Emerging Technologies in Healthcare and Medicine, Vol. 116, 2023, 221–230
https://doi.org/10.54941/ahfe1004375

Unveiling Readiness of Medical First Responders in Simulation Trainings: Insights Beyond Queries

Annelien Malfait[1,2], Martine Van Puyvelde[1,3,4], Frédéric Detaille[1], François Waroquier[5], Xavier Neyt[1], and Nathalie Pattyn[1,6]

[1]VIPER Research Unit, LIFE Department, Royal Military Academy, Brussels, Belgium
[2]University Hospital Brussels, Brussels, Belgium
[3]Brain, Body and Cognition, Department of Psychology, Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Brussels, Belgium
[4]Faculty of Science, School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, United Kingdom
[5]Medical Component Competence Center, Belgian Defence, Brussels, Belgium
[6]Center for Advanced Research in Sleep Medicine, Hôpital du Sacré-Coeur de Montréal, CIUSSS NÎM & Université de Montréal, Montréal, QC, Canada

ABSTRACT

First responders working in emergency medicine (medics, nurses and physicians), both within emergency departments and in extra-hospital care, face complex operational situations and therefore undergo extensive trainings. Besides enhancing medical technical and non-technical skills, these trainings also need to encompass the recognition and the management of emotions and stress. We designed a methodology to synchronously quantify both the technical performance of medical skills and the psychophysiological load experienced by the participant, the latter via facial expression analysis and voice stress analysis. The integration and synchronisation of these analyses will offer a unique additional insight in health care providers' performance during simulation trainings.

Keywords: Emergency medicine, Simulation training, Medical technical skills, Psychophysiological load, Facial expression analysis, Voice stress analysis

INTRODUCTION

Whatever distinctions may be made between military and civilian emergency medicine in terms of both morbidity, mortality and environment, we agree that medical professionals in both fields must be able to operate in stressful, difficult and sometimes extreme circumstances (e.g., Van Puyvelde et al., 2023). Where risk to personal life for the first responders used to be a feature of military environments, it is a growing area of concern for civilian interventions, both in exceptional circumstances like terror attacks (e.g., Amiresmaili et al., 2022) or in the more frequently occurring urban violence (e.g., Verwee et al., 2022).

Simulation trainings are particularly valuable to train medical proficiency in the management of these complex, time-critical and life-threatening emergencies (e.g., von der Heyden and Meissner, 2015). Clinical knowledge and skills alone are not sufficient to provide optimal medical care in this context. In addition to their technical expertise, medical professionals need to master non-technical skills (e.g., Greif et al., 2021). Recognizing the critical role and negative impact stress can have on performance (e.g., Kent et al., 2020) and on personal level (e.g., Van Puyvelde et al., 2023), they also need be able to acknowledge and manage their stress responses, both individually and at team level (e.g., Brazil et al., 2023). Therefore, besides focussing on technical and non-technical skills (e.g., L'Her et al., 2020), simulation trainings should also encompass the recognition and regulation of stress responses (Brazil et al., 2023).

Previously applied methods to map stress and emotions during simulation trainings often rely on subjective measurements as questionnaires. More objective physiological measurements, e.g., measuring salivary and serum cortisol levels, may have an invasive character and are difficult to obtain in real-time, besides having a time resolution that does not allow a fine-grained analysis of performance. Other measurements, e.g., heart rate variability, can be obtained in real time, however they require wearable sensors and can be confounded due to physical activity, medication... (e.g., Kent et al., 2020).

In order to fill this gap, we designed a methodology to quantify the technical performance of the medical skills as well as the psychophysiological load during simulation trainings. Both synchronously, independently of each other and without the need to wear any sensors or perform invasive sampling.

METHODS

The methodology we developed comprises a realistic simulation training with two objectives. On the one hand, quantifying the technical performance by analysing the outcome and the details of the performance of the medical skill. On the other hand, the analysis of facial expressions as well as voice stress analysis to quantify the allostatic load.

The simulation lab is equipped with eight high-quality cameras, four Axis M5525-E pan-tilt-zoom cameras attached on the ceiling corners and four AXIS M5075 pan-tilt-zoom cameras attached on the four walls of the room at eye-level. All cameras are synchronized but can be controlled separately during recording in terms of orientation and zooming. Video recordings obtained with this equipment have a resolution from 1920x1080 to 320x180 (16:9) pixels. The model AXIS M5525-E obtains a frame rate up to 20/30 frames per second and the model AXIS M5075 obtains a frame rate up to 50/60 frames per second.

Audio was synchronously recorded at a sampling frequency of 16KHz by two omnidirectional microphones, Sennheiser MEB102, attached on the ceiling.

The complete technical equipment is integrated through the Viso software (Noldus Information Technology BV, NL).

Via 'The Observer XT' (Version 16, Noldus Information Technology BV, NL) not only live observation but also retrospective microanalysis is available.

RESULTS

TECHNICAL SKILL ANALYSIS

We exposed our participants to a clinical relevant case/scenario in which they had to perform several medical technical skills. We observed and recorded their performance. Following the 'European Resuscitation Guidelines 2021' with regard to teaching and learning (Greif et al., 2021), formative assessment and corrective feedback were given by technical observers during a debriefing immediately after the simulation training. This macro-analysis was guided by detailed score-sheets for each core skill, based on the guidelines of 'Deployed Medicine' (Joint Trauma System, 2022).

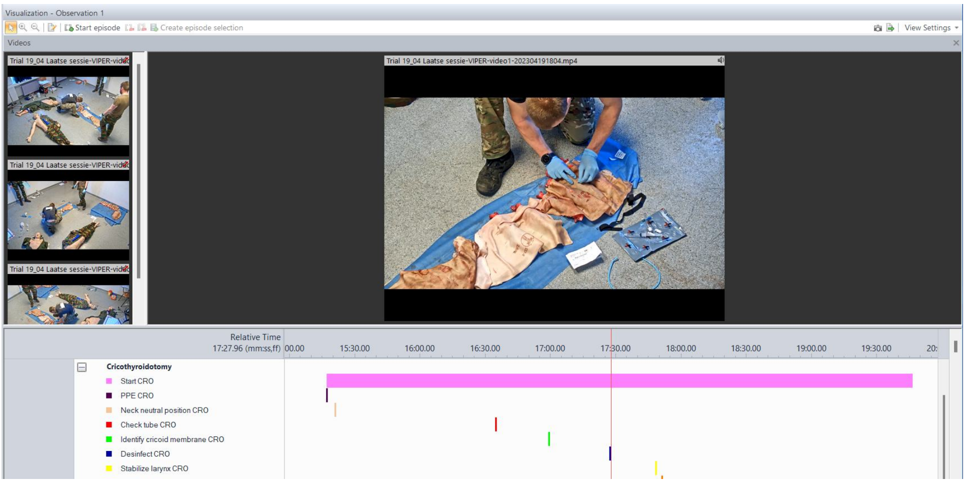

Figure 1: Detailed technical analysis of medical skills with 'The observer XT' (Version 16). The different required components of the skill (here cricothyroidotomy) are manually scored by an expert. Footage of different cameras make it possible to observe the performance from different angles.

Afterwards, a more comprehensive technical analysis occurred. A senior medical expert reviewed and scored the recorded performance of the medical skills with the help of 'The Observer XT' (Version 16). Not only the outcome (e.g. application of tourniquet and subsequent stop of the bleeding) was scored, but also a detailed micro-analysis of the skill was executed based on in-depth coding of interventions and skill performance. (See Figure 1). Each to be evaluated skill was broken down to several components, some critical and some not-critical (Joint Trauma System, 2022). If all critical components of a skill were performed correctly and in a timely manner, a skill could be considered as succeeded. If the execution of a technical skill was wrong or any critical component was missing, incorrectly or not in time executed, the performance of the skill was considered failed.

Since video recording and subsequent analysis in 'The Observer XT' (Version 16) allows to pause and play back videos unimpededly at recorded speed, as well as at half-speed, at one fifth- and one twenty-fifth of the speed, measurement of time intervals and more in-depth analysis was facilitated. Because of the presence of more cameras, experts could view and review performances from different angles. Furthermore the possibility of adapting orientation and zooming during recording provided more detailed views.

PSYCHOPHYSIOLOGICAL ANALYSIS

In addition to assessing technical skills, we wanted to map the psychophysiological load and resource allocation during the simulation training, with a time resolution that allowed to investigate changes throughout the scenario. Two methods, relying on the systemic measures of arousal and emotion were used: recordings of facial expressions on the one hand, and voice analysis on the other hand.

Our choice for these methods was guided, from a practical point of view, by the non-invasiveness: no sensor needs to be worn to capture the information. Regarding a more conceptual point of view, combining both analyses allows for a quantification of arousal and valence, according to the circumplex model of emotions (Russel, 1980), and thus goes beyond a linear quantification of "stress".

Facial Expression Analysis

Theoretical Background

Based on his evolutionary theory Darwin opened the topic of universal facial expressions of emotions in 1872 (Darwin, 1872). Almost a century later Ekman described six primary emotions (anger, disgust, fear, happiness, sadness, and surprise) and their corresponding universal facial expressions (Ekman, 1970). Contempt was added later (Ekman and Friesen, 1986).

In order to objectify and quantify facial expressions, Ekman and Friesen developed the 'Facial Action Coding System' (FACS) in 1978. The FACS breaks down the facial expression into different units of muscular activity responsible for the smallest visually discriminable and observable facial movements, called Action Units (AUs) (Ekman and Friesen, 1978). This version was reviewed and updated in 2002 (Ekman et al., 2002). FACS is the most comprehensive, psychometrically rigorous and widely used system to describe facial activity with AUs. It is used in domains varying from behavioural science and psychology to computer graphics and animation (Cohn et al., 2007).

FaceReader by Noldus-Viso

Manual FACS coding is very time-consuming, requires extensive training and remains prone to observer bias, therefore automated FACS coding was developed (e.g. Hamm et al., 2011). FaceReader (Noldus Information Technology BV, NL) is a software designed for automated facial coding. It detects and locates faces, using a deep learning based face-finding algorithm (Zafeiriou et al., 2015). After detection of the face, a facial modeling technique based on neural networks (Bulat and Tzimiropoulos, 2017) is used to synthesize an artificial face model, describing the location of 468 key points in the face. Next, classification of the facial expressions is done by a trained neural network to recognize patterns in the face (Gudi et al., 2015). The neural network was trained by over 20 000 manually annotated images to classify the universal emotions: happiness, sadness, anger, surprise, fear, disgust, contempt, a neutral state and the 20 most common AUs in the face (See Table 1) (Loijens et al., 2023).

Table 1. Action units that can be analysed by the action unit module in FaceReader (version 9).

Action Unit	FACS Name	Action Unit	FACS Name	Action Unit	FACS Name
1	Inner Brow Raiser	10	Upper Lid Raiser	23	Lip Tightener
2	Outer Brow Raiser	12	Lip Corner Puller	24	Lip Pressor
4	Brow Lowerer	14	Dimpler	25	Lips Part
5	Upper Lid Raiser	15	Lip Corner Depressor	26	Jaw Drop
6	Cheek Raiser	17	Chin Raiser	27	Mouth Stretch
7	Lid Tightener	18	Lip Pucker	43	Eyes Closed
9	Nose Wrinkler	20	Lip Stretcher		

Expressions, manually scored by certified annotators using the Amsterdam Dynamic Facial Expression Set (ADFES) (van der Schalk et al., 2011) and the Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) (Olszanowski et al., 2014), were used to validate the facial expression analysis of FaceReader (Version 9). FaceReader (Version 9) had an accuracy of respectively 99.3% and 95.7%. for facial expression classification on ADFES and WSEFEP. For AU classification on ADFES and WSEFEP FaceReader (Version 9) had an F1score[1] of respectively 0.775 and 0.762. Using an internal meta test set (Noldus Information Technology BV, NL) including more realistic/demanding expressions, the F1 score obtained for AU classification was 0.57 (Loijens et al., 2023). For our analysis we used FaceReader (Version 9.1.8). FaceReader analyses images with a maximum resolution of 1280 × 1280. If the resolution is higher then it is downscaled maintaining the aspect ratio.

As a standard setting for video analyses, videos are processed frame-by-frame. For camera analyses, analysing a new frame starts when the analysis previous frame is terminated (reaching 15 frames per second). If this is lower than the camera frame rate, FaceReader skips a frame to keep up with the camera frame rate but once the analysis is finished, the results of the bypassed frames are interpolated to the original frame rate of the recorded video (Loijens et al., 2023).

Normally, accurate analysis can be obtained by analysing 10–15 frames per second (Loijens et al., 2023).

Since our slowest parameter is the frame rate of 20/30 frames per second by the camera model AXIS M5525-E, accurate analyses could be acquired even in real-time.

[1]F1 score is the weighted average of precision and recall (2 x [(Precision x Recall)/(Precision + Recall)]

Variables

FaceReader generates the classification of facial expressions of the participant. Each expression receives a value between 0 (not visible) and 1 (fully present) on a continuous scale, indicating its intensity (again, based on training by human experts). Since different facial expressions can occur at the same time with varying intensities, the sum of the intensity values for the expressions at one specific moment is generally not equal to 1.

20 Action Units can also be classified (See Table 1). They are rated on a 5-point intensity scale (A = trace, B = slight, C = marked or pronounced, D = severe or extreme, E = maximum) defined by Ekman at al. (2002). Presence of a combinations of these AUs can be associated with emotional states as boredom, confusion...

Arousal can be calculated based on the activation values of AU 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 18, 20, 23, 24, 25, 26, 27 and the inverse of AU 43 (since AU 43 stands for eyes closed which indicates low arousal). The average AU activation values (AAV) are computed over the last 60 seconds (During the initial 60 s of the analysis, the AAV is calculated based on the analysis conducted up to that point). To correct for AUs that are constantly activated, the AAV are subtracted from the current AU activation values (AV), giving the Corrected Activation Values (CAV). Finally the arousal is calculated by taking the mean of the five highest values of the CAV values (Loijens et al., 2023).

Voice Analysis

Theoretical Background

Hans Seyle described stress as the non-specific response of the body to any demand made upon it (Selye, 1974). The stress response is a result of the interplay between the environment and the individual's resources to meet the demands (Lazarus and Folkman, 1984). Emotion and stress regulation are dependent from the balance between bottom-up arousal (responses induced by emotional stimulus) and top-down regulatory activity (responses after cognitive evaluation) (e.g., Flores-Kanter et al., 2021). Since voice production demands the complex coordination of breathing, phonation, and resonance (e.g., Morrison et al., 1994) involving both the central and peripheral nervous system (autonomous nervous system) (Câmara et al., 2015), the relation between voice output and psychophysiological processes of stress regulation is increasingly examined (e.g., Van Puyvelde et al., 2018). The 'Model for Voice and Effort' (MoVE) described by Van Puyvelde et al. (2018) illustrates the interplay between ongoing top-down and bottom-up regulation and their reflection on the phonation voice parameters F0, F0-range, and jitter (Van Puyvelde et al., 2018).

Technical Aspects

Audio, synchronously recorded with the videos, (cfr. supra) was analysed through Matlab and PRAAT (Boersma and Weenink), as applied in previous studies about voice reactivity in response to hypoxia (e.g. Van Puyvelde et al., 2020). Three phonation parameters were examined for our methodology: F0, F0-range and jitter.

The fundamental frequency (F0) of a speech signal is defined as the average number of oscillations of the vocal folds per second, expressed in Hertz. F0-range refers to the range of the measured fundamental frequencies.

The vocal fold oscillation is not exactly periodic but contains fluctuations, the amount of variation in period length is defined as jitter (Bäckström et al., 2022).

Variables and Their Correlation With Stress

As illustrated in Figure 2, the 'Model for Voice and Effort' shows that increased F0-ranges are linked with a decrease of cognitive top-down regulation. The model uses the example of sleep deprivation, where cognitive control is decreased, which causes increased F0-ranges. This is reaching an alarm zone when the cognitive top-down control is fully lost e.g., in life-threatening emergency situations such as an impeding flight crash. Decreased F0-ranges on the other hand correspond with a high cognitive load and top-down control.

Extremes in F0-values, i.e., strongly increased or decreased F0-values, suggest effort-depletion, also reaching an alarm zone in life-threatening emergency situations.

Jitter correlates inversely with bottom-up arousal (Van Puyvelde et al. 2018). However, the impact of real-life emergency situations is much bigger on the voice output than these induced in a laboratory (e.g., Van Puyvelde et al., 2018).

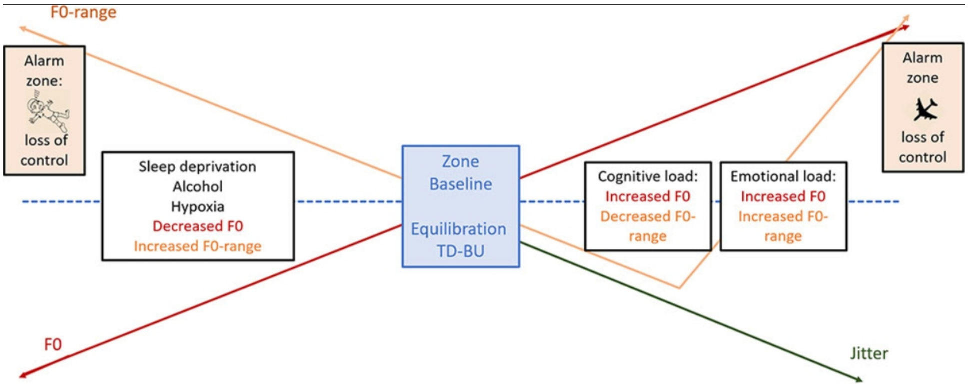

Figure 2: "Model for voice and effort" (MoVE) (Van Puyvelde et al. 2018).

DISCUSSION

The power and novelty of the method we developed, lies in the unique integration and synchronisation of a technical skill analysis and an analysis of the psychophysiological load at the same time, perfectly synchronized yet independent of each other.

The major added value resides in the non-invasive and objective character of the facial expression analyses and voice analyses, hence mapping the psychophysiological load without relying on self-reports and not hampering the actual execution of the simulation in any way.

This method allows to improve the understanding of how stress and emotions can affect technical performance. Given the objective character of the analyses, they can easily be repeated on future training sessions and thereby investigating the evolution of the technical and psychophysiological performance of the participants over time and in different circumstances, as well as investigating strategies to mitigate negative impacts of stress on performance.

CONCLUSION

The evaluation methodology designed for this project enables a detailed technical skill analysis by coupling the macro-outcome to micro-recordings of performance, together with mapping the psychophysiological load by facial expression analysis and voice stress analysis.

This combination will offer a unique insight in health care providers' performance during simulation trainings (or, if implemented in operational equipment, during operational engagement).

ACKNOWLEDGMENT

We thank Col Julie Delwick, Commanding Officer of the Medical Component Competence Centre of the Belgian Defense, for her continuous support to this project. Funding from the Royal Higher Institute for Defense of the Belgian Defense under grant HFM/19-08 is acknowledged. Finally, we gratefully acknowledge the willingness of all our participants: special operations medics and physicians in training.

REFERENCES

Amiresmaili, M., Talebian, A. and Miraki, S. (2020) 'Pre-hospital emergency response to terrorist attacks: A scoping review', *Hong Kong Journal of Emergency Medicine*, 29(1), pp. 56–62. doi: 10.1177/1024907920941620.

Bäckström, T. *et al.* (2022) *Introduction to speech processing*, *2nd Edition*. Available at: https:// speechprocessingbook.aalto.fi (Accessed: 01 August 2023).

Boersma, P. and Weenink, D. (no date) *Praat: Doing Phonetics by computer*. Available at: https://www.fon.hum.uva.nl/praat/ (Accessed: 01 August 2023).

Brazil, V. et al. (2023) 'Exploring participant experience to optimize the design and delivery of stress exposure simulations in emergency medicine', *AEM Education and Training*, 7(2). doi: 10.1002/aet2.10852.

Bulat, A. and Tzimiropoulos, G. (2017) 'How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230, 000 3D facial landmarks)', *2017 IEEE International Conference on Computer Vision (ICCV)* Venice, Italy, 2017, pp. 1021–1030. doi: 10.1109/iccv.2017.116.

Câmara, R. and Griessenauer, C. J. (2015) 'Anatomy of the vagus nerve', in R. S. Tubbs et al. (eds.) *Nerves and nerve injuries*. Amsterdam: Academic Press, pp. 385–397.

Cohn, J. F., Ambadar, Z. and Ekman, P. (2007) 'Observer-based measurement of facial expression with the facial action coding system', in J. A. Coan and J. J. B. Allen (eds.) *Handbook of emotion elicitation and assessment*. Oxford, New York: Oxford University Press, pp. 203–221.

Darwin, C. (1872) *The expression of the emotions in man and animals*. London, John Murray.

Ekman, P. (1970). Universal Facial Expressions of Emotions. *California Mental Health Research Digest*, 8(4), pp. 151–158.

Ekman, P. and Friesen, W. V. (1978). *Facial action coding system: A Technique for the measurement of facial movement*. Palo Alto. Consulting Psychologists Press.

Ekman, P. and Friesen, W. V. (1986) 'A new pan-cultural facial expression of emotion', *Motivation and Emotion*, 10(2), pp. 159–168. doi: 10.1007/bf00992253.

Ekman, P., Friesen, W. V., and Hager, J. C. (2002). *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. San Francisco, CA: Consulting Psychologists Press.

Flores-Kanter, P. E., Moretti, L. and Medrano, L. A. (2021) 'A narrative review of emotion regulation process in stress and recovery phases', *Heliyon*, 7(6). doi: 10.1016/j.heliyon.2021.e07218.

Greif, R. *et al.* (2021) 'European Resuscitation Council guidelines 2021: Education for Resuscitation', *Resuscitation*, 161, pp. 388–407. doi: 10.1016/j.resuscitation.2021.02.016.

Gudi, A. *et al.* (2015) 'Deep Learning based FACS action unit occurrence and intensity estimation', *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* 6: 1–5. doi: 10.1109/fg.2015.7284873.

Hamm, J. *et al.* (2011) 'Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders', *Journal of Neuroscience Methods*, 200(2), pp. 237–256. doi: 10.1016/j.jneumeth.2011.06.023.

Joint Trauma System (2022) *Combat Medic/Corpsman Skill Assessment Checklists*, *Deployed Medicine*. Available at: https://www.deployedmedicine.com/market/299/category/319 (Accessed: 22 July 2022).

Kent, J. *et al.* (2019) 'Acute provider stress in high stakes medical care: Implications for trauma surgeons', *Journal of Trauma and Acute Care Surgery*, 88(3), pp. 440–445. doi: 10.1097/ta.0000000000002565.

Lazarus, R. S. and Folkman, S. (1984) *Stress, appraisal, and coping*. New York, New York : Springer publishing company.

L'Her, E. *et al.* (2020) 'Simulation-based teaching in critical care, anaesthesia and emergency medicine', *Anaesthesia Critical Care & Pain Medicine*, 39(2), pp. 311–326. doi: 10.1016/j.accpm.2020.03.010.

Loijens, L. *et al.* (2023) *Reference Manual – FaceReader Version 9*. Wageningen, The Netherlands: Noldus Information Technology bv.

Morrison, M., *et al.* (1994) Anatomy and physiology of voice production. In: *The Management of Voice Disorders*. Springer, Boston, MA.

Olszanowski, M. *et al.* (2015) 'Warsaw set of emotional facial expression pictures: A validation study of facial display photographs', *Frontiers in Psychology*, 5. doi: 10.3389/fpsyg.2014.01516.

Russell, J. A. (1980) 'A circumplex model of affect.', *Journal of Personality and Social Psychology*, 39(6), pp. 1161–1178. doi: 10.1037/h0077714.

Selye, H. (1974). *Stress Without Distress*. Philadelphia, PA: Lippincott.

van der Schalk, J. *et al.* (2011) 'Moving faces, looking places: Validation of the Amsterdam dynamic facial expression set (ADFES).', *Emotion*, 11(4), pp. 907–920. doi: 10.1037/a0023853.

Van Puyvelde, M. *et al.* (2018) 'Voice stress analysis: A new framework for voice and effort in human performance', *Frontiers in Psychology*, 9. doi: 10.3389/fpsyg.2018.01994.

Van Puyvelde, M. *et al.* (2020) ‘Voice reactivity as a response to acute hypobaric hypoxia at high altitude’, *Aerospace Medicine and Human Performance*, 91(6), pp. 471–478. doi: 10.3357/amhp.5390.2020.

Van Puyvelde, M. *et al.* (2023) ‘Walk the line: A systemic perspective on stress experienced by emergency medical personnel by comparing military and civilian prehospital settings’, *Frontiers in Public Health*, 11. doi: 10.3389/fpubh.2023.1136090.

Verwee, I., Dormaels, A. and Vermeulen, C. (2022). Geweld en agressie tegen publieke dienstverleners. Geweld tegen ambulanciers, brandweerlui en werknemers op spoedgevallendiensten in kaart gebracht. Brussel: Vias Institute.

von der Heyden, M. and Meissner, K. (2015) ‘Simulation in preclinical emergency medicine’, *Best Practice & Research Clinical Anaesthesiology*, 29(1), pp. 61–68. doi: 10.1016/j.bpa.2015.01.001.

Zafeiriou, S., Zhang, C. and Zhang, Z. (2015) ‘A survey on face detection in the wild: Past, present and future’, *Computer Vision and Image Understanding*, 138, pp. 1–24. doi: 10.1016/j.cviu.2015.03.015.