Cognitive Computing and Internet of Things, Vol. 43, 2022, 64–72
https://doi.org/10.54941/ahfe1001840

Evaluation of Virtual Reality-Based First Responder Training Under Physiological and Cognitive-Emotional Strain

Lilian Reim[1], K. Wolfgang Kallus[1], Martin Pszeida[2], Michael Schneeberger[2], Christian Schönauer[3], Andreas Peer[3], Richard Feischl[4], Georg Aumayr[5], Dietmar Wallner[6], Amir Dini[2], Stefan Ladstätter[2], Anna Weber[2], Alexander Almer[2], and Lucas Paletta[2]

[1]Institut für Begleitforschung und Psychologisches Qualitätsmanagement, Graz, Austria
[2]Joanneum Research Forschungsgesellschaft mbH, Graz, Austria
[3]M[2]D MasterMind Development GmbH, Langenlebarn, Austria
[4]Ing. Richard Feischl GmbH, Gumpoldskirchen, Austria
[5]Johanniter Österreich Ausbildung und Forschung gemeinnützige GmbH, Wien, Austria
[6]FH JOANNEUM – University of Applied Sciences, Bad Gleichenberg, Austria

ABSTRACT

First responders are faced with various challenges and critical situations during their operations. Virtual Reality (VR) training environments offer a great opportunity to create realistic operational scenarios with appropriate levels of difficulty. To evaluate VR-based training regarding cognitive-emotional and physical strain, a pilot study with 13 first responders (firefighters and paramedics) was conducted. The study plan aimed to induce equivalent strain in both, real environment ('RC') and VR-based ('VR'), training scenarios. In summary, the results indicate that training sessions in real and in VR-based environments have a comparable impact on physical strain, on different dimensions of workload like physical and cognitive demand, and specific physical and emotional states. Most important, we could show that physical strain, which is a critical factor for appropriate training of first responders, cannot only be induced by traditional real-life training, but also by an immersive VR training environment. These results confirm and motivate the use of VR environments for the training of stress-resilient decision-making behaviour for emergency forces.

Keywords: Physical strain, Cognitive-emotional strain, Virtual reality, Training system, First responders

INTRODUCTION

For first responders, the increasing number of critical operations pose enormous challenges today. Environmental as well as technical disasters are about to increase due to climate changes, technical, and social developments. Being prepared for such critical situations through an appropriate training is essential for the safety and success of first responders (Bertram et al., 2015). Advanced training of first responders and emergency staff with typical

© 2022. Published by AHFE Open Access. All rights reserved.

operation scenarios is one of the upcoming challenges of the near future. Traditional training with real environment operation scenarios involves a high expenditure of resources and only allows limited training frequencies and thus limiting variability in the experience of critical and decision-relevant situations. VR-based training is cost-effective, safe to use and offers new perspectives to provide a wide variety of training environments (Engelbrecht et al., 2019). In VR and in augmented reality simulations, scenarios can be adaptively set to create stressful situations in a targeted manner and with appropriate levels of difficulty. For VR-based scenarios the core elements of the real-world scenario have to be represented, which can be easily obtained for the visual channel. However, as soon as psychomotor activity has to be included, it becomes more difficult, as VR is usually spatially limited. Modern VR-Equipment allows including physical activity like running or walking. However, the differences and similarities between VR training and real-life training are not fully understood yet. The study at hand throws some light on this area, as similar tasks were executed in a real-life scenario and a VR scenario.

A prerequisite for an efficient use of VR-based training is the measurement of the imposed strain of VR technology in comparison with real environment trainings. Specific areas like flight or automotive industry developed sophisticated simulator facilities to develop, promote and calibrate support systems for the operators. Interesting results could be obtained for exercises in flight training with psychophysiological multi-level assessments. In a study from Koglbauer et al. (2010) real flight scenarios were compared with equivalent scenarios in a flight simulator. Unusual attitude recoveries yielded comparable differences between different exercises in real flights and a simulator, which was set to the same flight characteristics. However, the psychophysiological change showed smaller amplitudes in all exercises in the simulator. Lessons learned from simulator trainings should be incorporated in the development of the scenarios, which should not provide negative training effects or negative effects on the trainees by causing simulator sickness or related problems. In addition, physical strain aspects like heat strain, physical exhaustion and management of physical load in adverse conditions like smoke, should be taken into account for the development of advanced training programs. Therefore, the aim of this study was to develop realistic VR-scenarios that contain relevant facets of first responder operations and to evaluate these scenarios under different aspects of physical, cognitive and emotional stress parameters as well as different psychological states.

VIRTUAL REALITY TRAINING SYSTEM

State-of-the-art. Recently, VR-based training in disaster preparedness has been increasingly recognized (Freeman et al. 2001) as an important additional modality to traditional real-life skill training. Multiple studies (Cone et al. 2011; Kurenov et al. 2009) have highlighted VR applications in disaster training. The increased realism in the practice enables first responders to reinforce their individual performance, in particular, to execute tasks appropriately under stress and apply decision making under conditions close

Figure 1: Physical and cognitive-emotional strain during the field study with first responders under 'Reality Condition' (RC) and 'VR-Condition' (VR), respectively.

to reality. The immersive capability of VR-based training offers more realism than classroom-based instructive teaching and substantially reduces the resource burden of real-life drills and tabletop exercises. Mills et al. estimated that a mass casualty triage training of paramedic students in a real-world simulation is about 13 times more expensive than in VR, while the simulation efficacy has been found near identical (Mills et al. 2019). Recent research has even indicated superior performance in simple search tasks following VR and augmented reality (AR) training of first responders as opposed to traditional classroom and real-world training in an ambulance bus (Koutitas et al. 2020).

VR training system. The VR training system in this work is based on the VROnSite platform (Mossel et al., 2020) that supports immersive training of first responder units' on-site squad leaders. This training platform is fully immersive, entirely untethered to ease use and provides two means of navigation—abstract and natural walking—to simulate physical stress and exhaustion. The development of the software has been closely interlocked with stakeholders from multiple fire brigades to gather early feedback in an iterative design process and is commercially available by the Austrian company M^{2}D MasterMind Development GmbH. In the context of the presented work, VROnSite was used to develop several scenarios that were specifically adjusted for the session on cognitive-emotional and physical strain in both conditions, RC and VR, respectively (see Figure 1).

Scenario development. In particular, four scenarios were prepared by means of videos with a duration 2 minutes and experience of virtual drive through the scenario. These scenarios were determined to be representative to evaluate the first responders' situation awareness. Each video was presented and then the first responder had to provide a report on the perceived scenario following the LEDVV command scheme with the key aspects of 'L' for 'Lage' (situation), 'E' for 'Entschluss' (decision making), 'D' for 'Durchführung' (execution), 'V' for 'Versorgung' (supply), and 'V' for 'Verbindung'

Table 1. Overview about crossover design: Groups, point of measurements and conditions.

	Points of Measurement (time)		
	Baseline (t0)	**Training procedure 1 (t1)**	**Training procedure 2 (t2)**
Group A (N = 6)	Baseline	Reality	VR
Group B (N = 7)	Baseline	VR	Reality

(communication). Psychophysiological recordings allowed the assessment of strain, as the procedure was designed in parallel to the Trier Social Stress Protocol (Kirschbaum et al., 1993). The first responder participants considered the videos highly suitable for training of decision making in complex first responder scenarios.

METHOD

Sample & Experimental design. Participants (N = 13, 11 males, 2 females, age: 17-70, M: 38.85 SD: 15.02) consisted of representative emergency forces of the Austrian volunteer fire brigade of Gumpoldskirchen (n = 7) as well as paramedics of the Johanniter organization Vienna (n = 6). The study plan aimed to induce equivalent strain in both, real environment training scenarios ('Reality Condition' – 'RC') and VR-based training scenarios ('VR Condition' – 'VR'). Therefore, two identical training procedures were developed, each of which was in turn divided into a (i) cognitive-emotional and (ii) a physical strain block. Both blocks were repeated twice in each procedure. Each participant carried out both training procedures, RC and VR. To avoid order effects, a crossover design was used, where participants were assigned to two groups, whereas group A received the first training procedure in RC and the second training procedure in VR, and group B the other way around (see Table 1). The study took ~ 160 minutes for each participant to complete.

Strain induction. Physical strain was induced by a 5-minute endurance run on the emergency forces' test site in the Reality Condition and on a VR-supported treadmill, using the commercially available immersive VR-training system described above, in the VR Condition. In both conditions, participants wore their operational clothing and a 20 kg back-package, which resembled usual operational equipment.

Cognitive-emotional strain was induced by the attentive perception of relevant events in operation scenarios and subsequent request for a situation report and time-critical reporting in the face of the team leader, as described above. After watching the scenarios, participants were asked to wait for one minute, before they could provide their report. This method was adapted from the Trier Social Stress Protocol (Kirschbaum et al., 1993), which has been successfully used to induce stress (Rodrigues et al., 2018).

Instruments. For the evaluation of physical strain, the Borg-Scale of perceived exertion (Borg, 1998; 2004) was used. The Borg-Scale was submitted in the introduction phase and after each endurance run. To measure cognitive performance a 3-minute version of the psychomotor vigilance task (PVT;

Dinges & Powell, 1985; Basner et al., 2011) was implemented on a touchscreen tablet. The PVT is a reaction time test and widely used measurement for cognitive performance, fatigue and reduced alertness (Grant et al., 2017; Basner et al., 2011). Lapsed stimuli, defined by reaction times > 400 ms were analysed. To reduce learning effects in the data, participants practiced the PVT at the beginning of the study. Then PVT was submitted three times, before and after both training sessions. Perceived workload was measured by NASA-TLX (Task Load Index; Hart & Staveland, 1988), which asses-ses six dimensions of workload: Mental demand, physical demand, temporal demand, performance, effort, and frustration. NASA-TLX was submitted after both training procedures. Further psychological variables were assessed to gain an understanding of the physical and psychological state and conditions of the participants. Physical symptoms were assessed by five dimensions of the "Multiple Physical Symptom List" (MKSL; Erdmann & Janke, 1984), e.g., adrenergic symptoms, pain and nausea. Psychological states were measured by the BSKE (Janke et al., 1986) which assess positive and negative states by 8 dimensions, e.g., activation, excitement, anxiety/sadness, irritation. MKSL and BSKE where submitted three times, before, and after each training procedure. Psychophysiological recordings were performed for heart rate and electrodermal activity (see Schneeberger et al., 2022 for results).

RESULTS

The study was a 2 (Group A / Group B) x 3 (Points of Measurement: t0 / t1 / t2) crossover design experiment with RC and VR as Training Conditions (see Table 1). Two-way repeated measure ANOVAs with group (A vs. B) as fixed factor and time (baseline (t0) vs. training procedure 1 (t1) vs. training procedure 2 (t2)) as repeated measures factor, were used. To examine dependent variables for time and order effects, main effects of time and interaction effect between time and groups were analysed. To analyse the effect of conditions only, control analyses were done. Therefore, data was restructured and conditions from both groups were combined and one-way repeated measure ANOVAs with conditions (Baseline (BL) / RC / VR) were calculated. Post-hoc analyses were performed using Bonferroni correction for multiple comparisons.

For measuring physical exertion, Borg-results were analysed (see Figure 2). A two-way repeated measure ANOVA showed a significant effect for time ($F_{1.5, 16.1} = 59.21$; $p = .00$; $\eta^2 = .84$). Posttests for time effects showed significant differences ($p = .00$) between baseline and each following point of measurement, but no other significant differences between points of measurement. There was no significant interaction effect between time and group ($F_{1.5, 16.1} = .97$; $p = .37$; $\eta^2 = .08$). Regarding the control analyses for conditions, the only effects we could find, was between baseline and both conditions ($p = .00$), but there was no significant difference between VR and RC ($p = 1$).

Cognitive Performance was analysed, using the number of lapsed stimuli of PVT (> 400 ms). Repeated measure ANOVAs revealed no significant main effect of time ($F_{2,22} = .07$; $p = .94$; $\eta^2 = .01$), no significant interaction effect

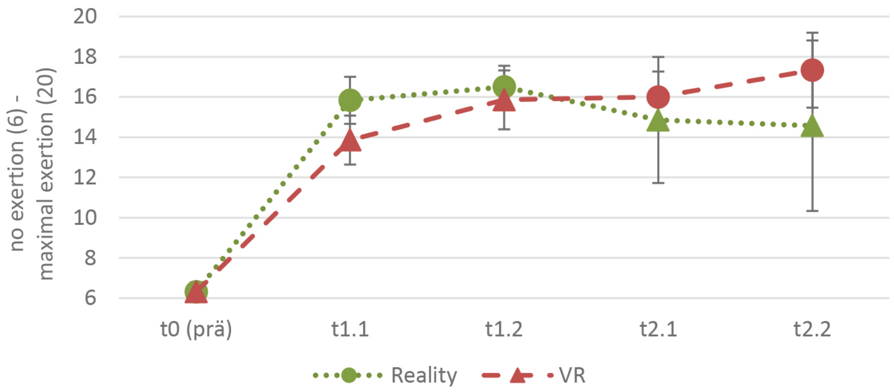

Figure 2: Borg Scale: Perceived Exertion, mean and standard deviation for points of measurement, conditions and groups (group A (N = 6): ●; group B (N = 7): ▲).

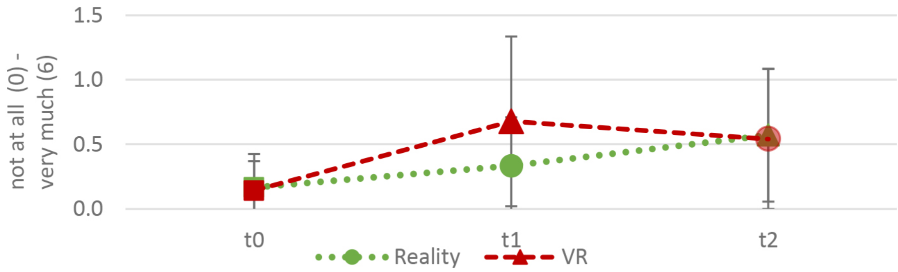

Figure 3: Physical states (MKSL): Adrenergic symptoms, mean and standard deviation for points of measurement, conditions and groups (group A (N = 6): ●; group B (N = 7): ▲).

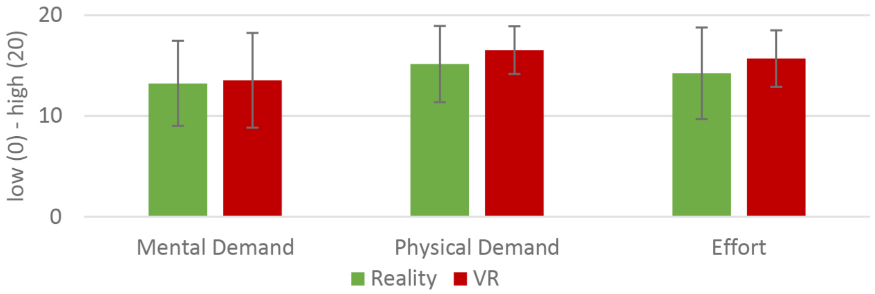

Figure 4: NASA-TLX: mental & physical demand, effort; mean and standard deviation for conditions, N = 13.

of time and group ($F_{2,22}$ = .08; p = .93; η^2 = .01) and also no significant effect for condition ($F_{2,24}$ = .07; p = .93; η^2 = .006; BL: M = 19.37; SD = 23.29; RC: M = 18.16; SD = 20.39; VR: M = 19.25, SD = 16.90).

Results of BSKE and MKSL indicated that training procedures in both conditions, VR and RC, had an impact on specific psychological states and physical symptoms. For example, a two-way repeated measure ANOVAs revealed significant time effects for adrenergic symptoms ($F_{2,22}$ = 7,29; p

= .00; $\eta^2 = .40$, see Figure 3) but no significant interaction effects between time and group ($F_{2,22} = 1.51$; $p = .24$; $\eta^2 = .12$). Post-tests for time effects showed only significant differences between baseline and following points of measurement after the training (t0-t1: $p = .04$; t0-t2: $p = .02$) but no significant difference between training procedures (t1-t2: $p = 1$). Control analyses for conditions, only showed a significant difference between baseline condition and both training conditions (BL-RC: $p = .03$; BL-VR: $p = .02$) but no significant differences between VR and RC ($p = .41$). Comparable results were also found for pain and excitement. No significant results were found for nausea/sickness, activation and irritability.

For NASA-TLX, two-way ANOVAs revealed no significant main effects of time and no significant interaction effects for all dimensions of workload (e.g., mental demand, time: $F_{1,11} = .01$; $p = .91$; $\eta^2 = .00$; time*group: $F_{1,11} = .06$; $p = .81$; $\eta^2 = .01$). For control analyses, dependent samples t-tests showed no significant differences between VR and RC (see Figure 4; e.g., mental demand: $t = -.46$, $p = .81$).

CONCLUSION

The aim of this study was to evaluate a virtual reality-based training for first responders. Therefore, the effects of physiological and cognitive-emotional strain for training sessions in real environment scenarios and VR-based operation scenarios were assessed and analyzed. The most important outcome of this study were the effects of perceived physical exertion measured by the Borg-Scale. Because first responders are confronted with different aspects of physical strain, like fire fighters that attain an intense level of physical activity quickly and maintain that level as long as they are actively engaged in fighting fire (Manning & Griggs, 1983; see also Sanquist, 2015), it is important to take physical strain into account when developing new training procedures. Results showed that the perceived exertion of the participants after 5-minute runs in reality was at a similar level as after 5-minute runs in an immersive VR-training system. The results of NASA-TLX regarding physical demand, point in the same direction, suggesting that both training environments had a comparable effect on the perceived physical demand.

Comparable effects were found for some physical symptoms like pain and adrenergic symptoms and mental states like excitement, indicating that both training conditions had an influence on some physical and psychological states of participants. As Virtual Reality sickness can be a critical factor in VR-based training (Chang et al., 2020) we assessed perceived feelings of sickness/nausea but could not find any significant effects of the VR Condition on sickness.

We expected that the physical, emotional and cognitive strain induced by training procedures would have a negative effect on cognitive performance, measured by PVT, but could not find any effects there. We assume that the physical exercises had a positive impact on cognitive performance. Meta-analyses (Lambourne & Tomporowski, 2010) suggest that physical exercises, due to a heightened level of arousal, lead to improved cognitive performance, but also point out to the complex relation between exercise and cognitive performance, which can be influenced by type and length of exercise or point of measurement. Regarding the perceived mental demand, measured by NASA-TLX, we could find comparable results, indicating that both training conditions, VR and RC, influenced the perceived mental workload of the participants.

In summary, the results indicate that both training sessions, in reality and in VR environments, have a comparable impact on physical strain, on different dimensions of workload like physical and cognitive demand, and specific physical and emotional states. Most important, we could show that physical strain, which is a critical factor for appropriate training of first responders, can be induced by an immersive VR training environment similarly to traditional real-life training. These results confirm and motivate the use of VR environments for the training of stress-resilient decision-making behaviour for emergency forces.

ACKNOWLEDGMENT

This work was partly funded from project VR-Responder under grant No 873522 of the Austrian FFG KIRAS programme, co-financed by the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK), and from the project SIXTHSENSE under grant No 883315 of the Horizon 2020 research and innovation programme of the European Commission.

REFERENCES

Basner, M., Mollicone, D., & Dinges, D. F. (2011). "Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation", Acta Astronautica, 69(11-12), 949–959.

Bertram, J., Moskaliuk, J., & Cress, U. (2015). "Virtual training: Making reality work?", Computers in Human Behavior, 43, 284–292.

Borg, G. (1998). "Borg's perceived exertion and pain scales", Human Kinetics.

Borg, G. (2004). "Anstrengungsempfinden und körperliche Aktivität", Deutsches Ärzteblatt, 101(15), 1016–1021.

Chang, E., Kim, H.T., Yoo, B. (2020) "Virtual Reality Sickness: A Review of Causes and Measurements", International Journal of Human–Computer Interaction, 36:17, 1658–1682, DOI: 10.1080/10447318.2020.1778351

Cone, D.C., Serra, J., Kurland, L. (2011) "Comparison of the SALT and Smart triage systems using a virtual reality simulator with paramedic students", Eur J Emerg Med 18(6):314–21. https://doi.org/10.1097/MEJ.0b013 e3283 45d6f d

Dinges, D. F., & Powell, J. W. (1985). "Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations", Behavior research methods, instruments, & computers, 17(6), 652–655.

Engelbrecht, H., Lindeman, R. W., & Hoermann, S. (2019). "A SWOT analysis of the field of virtual reality for firefighter training", Frontiers in Robotics and AI, 101.

Erdmann, G., & Janke, W. (1984). "Die mehrdimensionale körperliche Symptomliste (MKSL)", Berlin, Germany: TU Berlin.

Freeman, K. M., Thompson, S.F., Allely, E.B., Sobel, A.L., Stansfield, S.A., & Pugh, W.M. (2001). "A virtual reality patient simulation system for teaching emergency response skills to US Navy medical providers", Prehospital and Disaster Medicine, 16(1), 3–8.

Grant, D. A., Honn, K. A., Layton, M. E., Riedy, S. M., & Van Dongen, H. (2017). "3-minute smartphone-based and tablet-based psychomotor vigilance tests for the assessment of reduced alertness due to sleep deprivation", Behavior Research Methods, 49(3), 1020–1029.

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, 139–183, North-Holland.

Janke, W., Debus, G., Kallus, K. W., Hüppe, M., & Schmidt-Atzert, L. (1986). "Befindlichkeitsskalierung nach Kategorien und Eigenschaftswörtern (BSKE (EWL)-ak-24)", Julius-Maximilian-Universität Würzburg.

Kirschbaum, C., Pirke, K.M., & Hellhammer, D.H. (1993). "The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting", Neuropsychobiology, 28(1-2), 76–81.

Koglbauer, I., Kallus, K.W., Braunstingl, R., & Boucsein, W. (2011). "Recovery training in simulator improves performance and psychophysiological state of pilots during simulated and real visual flight rules flight", The international journal of aviation psychology, 21(4), 307–324.

Koutitas, G., Smith, S., Lawrence, G. (2020). "Performance evaluation of AR/VR training technologies for EMS first responders", Virtual Real. https://doi.org/10.1007/s1005 5-020-00436 -8

Kurenov, S.N., Cance, W.W., Noel, B., Mozingo, D.W. (2009). "Game-based mass casualty burn training", Studies Health Technol Inf 142:142–144.

Lambourne, K., & Tomporowski, P. (2010). "The effect of exercise-induced arousal on cognitive task performance: a meta-regression analysis", Brain Research, 1341, 12–24.

Manning, J.E., & Griggs, T.R. (1983). "Heart rates in firefighters using light and heavy breathing equipment: similar near-maximal exertion in response to multiple work load conditions", J of Occupational Health, 25(3), 215–218.

Mills, B., Dykstra, P., Hansen, S., Miles, A., Rankin, T., Hopper, L., Brook, L., Bartlett, D. (2019). "Virtual reality triage training can provide comparable simulation efficacy for paramedicine students compared to live simulation-based scenarios", Prehospital Emerg Care. https://doi.org/10.1080/10903 127.2019.16763 45.

Mossel, A., Schoenauer, C., Froeschl, M., Peer, A., Goellner, J., Kaufmann, H. (2021). "Immersive training of first responder squad leaders in untethered virtual reality", Virtual Reality, 25:745–759, https://doi.org/10.1007/s10055-020-00487-x.

Rodrigues, S., Paiva, J. S., Dias, D., Pimentel, G., Kaiseler, M., & Cunha, J. P. S. (2018). "Wearable biomonitoring platform for the assessment of stress and its impact on cognitive performance of firefighters: an experimental study", Clinical Practice and Epidemiology in Mental Health: CP & EMH, 14, 250.

Sanquist, T.F. (2015). "Remote Physiological Health and Status Monitoring of First Responders: Promises, Practicalities, and Prospects", Technical Report of Pacific Northwest National Laboratory, February 2015.

Schneeberger, M., Paletta, L., Kallus, K.W., Reim, L., Schönauer, C., Peer, A., Feischl, R., Aumayr, G., Pszeida, M., Dini, A., Ladstätter, S., Weber, A., Almer, A., Wallner, D. (2022). "First Responder Situation Reporting in Virtual Reality Training with Evaluation of Cognitive-emotional Stress using Psychophysiological Measures", In: Ayaz H., Paletta, L., eds., Advances in Neuroergonomics and Cognitive Engineering, Taylor & Francis, in print.