MASTER - A multimethod system for the assessment and training of teamwork in simulated scenarios

Contract n° AO2016-07

Partners

Disfor - Department of Education Sciences, University of Genoa	Italy
ENEL DISTRIBUZIONE spa (ENEL D)	Italy
Neuroscience of Emotion and Affective Dynamics Lab (NEAD), affiliated to the Swiss Center for Affective Sciences (SCAS), University of Geneva	Switzerland

Project coordinator

Bracco Fabrizio PhD
Department of Education Sciences, University of Genova, corso Andrea Podestà, 2, 16128, Genova (Italy)
tel: +39 010-20953704
fax: +39 010-20953728
fabrizio.bracco@unige.it

Contributors to the deliverable

Fabrizio Bracco, Michele Masini, Tommaso F. Piccinno, Donald Glowinski, Simon Schaerlaken

Deliverable 2

- Resume of deliverable 1
- Data analysis method
- Results and comments
- Conclusions

Resume of deliverable 1

In Deliverable 1 the following topics were addressed:

- *simulation as a training method for safety*: we described the method of simulation as a valid tool to improve safety in complex sociotechnical systems. In particular, we described the non-technical skills (NTS) as part of the broader situated professional skills (SPS)
- *simulation based training for SPS in the electricity domain*: we described how simulation has been adopted as a training method for electricity workers, in particular concerning SPS
- *definition of SPS profile of electrical workers*: we described the development of the SPS questionnaire to assess the capacity of workers to think about safety procedures during their daily working activity
- *development of simulation scenarios*: we described the procedure adopted to develop the scripts of the simulated scenarios, focusing on the SPS addressed by each scenario, and the issue of the standardized client
- *development of non-verbal cues (NVC) tracking system*: we described the setting of the training ground, the remote control room and the complete device list
- *procedure*: we described the scenario's setup, the setting, and the debriefing phase

Data analysis method

The five scenarios have been assessed with three methods:

- Training efficacy evaluation
- subject matter expert (SME) assessment of the performance through video analysis
- NVC assessment through sensors recording physiological data

Training efficacy evaluation

We adopted the Kirkpatrick's Evaluation of Training Model (1959) to assess the efficacy of the training method. The model states there are different outcomes of a training. These outcomes are classified into 4 levels:

- Level 1 - Reaction: to what degree participants react favorably to training
- Level 2 - Learning: to what degree participants acquire intended knowledge, skills or attitudes
- Level 3 - Behaviour: to what degree participants apply what they learned during the training when they are back on the job
- Level 4 - Results: to what degree targeted outcomes occur as a result of the training

Early outcome levels (Reaction, Learning and Behaviour) are mandatory prerequisites to accomplish the long term effects on organizational results (level 4) such as increasing safety, which is usually the real goal of a training. In the present study level 4 outcomes are increasing safety, reducing accidents and, thus, reducing injuries/economic losses. These outcomes can not be seen immediately after the training. Usually it takes a quite long period of time to be able to assess the impact of an intervention on safety. Furthermore, the training should be implemented on all the population of workers and adopted as recurrent training to be more effective. For example, simulation training is mandatory and repeated over

a specific amount of time for pilots in the aviation domain. This approach contributed to heavily reducing accidents (de Winter, Dodou, Mulder, 2012) . The study conducted is a simple pilot study, which involved a small number of people (N = 26) in four different training sessions. None of the participants repeated the training. Therefore it was not possible to assess the organizational outcome for both reason of time (it would have been necessary a much longer period to collect enough data) and statistical power (it would need many more participants to test the efficacy on the number of accidents/incidents).
We limited our evaluation of the efficacy of the training on the first level of Kirkpatrick's model by asking the participants to fill in a questionnaire about their experience during the training, its perceived usefulness and the satisfaction of attending it.
We did not evaluate the second level of Kirkpatrick's model because we did not want to include an evaluation of the competence/performance of the participants in the simulation training. We assume that the efficacy of the simulation method largely depends on creating a safe and non-judgemental place to try, discuss and, even fail.
Furthermore we wanted to evaluate the third level of Kirkpatrick's model (behaviour), but it was not possible to observe the participants on the job after the training. The most suitable way to collect this data was a self-report questionnaire about the application of the Situated Professional Skills of electrical workers. Unfortunately there was no such validated questionnaire. So we decided to begin the validation process of the SPS questionnaire.

Perceived usefulness and satisfaction of participants

An 11-item questionnaire was administered immediately after the end of each training session. The questionnaire aimed to evaluate the experience of the participants. The items measured three main areas (1) quality of the training, (2) Perception and feeling of the participants and (3) perceived usefulness. The questions used in the questionnaire are listed below.

1. Quality of the training
    a. Do you feel the trainers adopted a non-judgemental and open approach?
    b. Do you think simulation scenarios were similar to your typical working situation?
    c. To what extent simulation scenarios were realistic?
2. Perception and feelings of the participants
    a. How much did you feel committed and inspired by the training?
    b. How satisfied are you with having attended the training?
    c. Are you interested in attending other similar training in the future?
3. Perceived usefulness of the training
    a. Overall, to what extent do you think the training is useful for your work?
    b. Usefulness of the first part of the training (introduction)
    c. Usefulness of participating in simulated activities
    d. Usefulness of watching simulated activities
    e. Usefulness of discussion after simulated activities (debriefing)

Participants were asked to answer on a 5-point Likert scale (Not at all, A few, Moderately, A lot, Completely).
A one sample t test was performed for each item to evaluate if the mean score was significantly different from the median point of the response scale (3). Table 1 shows all items had an average score statistically higher than 3. Participants considered the training as a valuable event: scenarios were perceived as realistic, participants felt engaged and satisfied and the training was evaluated as useful.

Item	*N*	*Mean*	*SD*	*Median*	*t*	*df*	*sig.*	*d*
Non-judgmental approach	26	4.81	0.40	5	22.934	25	< 0.001	4.53
Scenarios are similar to working situations	25	3.72	0.84	4	4.272	24	< 0.001	0.86
Realism of scenarios	26	3.77	0.59	4	6.682	25	< 0.001	1.31
Commitment	26	4.42	0.76	5	9.579	25	< 0.001	1.87
Satisfaction for the training	26	4.35	0.80	5	8.611	25	< 0.001	1.69
Intention to participate to other similar trainings	26	4.00	1.10	4	4.655	25	< 0.001	0.91
Usefulness of the training	26	4.00	0.85	4	6.001	25	< 0.001	1.18
Usefulness of the introduction	26	4.08	0.69	4	7.977	25	< 0.001	1.57
Usefulness of participating in simulated activities	25	4.40	0.87	5	8.083	24	< 0.001	1.61
Usefulness of watching simulated activities	26	4.35	0.75	4	9.211	25	< 0.001	1.80
Usefulness of debriefing	26	4.46	0.76	5	9.799	25	< 0.001	1.91

Table 1 - statistics of the items of the post-training questionnaire

SPS questionnaire

We developed the items of the SPS questionnaire starting from the definition of the SPS profile of electrical workers. The items of the questionnaire were partially derived from the SPS checklist. Other items were newly developed because using a self report questionnaire allows to ask information about the cognitive processes, whereas the checklist used to evaluate the performance must refer only to observable behaviours.

The first version of the questionnaire counted 28 items with a 5-point frequency Likert scale (Never, Rarely, Sometimes, Often, Always). Each item asked how often the participant performed a specific behaviour in his work. Each behaviour referred to the application of one or more SPS of the electrical workers.

An example of the operationalization of the SPS “Risk Assessment” is the item “I check at least twice before performing a potentially dangerous operation”. Annex XX is the complete set of items.

The questionnaire was administered to 188 electrical workers employed by e-distribuzione

Four items did not receive an answer on some response point, thus they were deleted from the questionnaire.

An Exploratory Factor Analysis was conducted on the polychoric correlation matrix of the 24 items.

Different factorial solutions were considered (two, three, four factors). Four factors were considered as the optimal number of factors, based on the RMSEA and TLI. Unfortunately, it was not possible to have a clear loadings pattern; furthermore, it was not possible to identify a theoretical model that could explain how the items clustered together in factors. Thus, this preliminary study could not lead to a definitive form of the questionnaire. The items have to be reviewed to assess their adequacy. A new study has to be conducted with a new sample of participants to test a revised version of the questionnaire.

SME assessment of performance

In order to have a detailed analysis of the behavior of each participant to the simulation, we performed a further assessment with the support of an expert of the electrical operations, provided by e-distribuzione. The assessment consisted in the analysis of each video recording, tracking each behavior and classifying it as compliant, partially compliant or non-compliant.
By compliant we meant that the observed behaviour was completely adherent to the operational rules and procedures for that specific activity. The behaviour was rated as partially compliant when the actions were not completely adherent with the procedures, because some steps in the procedure had been missed or switched. The behaviour was rated as non compliant when the procedure was completely neglected or the workers skipped some safety-critical steps in the procedure.
The analysis of each scenario lasted between 4 and 5 hours, and consisted in the observation step by step of the simulation, and deep discussion with the SME, facilitated by a psychologist. In table 2 we reproduce an example of analysis; the complete set of analyses will be annexed to the deliverable.
Each assessment was composed of:

- Time in the video where the specific behavior was initiated
- Phase of the task
- Description of the activity in the phase (when necessary)
- SME evaluation of the behavior about its compliance with operational procedures
- Explanation of the evaluation
- Description of the behavior (in case of multiple-operator scenarios, this column was replicated for each participant)
- Client action description, when relevant
- Notes in addition to the assessment

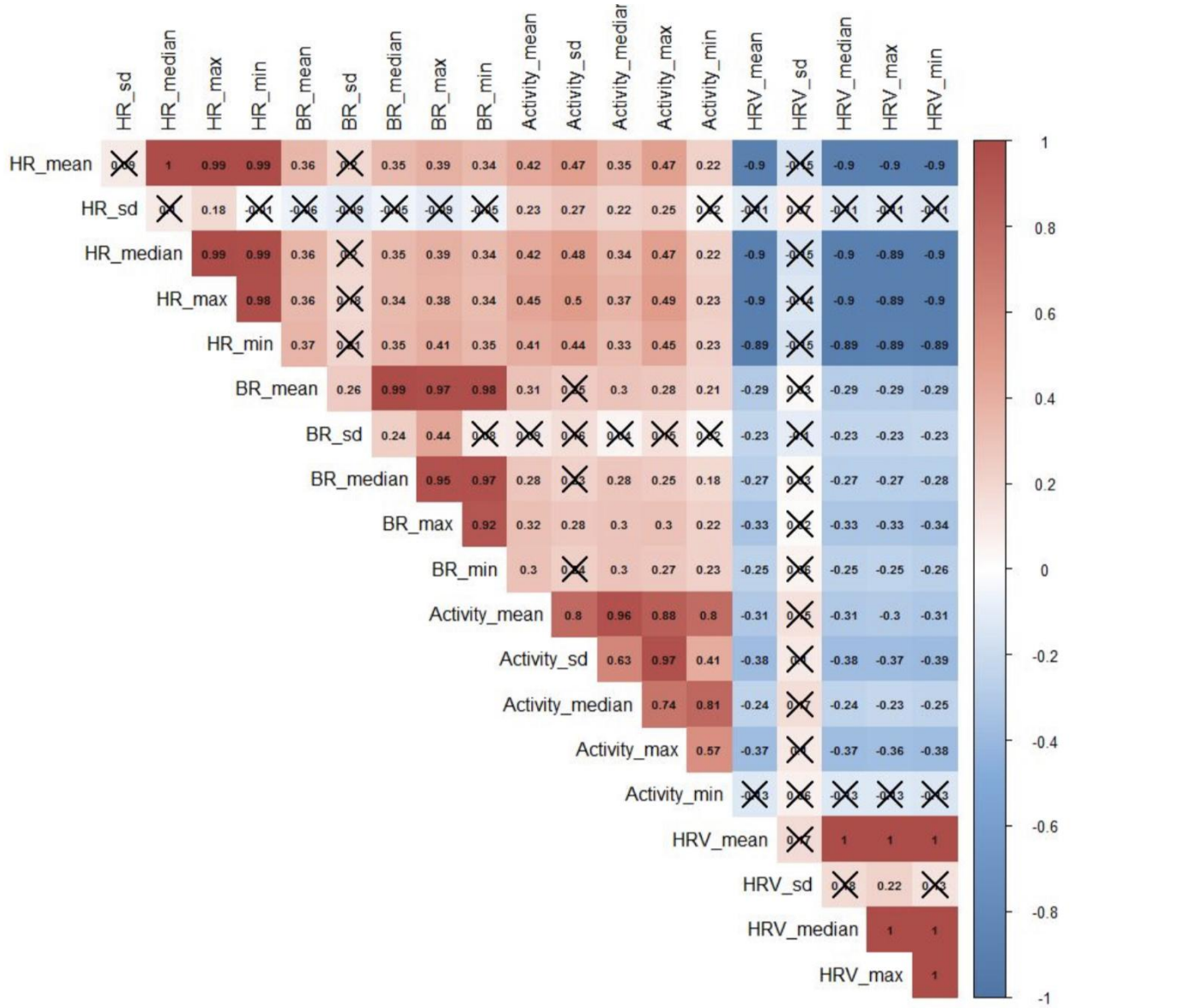
Time	Phase	Description	Evaluation	Evaluation Explanation	1° operator	Client action	Notes
0'00	Analysis	Situation assessment	compliant		RDL analysis		
0'35	Analysis	Situation assessment	compliant		Pushes button on CE		
0'42	Preparation		partially compliant	non-delimited working area	Takes the tools		
1'00	Preparation		partially compliant	shirt out of the pants	Buttons up jacket		
1'36	Preparation		partially compliant		Explains the problem to the client	Start asking question about the intervention; he didn't respect the safety distance (due to the lack of delimitation)	

1'56	Preparation		partially compliant		Asks the client to stay at a safe distance		
2'00	Deprogrammation		non-compliant		Starts deprogramming (using the specific device)		
2'14	Deprogrammation		non-compliant	forgot to perform the evaluation	1st isolation glove		he takes off the glove after some seconds
2'16	Deprogrammation		non-compliant		Asks the client to stay at safe distance		
2'56	Deprogrammation		non-compliant	forgot to perform the evaluation	1st isolation glove (again)		
3'24	Deprogrammation		non-compliant	forgot to perform the evaluation	Puts on both isolation gloves		
3'28	Deprogrammation		non-compliant		lowers the visor		
4'10	Work under electric voltage		non-compliant		starts unscrewing the EC		
5'38	Work under electric voltage		non-compliant	forgot to check the detector	uses the voltage detector		
13'19	Work under electric voltage		non-compliant		End of scenario		

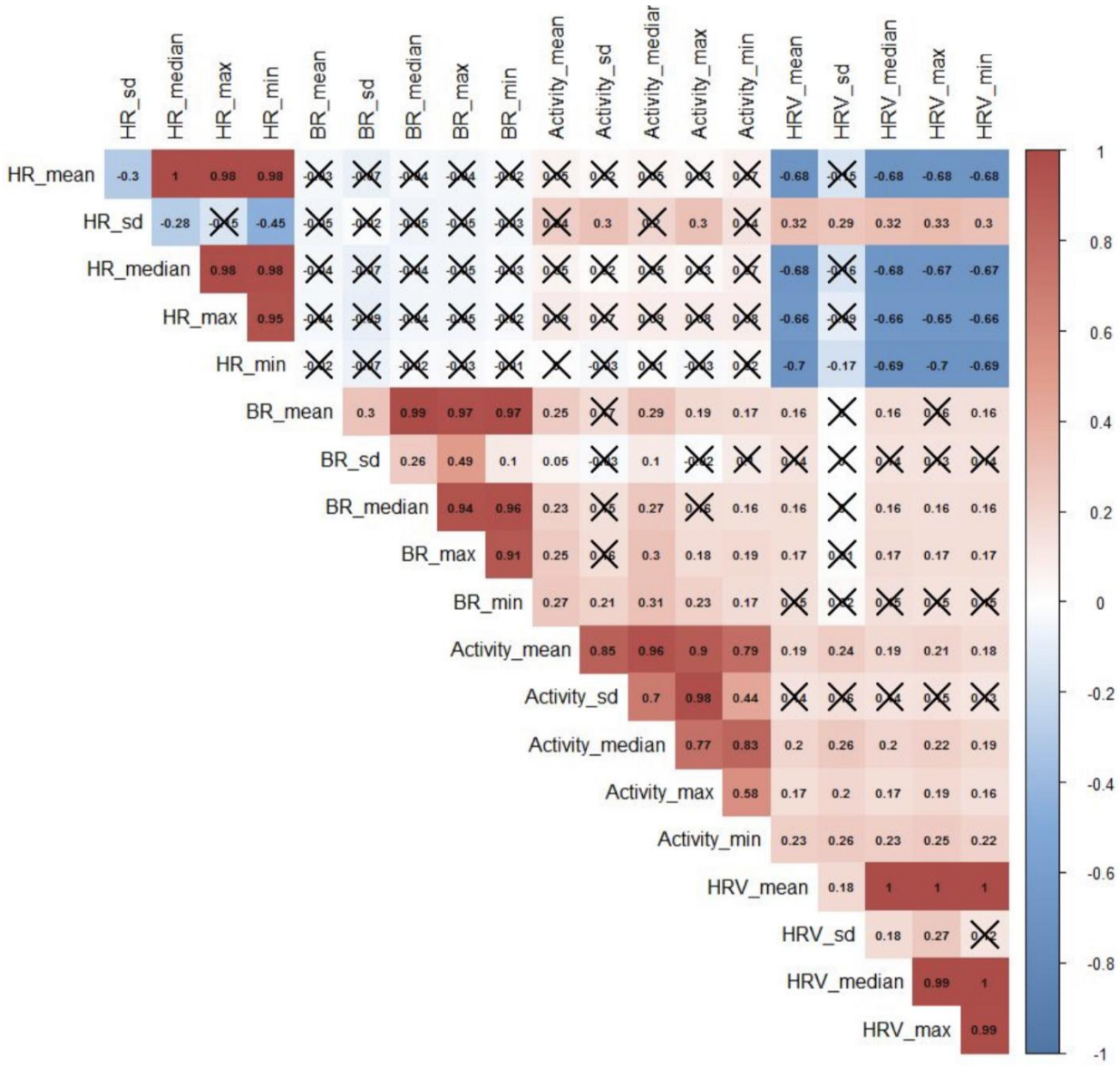
Table 2 – an example of SME's assessment of a scenario

NVC physiological data analysis

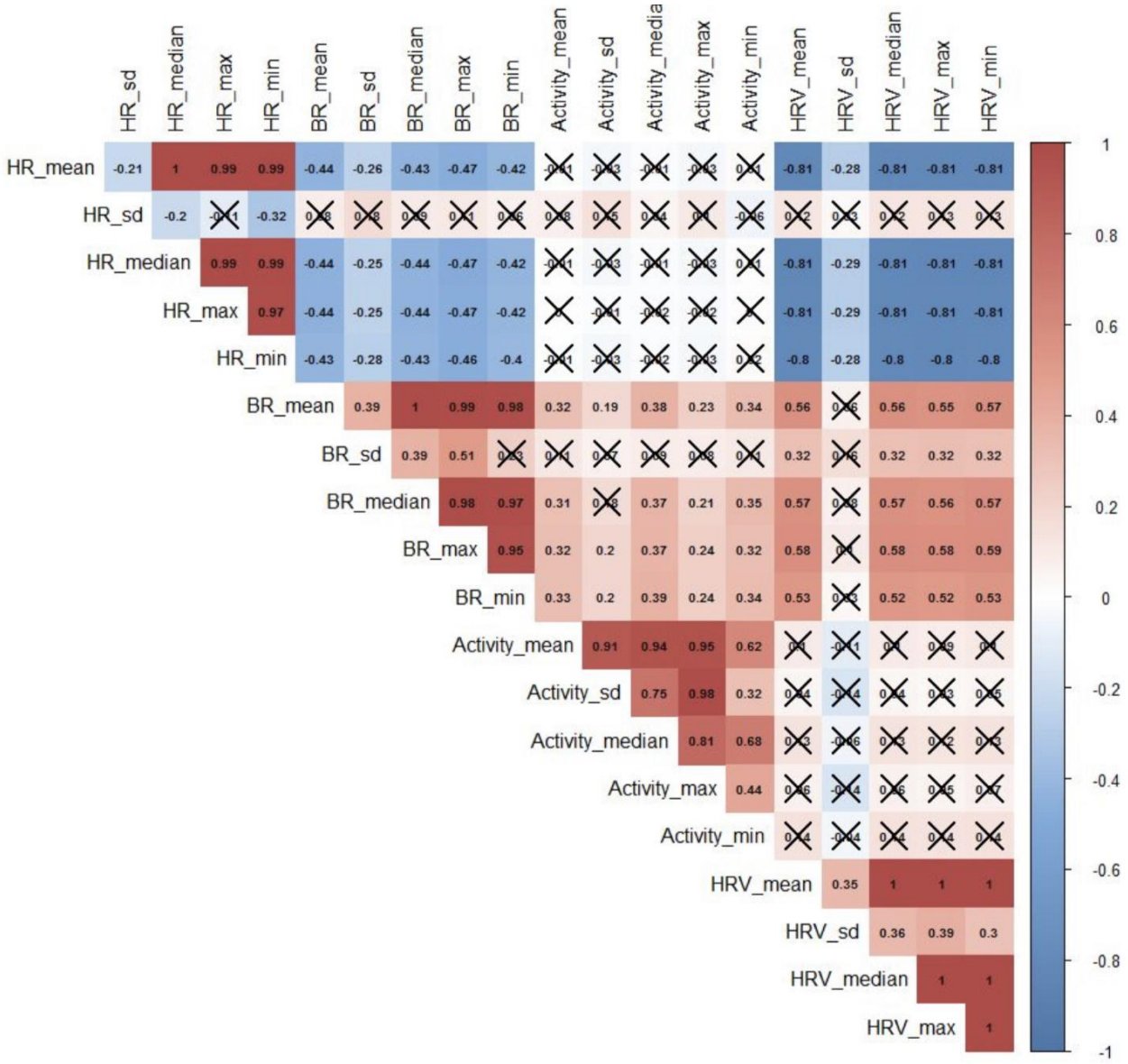
Multiple analyses were conducted in order to analyse how the dependent variables (the NVC physiological data) varied over the different independent variables (the conditions Compliant, Partially Compliant, and Non-Compliant). We described in the first deliverable how the NVC physiological data was acquired and kindly ask the reader to refer to it if needed. The analyses were conducted on the statistical software R. They were of two kinds: correlation analyses and general linear mixed models (GLMM). Correlations are used in research to show how two variables vary in regard to each other. A correlation of value 1 between the variables A and B means that when the variable A is changing in a certain way the variable B changes in exactly the same way. A correlation of value -1 means that it varies in the opposite direction. A correlation of value 0 means that there is no relationship between the two variables. Generalized linear mixed models (GLMM) are used to estimate the value of each NVC physiological data based on the different compliance conditions. GLMMs take advantage of the modelling of random effects to improve precision of the model. For example, in our case, we used the participants (electricians in training) as a random effect since each has a different way to handle the situation and different baseline. This helps measure the true difference between the compliance conditions while taking care of inter-individual differences.



Correlation between NVC physiological data during the compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)

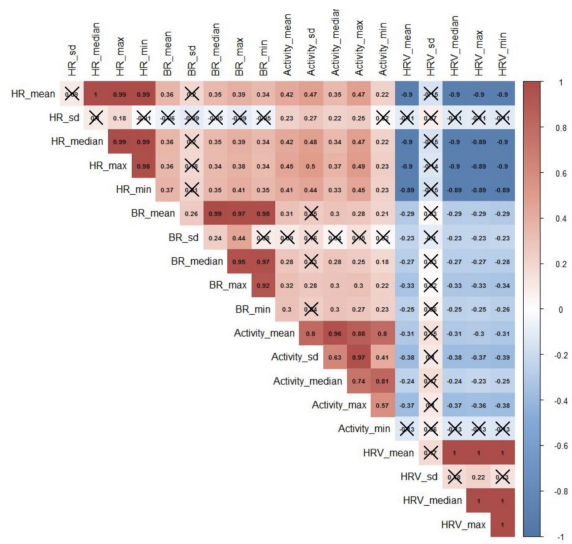
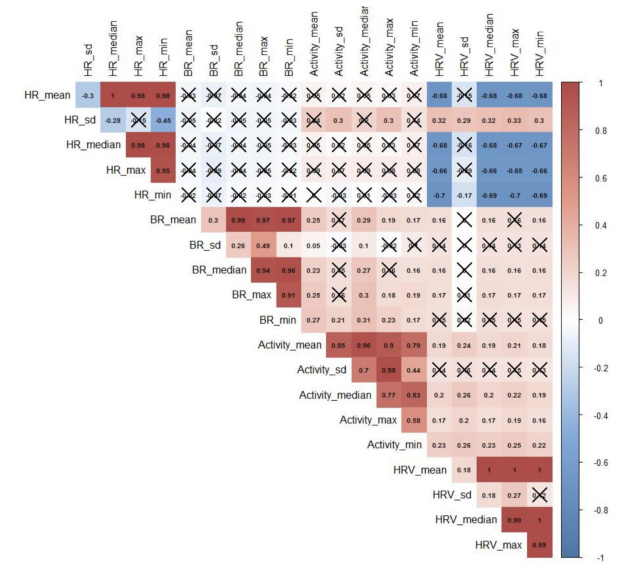
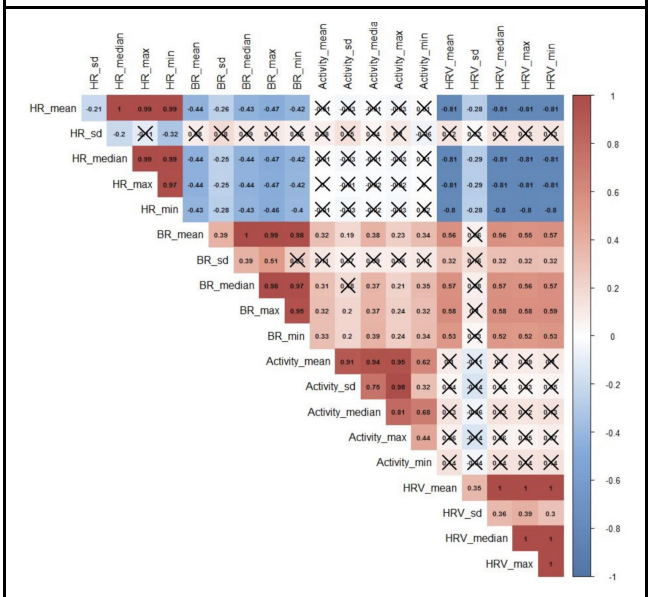


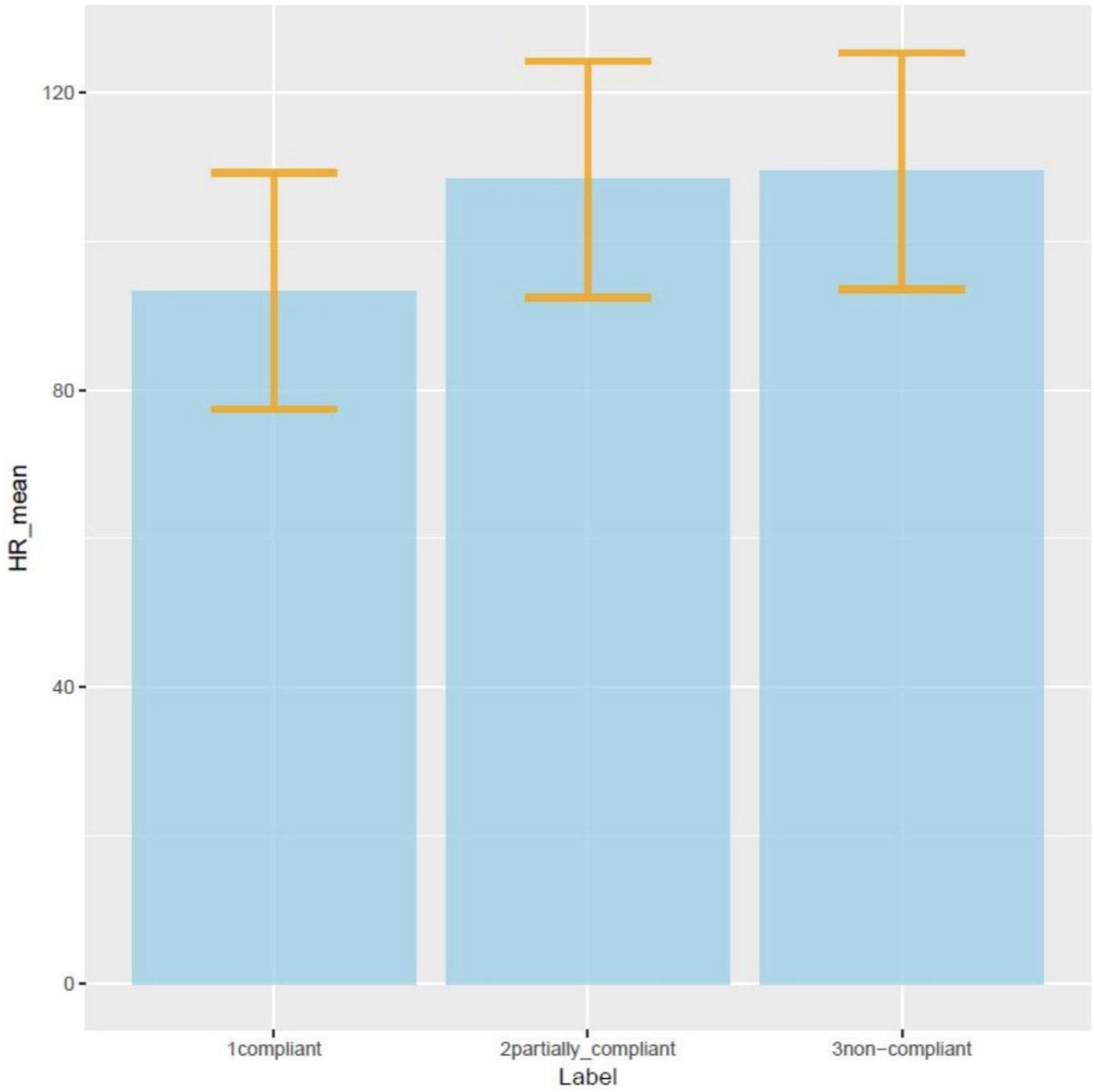
Correlation between NVC physiological data during the partially compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)



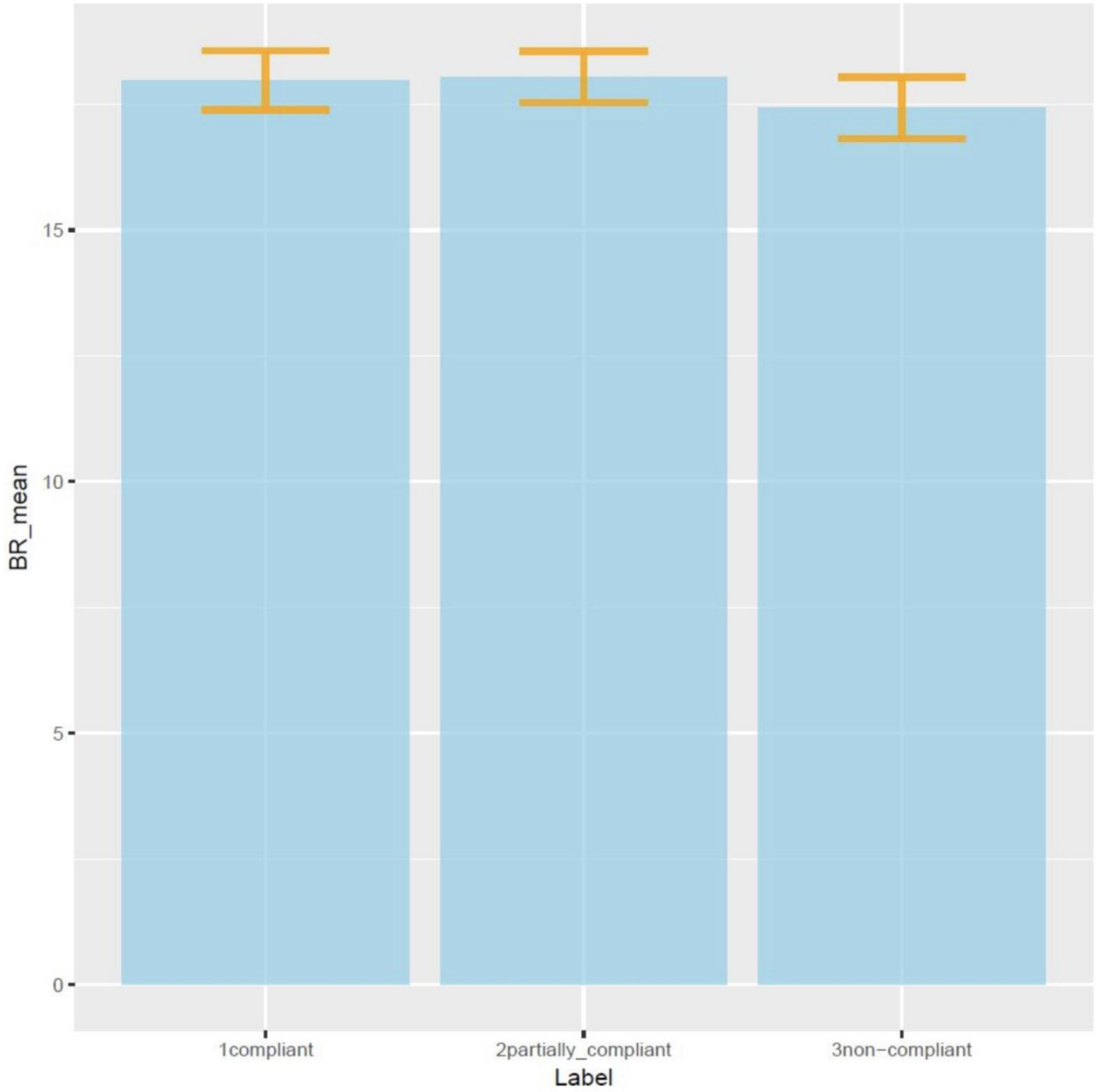
Correlation between NVC physiological data during the non compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)

In order to facilitate the comparison we display the patterns of NVC in the three conditions

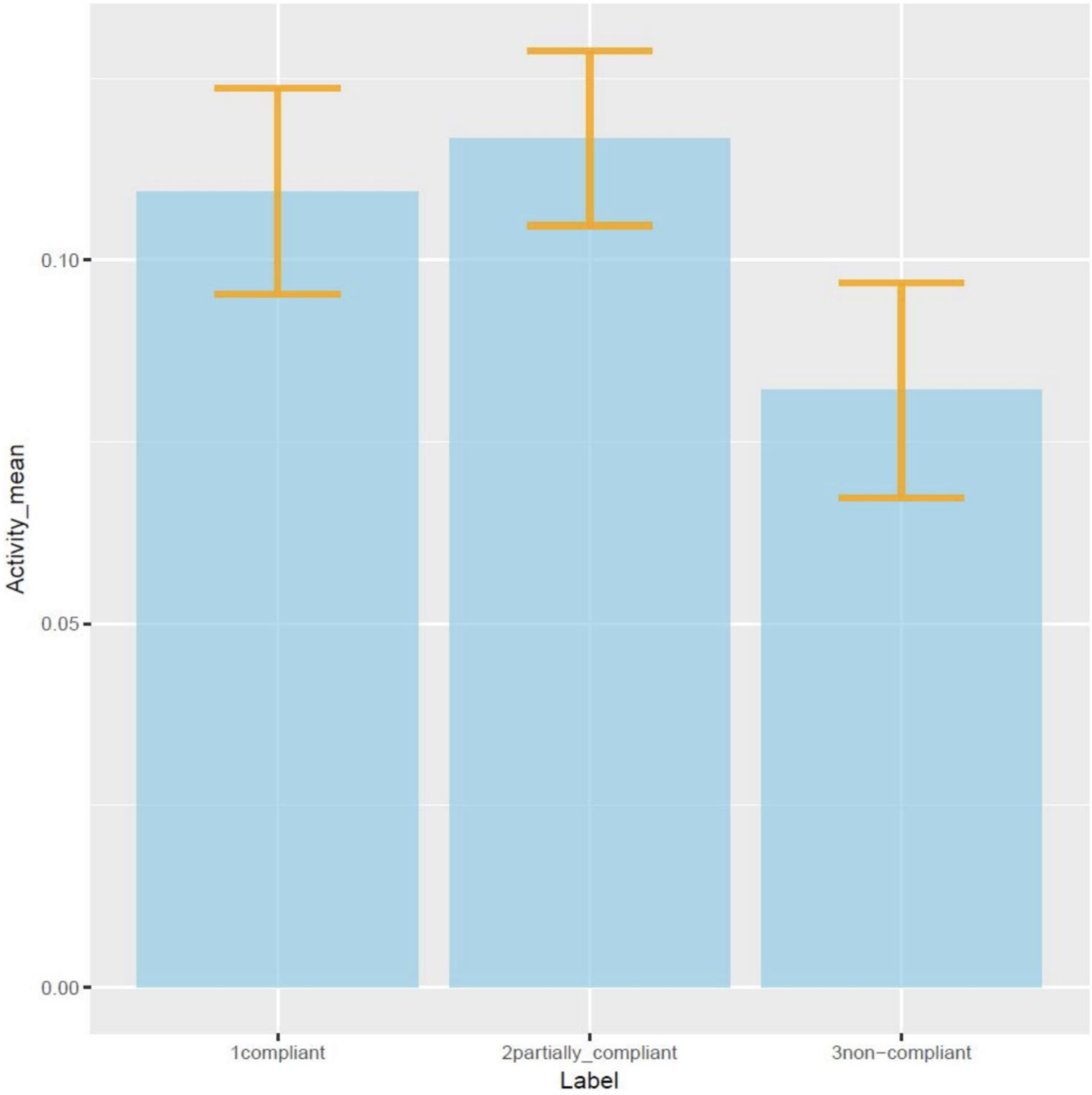
	*Correlation between NVC physiological data during the compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)*
	*Correlation between NVC physiological data during the partially compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)*
	*Correlation between NVC physiological data during the non compliant condition (HR = heart rate, BR = breathing rate, HRV = heart rate variability, sd = standard deviation)*



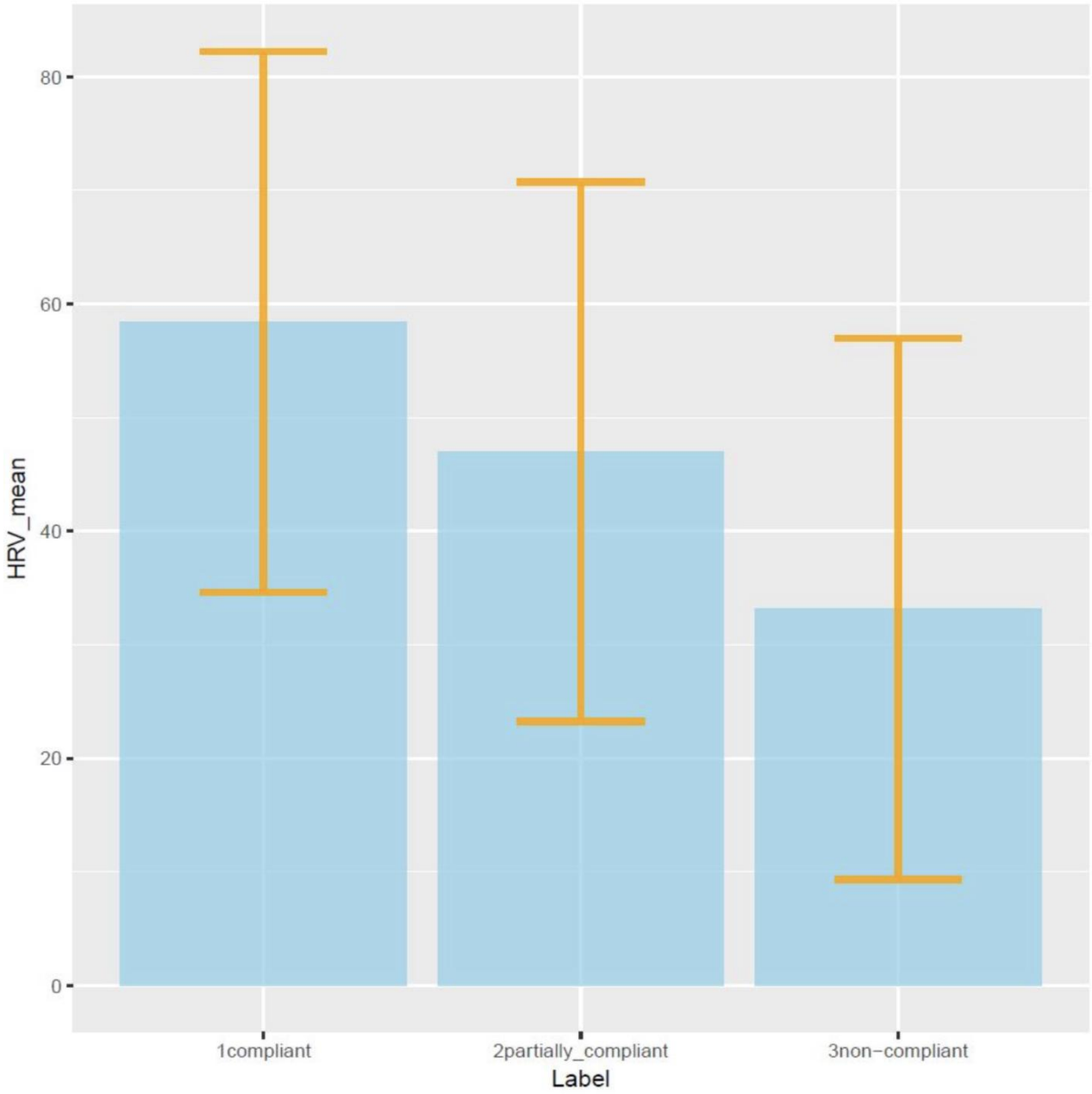
Estimate of heart rate values for all 3 conditions; compliant, partially compliant, and non compliant.



Estimate of breathing rate values for all 3 conditions; compliant, partially compliant, and non compliant.



Estimate of activity values for all 3 conditions; compliant, partially compliant, and non compliant.



Estimate of heart rate variability values for all 3 conditions; compliant, partially compliant, and non compliant.

Comments

The results are potentially very promising, since they show correlation patterns clearly different according to the compliance to procedures. In addition, the physiological data show different trends between compliant and non-compliant behaviors.
Specifically, we comment on the main results.

Breath rate and Heart rate

When the workers are compliant, BR and HR are positively correlated and are decreasing. This result is interesting, since it is a cue suggesting that the workers, in that moment, were in a particularly calm and

mindful state of mind. A decrease in BR, associated with a decrease in HR, seems to be associated with a psychological condition of mindfulness and well being. As suggested by Zaccaro et al. (2018) in their systematic review of the literature about the psychological effects of slow breathing techniques, a decrease in breath rate can "enhance interactions between autonomic, cerebral and psychological flexibility, linking parasympathetic and Central nervous System activities related to both emotional control and well-being" (Zaccaro et al., 2018, 10).
These physiological conditions are commonly considered as a marker of a mindful and relaxed state (May et al., 2016), and it is, in turn, a condition that could enhance cognitive performance (Chiesa, Calati and Serretti, 2011).
The pattern of physiological parameters changes in an interesting way, when the workers are non compliant: BR is low and is negatively correlated with HR. This kind of asynchrony may be correlated with lack of mindfulness and presence of stress. In addition, HR is high also when there is a decrease in the activity, which may be interpreted as a sign of anxiety and stress.

Breath rate and Heart rate variability

When the workers are compliant, HRV is high, BR is low, notwithstanding the increased activity. The activity index during the scenario revealed moments where the workers had to climb the stairs, operate in uncomfortable positions, carry heavy toolboxes, etc. Notwithstanding the increase in the workload, the BR was low when the workers were compliant to procedures. According to literature, high HRV has been correlated to a wide variety of psychophysiological states related to relaxation, emotional regulation, constructive coping with stressors, effective attention allocation (Appelhans & Luecken, 2006). This kind of psychological state is relevant in terms of safety, since it allows the worker to effectively manage his/her cognitive and emotional resources.
When the workers are non compliant, BR is positively correlated with HRV. As reported before, during non-compliant activities, workers' BR is low and, since it is positively correlated with HRV, we can interpret this result as a sign of stress. As stated by Lo, Wei and Hwang (2020), "the sudden decrease of HRV can be treated as a warning from the body systems [...] and is an objective method to evaluate occupational burnout".

Conclusions

The aim of this study was to develop a tool for the peer observation of specific behavioural markers for electricians as single operators and as a team during simulated scenarios. In addition, we wanted to add a new sensor-based method for observing behaviours based on non-verbal cues (NVC), like movement in the operational environment.
The basic tenets of the project were:

- training method: simulation is an effective method for competences development in many domains;
- metrics: self and peer assessment tools and a sensor-based system for monitoring team coordination during the simulation;
- key factors: learning by doing, peer-assessment, metacognition and reflection on work activities, visual and concrete representation of teamwork non-verbal dynamics;
- New mechanism for learning: simulation and debriefing, sensor-based system for monitoring team coordination;

- Integration with top-down safety management: simulation is a very effective tool for behaviour change for safety's sake, the training of new procedures, the discussion of "hidden practices", the clarification of attitudes and beliefs.

The stakeholders interested in these results will be practitioners working in electricity distribution belonging to other local units of e-distribuzione or other distribution companies. However, once validated, the method could be easily disseminated to other technical domains (e.g., High Voltage workers, Power plant workers, etc.), or even other work domains (e.g., team of fire-fighters). For the first time in the electric industry, it is possible to explicitly analyse SPS by means of a structured form, based on the specific activities, and, most of all, NVC can be tracked and used for addressing often vague and ill-defined concepts like teamwork and coordination. Moreover, the analysis of NVC will reveal specific patterns of behaviours that are correlated with safe performance and, therefore, could be explicitly addressed during the professional training.

Future developments of this research could aim at developing a new generation of sensor-based systems for monitoring team coordination in both routine and extreme situations, namely a context simultaneously marked by high levels of uncertainty, change and risk. The analysis of social signals and face-to-face communication patterns (e.g., kinesics, proxemics, interpersonal synchronization), could be combined with other sources of information such as survey and performance metrics (e.g., feedback on the SPS, self and peer assessment of performance efficiency, etc.). We argue that these systems could help teams to design interventions aimed at enhancing individual and group performance, especially for coordinating efficiently and ensure their resilient capacity to face risk and overcome perturbations, a critical aspect of extreme situations. By automatically quantifying human behaviour using wearable and non-invasive sensors, we could find relationships between sensor data and team performance and thus identify optimal behaviour patterns that would lead to improved performance.

References

Appelhans, B. M., & Luecken, L. J. (2006). Heart Rate Variability as an Index of Regulated Emotional Responding. Review of General Psychology, 10(3), 229–240. https://doi.org/10.1037/1089-2680.10.3.229

Chiesa, A., Calati, R., and Serretti, A. (2011). Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. Clin. Psychol. Rev. 31, 449–464. doi: 10.1016/j.cpr.2010.11.003

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. Journal of ASTD, 11, 1–13.

Lo EV, Wei YH, Hwang BF. (2020). Association between occupational burnout and heart rate variability: A pilot study in a high-tech company in Taiwan. Medicine (Baltimore). Jan;99(2):e18630. doi: 10.1097/MD.0000000000018630.

May, Ross & Bamber, Mandy & Seibert, Gregory & Sanchez-Gonzalez, & Leonard, & Salsbury, & Fincham, Frank. (2016). Understanding the Physiology of Mindfulness: Aortic Hemodynamics and Heart Rate Variability. Stress: The International Journal on the Biology of Stress .. 19. 10.3109/10253890.2016.1146669.

de Winter, J. C. F., Dodou, D., Mulder, M. (2012) Training Effectiveness of Whole Body Flight Simulator Motion: A Comprehensive Meta-Analysis, The International Journal of Aviation Psychology, 22:2, 164-183, DOI: 10.1080/10508414.2012.663247

Zaccaro, A., Piarulli, A., Laurino, M., Garbella, E., Menicucci, D., Neri, B., & Gemignani, A. (2018). How Breath-Control Can Change Your Life: A Systematic Review on Psycho-Physiological Correlates of Slow Breathing. Frontiers in human neuroscience, 12, 353. https://doi.org/10.3389/fnhum.2018.00353