on behavioral entropy, Reinhardt and colleagues (2019, 2020) always used some inhibition tasks to increase the levels of demand on the user, while no instances of ECM under VR dual-tasking were ever provided. It is thus possible that the presence of a secondary task changes the nature of the primary task, particularly when human physical motion with higher degrees of freedom is involved, like in our VR contexts. This underlines the necessity of a better understanding of behavioral entropy sensitivity in VR, by also considering the impact of secondary tasks on human motion trajectories.

As a second hypothesis (H2), we then expected higher ECM values in the senior group, possibly reflecting their age-related alterations of movement trajectories (Seider et al., 2002). In contrast, we observed consistently higher behavioral entropy in young respect to senior participants. This result is completely reversed compared to our expectations, and therefore, does not allow to define whether it drawn the absence of age-related motor trajectory alterations, or a different behavioral approach to task execution compared to young users.

Nonetheless, as young and senior participants reported similar levels of workload at the NASA-TLX, such findings don't suggest that workload was the cause of differences in motor behavior between young and senior participants. Other factors, such as their particular VR inexperience and low technological literacy, may have influenced their motor behavior. In this view, senior users, being completely new to virtual environments, may have been more cautious when maneuvering the virtual robot, while young users which are typically more familiar with advanced technologies were more impulsive. The study also found that senior users had higher ECM during the dual-task, suggesting they used different strategies than younger users. Furthermore, only senior users demonstrated slightly higher ECM in the dual- compared to the single-task specifically in the place phase. This also indicate that young and senior users have likely adopted different strategies for the task: for easier operations like the place action, senior users actually moved at a more hectic pace under high mental demands, while being smoother under lower demands. The same does not apply to younger individuals, who always kept highly frenetic movement patterns under lower demands.

Overall, behavioral entropy, and ECM in VR, represent valuable unobtrusive and easily computable measures to trace the dispersion of human motion trajectories in digital environments. However, new questions were raised about its effectiveness in reflecting users' workload in VR. Future research should delve into these questions and better understand which are the factors that may affect human motion strategies in VR and to what extent behavioral entropy can inform on users' workloads across different task scenarios.

REFERENCES

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Boer, E. R. (2000, July). Behavioral entropy as an index of workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, No. 17, pp. 125-128). Sage CA: Los Angeles, CA: SAGE Publications.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8,* 3-62.

Borchers, H. W., & Borchers, M. H. W. (2022). Package 'pracma'. *accessed on, 4.*

Chatzithanos, P., Nikolaou, G., Stolkin, R., & Chiou, M. (2021, October). Fessonia: a method for real-time estimation of human operator workload using behavioural entropy. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1325-1331). IEEE.

Chen, C., Sun, S., Cao, Z., Shi, Y., Sun, B., & Zhang, X. D. (2019). A comprehensive comparison and overview of R packages for calculating sample entropy. *Biology Methods and Protocols, 4(1),* bpz016.

Coronado, E., Kiyokawa, T., Ricardez, G. A. G., Ramirez-Alpizar, I. G., Venture, G., & Yamanobe, N. (2022). Evaluating quality in human-robot interaction: A systematic search and classification of performance and human-centered factors, measures and metrics towards an industry 5.0. *Journal of Manufacturing Systems*, *63*, 392-410.

Dautenhahn, K. (2018). Some brief thoughts on the past and future of human-robot interaction. *ACM Transactions on Human-Robot Interaction (THRI), 7(1),* 1-3.

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of statistical software, 64,* 1-34.

Dianatfar, M., Latokartano, J., & Lanz, M. (2021). Review on existing VR/AR solutions in human–robot collaboration. *Procedia CIRP, 97,* 407-411.

Diaz-Piedra, C., Rieiro, H., Cherino, A., Fuentes, L. J., Catena, A., & Di Stasi, L. L. (2019). The effects of flight complexity on gaze entropy: An experimental study with fighter pilots. *Applied ergonomics, 77*, 92-99.

Duguleana, M., Barbuceanu, F. G., & Mogan, G. (2011). Evaluating human-robot interaction during a manipulation experiment conducted in immersive virtual reality. In *Virtual and Mixed Reality-New Trends: International Conference, Virtual and Mixed Reality 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part I 4* (pp. 164-173). Springer Berlin Heidelberg.

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types?. *International journal of psychophysiology, 83(3),* 269-275.

Goodrich, M. A., Boer, E. R., Crandall, J. W., Ricks, R. W., & Quigley, M. L. (2004). *Behavioral entropy in human-robot interaction*. BRIGHAM YOUNG UNIV PROVO UT.

Guo, Y., Freer, D., Deligianni, F., & Yang, G. Z. (2021). Eye-tracking for performance evaluation and workload estimation in space telerobotic training. *IEEE Transactions on Human-Machine Systems*, *52*(1), 1-11.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology (*Vol. 52, pp. 139-183). North-Holland.

Havard, V., Jeanne, B., Lacomblez, M., & Baudry, D. (2019). Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations. *Production & Manufacturing Research*, *7*(1), 472-489.

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes & Intergroup Relations*, *18*(3), 384-401.

Leng, J., Sha, W., Wang, B., Zheng, P., Zhuang, C., Liu, Q., ... & Wang, L. (2022). Industry 5.0: Prospect and retrospect. *Journal of Manufacturing Systems, 65*, 279-295.

McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science, 19(1),* 22-24.

Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). Development of a steering entropy method for evaluating driver workload. *SAE transactions*, 1686-1695.

Nenna, F., Orso, V., Zanardi, D., & Gamberini, L. (2022). The virtualization of human–robot interactions: a user-centric workload assessment. Virtual Reality, 1-19.

Nenna, F., Zanardi, D., & Gamberini, L. (2023). Enhanced Interactivity in VR-based Telerobotics: An Eye-tracking Investigation of Human Performance and Workload. International Journal of Human-Computer Studies, 103079. https://doi.org/10.1016/j.ijhcs.2023.103079

Panagou, S., Neumann, W. P., & Fruggiero, F. (2023). A scoping review of human robot interaction research towards Industry 5.0 human-centric workplaces. *International Journal of Production Research*, 1-17.

Reinhardt, D., & Hurtienne, J. (2018, March). Cursor entropy reveals decision fatigue. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion* (pp. 1-2).

Reinhardt, D., Haesler, S., Hurtienne, J., & Wienrich, C. (2019, March). Entropy of controller movements reflects mental workload in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR*) (pp. 802-808). IEEE.

Reinhardt, D., Hurtienne, J., & Wienrich, C. (2020, March). Measuring Mental Effort via Entropy in VR. In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion* (pp. 43-44).

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology-heart and circulatory physiology.*

RStudio, T. (2020). RStudio: integrated development environment for R. *(No Title).*

Seidler, R. D., Alberts, J. L., & Stelmach, G. E. (2002). Changes in multi-joint performance with age. *Motor control, 6(1),* 19-31.

Stillman, P. E., Shen, X., & Ferguson, M. J. (2018). How mouse-tracking can advance social cognitive theory. *Trends in cognitive sciences, 22(6),* 531-543.

Van der Valk, H., Haße, H., Möller, F., Arbter, M., Henning, J. L., & Otto, B. (2020, August). A Taxonomy of Digital Twins. In *AMCIS*.

Villani, V., Capelli, B., & Sabattini, L. (2018, August). Use of virtual reality for the evaluation of human-robot interaction systems in complex scenarios. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 422-427). IEEE.

Wehrl, A. (1978). General properties of entropy. *Reviews of Modern Physics, 50(2),* 221.

Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). *International journal of medical informatics, 114*, 66-75.

WM Van Gerven, Fred GWC Paas, Jeroen JG Van Merriënboer, Henk G. Schmidt, P. (2000). Cognitive load theory and the acquisition of complex cognitive skills in the elderly: Towards an integrative framework. *Educational gerontology*, *26*(6), 503-521.

Wu, C., Cha, J., Sulek, J., Zhou, T., Sundaram, C. P., Wachs, J., & Yu, D. (2020). Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Human factors, 62(8),* 1365-1386.

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, *58*(1), 1-17.