Some Major Human Issues in Aerospace Engineering: Review and Extension

Ephraim Suhir[1,2,3,4](✉)

[1] Portland State University, Portland, OR, USA
`suhire@aol.com`
[2] James Cook University, Townswille, QLD, Australia
[3] Technical University, Vienna, Austria
[4] ERS Co., Los Altos, CA, USA

Abstract. By employing quantifiable and measurable ways to assess the role of various uncertainties associated with the mental workload (MWL) and human capacity factor (HCF), and treating a human-in-the-loop (HITL) as a part, often the most critical part, of the complex man-instrumentation-equipment-vehicle-environment system, one can improve dramatically the human's performance, achieve the best human-system-integration (HSI) possible and predict, minimize and, when appropriate, even specify the probability of the occurrence of a casualty. The ultimate objective of the analysis is to develop effective predictive modelling techniques that would enable quantifying, on the probabilistic basis, the role of the human factor (HF) and improve his/hers HCF, so that he/she would be able to successfully cope, when necessary, with an elevated MWL and minimize the probability of an accident, when reliability of instrumentation and human performance contribute jointly to the outcome of an HSI related mission or an off-normal situation. It is concluded that the suggested MWL/HCF models and their possible modifications and generalizations can be helpful when developing guidelines for personnel selection and training; and/or when there is a need to decide, if the existing methods of reliability and ergonomics engineering are adequate in various off-normal situations, and if not, whether additional and/or more advanced and, perhaps, more expensive equipment or instrumentation should be developed, tested and installed to meet the safety requirements.

Keywords: Human-in-the-loop · Human-system-integration · Mission outcome · Mental workload · Human capacity factor

Acronyms

AG = Automated Driving
BAZ = Boltzmann-Arrhenius-Zhurkov's (model)
DD = Driver Drowsiness
DEPDF = Double Exponential Probability Distribution Function
EVD = Extreme Value Distribution
FOAT = Failure Oriented Accelerated Testing
HALT = Highly Accelerated Life Testing

N. L. Black et al. (Eds.): IEA 2021, LNNS 221, pp. 169–177, 2021.
https://doi.org/10.1007/978-3-030-74608-7_22

HCF =	Human Capacity Factor
HE =	Human Error
HF =	Human Factor
HITL =	Human in the Loop
HnF =	Human non-Failure
HSI=	Human-System Integration
MTTE =	Mean Time to Error
MTTF =	Mean Time to Failure
MWL =	Mental Workload
PDfR =	Probabilistic Design for Reliability
PoF =	Probability of Failure
PPM =	Probabilistic Predictive Modelling
SA =	Sensitivity Analysis
SoH =	Symptom of Health
TTF =	Time to Failure

1 Introduction

The state-of-the-art in aerospace ergonomics cannot be improved, if the outcomes of its critical missions or possible anticipated off-normal situations are not quantified in advance, and since nobody and nothing is perfect and the probabilities of failure (PoF) of a navigational device or a HITL or both are never zero, such quantification should be done on a probabilistic basis. Then the anticipated PoF and the corresponding times-to-failure (TTF) could be reviewed and, if possible, made adequate for a particular application. The PoF cannot be high, of course, but should not be lower than necessary either: it should be adequate for the particular instrumentation, mission and application. If the assessed operational PoF is very low, that could very well be an indication that the employed instrumentation is "over-engineered", i.e. is too robust and, most likely, too expensive for the given application, and/or that the HITL is overqualified for the given mission or an off-normal situation. A sensitivity analysis (SA) based on the developed methodologists and algorithms could be used, if necessary, to improve, the planned undertaking. The objective of the analyses that follow is to demonstrate how analytical ("mathematical") PPM can be fruitfully and effectively employed to predict the PoF of some more or less typical aerospace missions and extraordinary situations. Human error (HE) was addressed in numerous human psychology problems (see, e.g., [1]), including aerospace ergonomics (see, e.g., [2]), but it was done on a qualitative, rather than on a quantitative bases, not to mention using a probabilistic approach.

2 Review

An analytical convolution based PPM (see, e.g., [3]) was applied, to the authors' knowledge, for the first time in aerospace ergonomics engineering [4] in order to assess the roles of the decision making times of the two humans involved, the officer on the ship's board and the helicopter pilot, in making landing of a Navy helicopter on a ship deck successful and safe. Safe landing will take place, i.e., the helicopter undercarriage will not be damaged, if the probability that the random duration of the calm "widow" in the sea wave conditions exceeds appreciably the random sum of the two decision making times and the random time of actual landing is sufficiently large. Probabilistic assessment of the likelihood of a casualty if one of the two aircraft pilots becomes incapacitated [5] has indicated that the likelihood of a safe fulfillment of the aircraft mission might still be quite high, and will naturally depend on the duration of the flight and the moment of the mishap during the flight. The "quantitative aftermath" of the famous "miracle on the Hudson" event vs. infamous "UN shuttle" disaster has been undertaken in [6]. A way to evaluate the role of the Human-Capacity-Factor (HCF) vs. Mental-Workload (MWL) and its effect on the probability of the safe outcome of what seemed to be a "miracle" was suggested and it was argued that, in effect, the "miracle" was not that Captain Sullenberger managed to ditch his aircraft on Hudson river successfully, but because an individual with an exceptionally high hypothetical HCF/MWL ratio, like Captain "Sully", turned out to be in control, when the problem occurred. A methodology for the evaluation of the likelihood of a vehicular mission success and safety, based on a route segmentation model, was suggested [7] in application to an aircraft, with consideration of the roles of the reliability of the navigation equipment, human performance and most likely anticipated environmental conditions at each segment. The model is, generally speaking, applicable to any vehicular engineering field [8, 9], whether automotive, railway, maritime or even outer space. Several effective PPMs were developed in connection with the short- and long-term anticipation challenge in aeronautics [10]. It was shown [11] that a PPM based on the application of a double-exponential-probability-distribution function (DEPDF) can effectively quantify the role of the human factor (HF) in various HITL related tasks and problems. Several advanced probabilistic design-for-reliability (PDfR) techniques were addressed in [12–15] in application to the prediction, quantification and assurance of the reliability of aerospace electronics. Particularly, the suggested multi-parametric Boltzmann-Arrhenius-Zhurkov's (BAZ) kinetic constitutive model enabled developing methodologies for the evaluation of the PoF of an electronic device after the given time in operation at the given temperature and under the given (anticipated) stress (not necessarily mechanical). An Extreme Value Distribution (EVD) technique, which could be viewed as a special case of the DEPDF model, can be used to account for the number of repetitive loadings that eventually lead to the material/device failure by closing, in a step-wise fashion, the gap between its strength/capacity (characterized and quantified by its stress-free activation energy) and the applied stress/demand.

3 Extension

Possible Role of Failure Oriented Accelerated Testing (FOAT) in Ergonomics

Let us use as an example, aviation ergonomics [16, 17]. As is known, accelerated testing, such as highly accelerated life testing (HALT), is widely used in electronic and photonic engineering. It was shown [14–16] that highly focused and highly cost effective failure-oriented accelerated testing (FOAT) can often successfully and effectively complement

HALT in many reliability endeavors in electronics, photonics, MEMS and MOEMS, and that various FOAT means can be employed in high technologies, when failure-free operation is needed. Flight simulator could be employed as an appropriate FOAT vehicle to quantify, on the probabilistic basis, when fulfilling a particular mission, the required level F of the HCF vs the expected MWL G.

The probability of non-failure of a HITL could be sought in such tests as, say, $P = \exp\left[-\gamma t I_* \exp\left(-\frac{F}{G}\right)\right]$. Here I_* is the agreed upon high value of the continuously monitored, measured and recorded MWL characteristic I (electro-cardiac activity, respiration, skin-based measures, blood pressure, ocular measurements, brain measures, etc.), t is time and γ is the sensitivity parameter. When flight simulator is used as an appropriate FOAT test vehicle, a group of more or less equally (preferably highly) qualified individuals should be tested. The HCF is a characteristic that remains more or less unchanged for these individuals during the relatively short time of the FOAT. The MWL, on the other hand, is a short-term characteristic that can be tailored, in many ways, depending on the anticipated MWL conditions. The above equation can be written as $-G \ln\left(\frac{n}{\gamma}\right) = F = Const$, where $n = -\frac{\ln P}{I_* t}$. Let the FOAT is conducted at two MWL levels, G_1 and G_2, and the criterion I_* was observed and recorded at the times of t_1 and t_2 for the established percentages of $Q_1 = 1 - P_1$ and $Q_2 = 1 - P_2$ of failure, respectively. Then the parameter γ can be evaluated as $\gamma = \exp\left(\frac{\ln n_2 - \frac{G_1}{G_2} \ln n_1}{1 - \frac{G_1}{G_2}}\right)$. The HCF of the individuals that underwent the accelerated testing can be determined as: $F = -G_1 \ln\left(\frac{n_1}{\gamma}\right) = -G_2 \ln\left(\frac{n_2}{\gamma}\right)$.

Let, e.g., the same group of individuals was tested at two different MWL levels, G_1 and G_2, until failure (whatever its definition and nature might be), and let the MWL ratio was, say, $\frac{G_2}{G_1} = 2$. Because of that the TTF was considerably shorter and the number of the failed individuals was considerably larger, for the same I_* level (say, $I_* = 120$) in the second round of tests. Let the FOAT shows that $P_1 = 0.8$, $P_2 = 0.5$, $t_1 = 2.0\,\mathrm{h}$, and $t_2 = 1.5\,\mathrm{h}$. Then we obtain the following data: $n_1 = -\frac{\ln P_1}{t_1 I_*} = -\frac{\ln 0.8}{2x120} = 9.2976x10^{-4}$, $n_2 = -\frac{\ln P_2}{t_2 I_*} = -\frac{\ln 0.5}{1.5x120} = 38.5082x10^{-4}$, $\gamma = \exp\left(\frac{\ln n_2 - \frac{G_1}{G_2} \ln n_1}{1 - \frac{G_1}{G_2}}\right) = \exp\left(\frac{\ln 38.5082x10^{-4} - 0.5 \ln 9.2976x10^{-4}}{1 - 0.5}\right) = 0.015948$, $\frac{F}{G_1} = -\ln\left(\frac{n_1}{\gamma}\right) = \ln\left(\frac{9.2976x10^{-4}}{0.015948}\right) = 2.8422$, $\frac{F}{G_2} = -\ln\left(\frac{n_2}{\gamma}\right) = \ln\left(\frac{38.5082x10^{-4}}{0.015948}\right) = 1.4210$.

The calculated required HCF-to-MWL ratios $\frac{F}{G} = -\ln\left[62.7038\left(\frac{-\ln P}{t}\right)\right]$ for the given/required probabilities P of non-failures are shown in the table:

P	0.95	0.99	0.999	0.9999	0.99999
t, h	×	×	×	×	×
48	2.7030	4.3329	6.6400	8.9431	11.2457
240	4.3124	5.9424	8.2495	10.5525	12.8551
720	5.4110	7.0410	9.3481	11.8511	13.9537
8760	7.9097	9.5397	11.8468	14.1498	16.45424

As evident from the calculated data, the level of the HCF in this example should exceed considerably the level of the MWL, so that a high enough value of the probability of human-non-failure is achieved, especially for long operation times.

Adequate Trust as an Important Part of the HCF and HSI

Since classical Shakespearian "love all, trust a few" and "don't trust the person who has broken faith once" and to the today's ladygaga's "trust is like a mirror, you can fix it if it's broken, but you can still see the crack in that mother f*cker's reflection", the importance of human-human trust was addressed by numerous writers, politicians and human psychologists. It was the 19th century South Dakota politician Frank Craine who seems to be the first one who indicated the importance of an *adequate trust* in human relationships: "You may be deceived if you trust too much, but you will live in torment unless you trust enough".

It is shown [18] that the entropy of the DEPDF distribution, when applied to the *trustee* (a human, a technology, a methodology, a concept, etc.), can be viewed as an appropriate quantitative characteristic of the propensity of a decision maker in a HSI situation to an under-trust or an over-trust judgment and, as a consequence of that, to the likelihood of making a mistake or an erroneous decision. The analysis that follows addresses some important aspects of a HITL problem for safety-critical missions and extraordinary situations. It is argued that the role and significance of trust can and should be quantified when preparing such missions. Certainly, it should be considered in any HSI activity, and the concept of an adequate trust should be included into an engineering technology, design methodology or a human activity, when there is a need to assure a successful and safe outcome of a particular engineering effort or an aerospace or a military mission.

A suitable modification of the DEPDF for the human non-failure, whether it is the performer (decision maker) or the trustee, is assumed here in the following simple form $P = \exp\left[-\gamma t \exp\left(-\frac{F}{G}\right)\right]$, where P is the probability of non-failure, t is time, F is the HCF, G is the MWL, and γ is the sensitivity factor for the time. This expression makes physical sense. Indeed, the probability P of human non-failure, when fulfilling a certain task, decreases with an increase in time and increases with an increase in the HCF/MWL ratio. At the initial moment of time ($t = 0$) the probability of non-failure is $P = 1$ and exponentially decreases with time, especially for low F/G ratios. The above expression, depending on a particular task and application, could be applied either to the performer (the decision maker) or to the trustee, who could be a human, a technology, a concept, an existing best practice, etc.

The ergonomics underlying the above distribution could be seen from the time derivative $\frac{dP}{dt} = -\frac{H(P)}{t}$, where $H(P) = -P \ln P$ is the entropy of this distribution. Thus, the accepted distribution reflects an assumption that the time derivative of the probability P is proportional to the entropy $H(P)$ of this distribution and decreases with an increase in time. This entropy, when applied to the distribution in question is the probability of non-failure of the trustee's performance, and is zero for both extreme values of this performance: when the probability of non-failure is zero, it should be interpreted as an extreme under-trust in someone else's authority or expertise, which is the case of a "not invented here (NIH)" attitude; when the probability of the trustee's non-failure is one, that means that there is an extreme over-trust in an NIH technology: as is known, "my neighbor's grass is always greener" and "no man is a prophet in his own land". The entropy $H(P)$ reaches its maximum value $H_{\max} = e^{-1} = 0.3679$ for a rather moderate probability $P = e^{-1} = 0.3679$ of non-failure of the trustee. Note that in the well-known and still widely used Arrhenius equation $\tau = \tau_0 \exp\left(\frac{U_0}{kT}\right)$ MTTF τ, the time τ is, in effect, the time t needed, in accordance with the exponential law of reliability $P = \exp(-\lambda t)$, to reach the maximum entropy $H_{\max}$ of the probability of the Arrhenius (actually, Boltzmann's) distribution. Indeed, by replacing the failure rate λ in the exponential law of reliability with its reciprocal value $\frac{1}{\tau}$ the following expression for the entropy can be obtained: $H(P) = -P \ln P = \exp\left[-\frac{t}{\tau_0} \exp\left(-\frac{U_0}{kT}\right)\right]\left[-\frac{t}{\tau_0} \exp\left(-\frac{U_0}{kT}\right)\right] = \frac{t}{\tau} \exp\left(-\frac{t}{\tau}\right)$. The time t, when the entropy of the distribution $P = \exp(-\lambda t)$ reaches its maximum value $H_{\max} = e^{-1}$, can be found from the equation $e^{-1} = \frac{t}{\tau} \exp\left(-\frac{t}{\tau}\right)$. This equation yields: $t = \tau$. Trust is an important HCF quality.

Captain "Sully", the hero of the miracle-on-the-Hudson event, did possess such a quality. He "*avoided over-trust*" in the ability of the first officer, who ran the aircraft when it took off La Guardia airport, to successfully cope with the emergency situation and took over the control, as well as in the possibility, with the help of the air traffic controllers at LaGuardia and at Teterboro, to land the aircraft safely at these airports. What is even more important, is that Captain "Sully" also *"avoided under-trust"*: 1)in his own skills, abilities and experience that would enable him to successfully cope with the situation (57-year-old Captain "Sully" was a former fighter pilot, a safety expert, an instructor and, most importantly, a glider pilot), and that was the case when "team work" was not the right thing to pursue (quite often, as is known, "too many cooks spoil the broth"); 2) in the aircraft structure that would be able to withstand the slam of the water during ditching and, in addition, would enable slow enough flooding after ditching; in the aircraft safety equipment that was carried in excess of that mandated for the flight; 3) in the outstanding cooperation and excellent cockpit resource management among the flight crew who *trusted* their captain and exhibited, after actual ditching, outstanding team work, when such work was needed during the rescue operation; 4) in the fast response from and effective help of the various ferry operators located near the USS Intrepid museum and in the ability of the rescue team to provide timely and effective help; and 5) in the exceptionally good visibility, an important contributing factor to the success of the accident.

Here is another indication on the role that the maximum entropy can be employed as a suitable quantitative criterion of the adequate trust. It has been determined [19] that the age groups of 20–25 and 65–70 years old are more prone, for the same driving time, to driver drowsiness (DD) than the 26–64 age group. Of course, this happened, first of all, because the middle-aged group possesses the best combination of experience and personal qualities, i.e., has, in general, a higher HCF. Fifty-four years old Captain "Sully" is a good example. The following human qualities have to do with age: ability to concentrate, to anticipate, to withstand fatigue (both physical and mental) for a long time (tolerance to stress), to act in cold blood in off-normal and even life-threatening situations, to make well substantiated decisions in the conditions of uncertainty, to operate the vehicle effectively under time pressure; situation awareness; self-control; mature (realistic) thinking; swiftness in reaction, when necessary; ability to maintain an optimum level of psychological arousal. All these qualities are naturally stronger in mature middle-aged individuals than in younger ones, not to mention elderly folks.

But what is much less obvious is that adequate trust in a system is equally important. The driver should definitely and first of all have adequate trust in himself/herself (again, Captain "Sully" is a good example), but, at the same time, have a reasonable, but a moderate, trust in the automated driving (AD) system. The system should be subjected FOAT for the most important anticipated missions and possible extraordinary situations, and the vehicle operator should be informed of these tests and their results for the given system. Such an information should be included into his/hers education as a driver of an automated driven vehicle and reflected in his/hers driver license. As to the human propensity to driver drowsiness (DD), we suggest, based on the study [19], that the area under the corresponding portion of the entropy curve is used as a suitable quantitative measure of the propensity of a particular age group to the DD "syndrome".

The entropy curve area located between two arbitrary probabilities is as follows: $A(P) = \int_{P_1}^{P_2} H(P)dP = \int_{P_2}^{P_1} P\ln P dP = \frac{1}{2}\left[P_1^2\ln P_1 - P_2^2\ln P_2 + \frac{1}{2}\left(P_2^2 - P_1^2\right)\right]$. The area under the entire entropy curve can be found by putting $P_1 = 0$ and $P_2 = 1$ in this expression and is $A_0 = A(1) = 0.25$. The age of drivers under test [19] ranged from 20 to 70, i.e., for 50 years, and since the total length of the probability axis is 1.0, each year corresponds therefore to the segment 1/50 = 0.02 on the P axis. The lengths of the segments for the 20–25 "agers" and 65–70 "agers" are the same and are equal to 0.02 × 5 = 0.1. The 20–25, 65–70 and 26–64 areas under the H(P) curve are:

$$A_{65-70} = A(P_1 = 0.9; P_2 = 1) = \frac{1}{2}\left[-0.9^2\left(\ln 0.9 - \frac{1}{2}\right) + \frac{1}{2}\right] = 0.00483$$

$$A_{26-64} = A(P_1 = 0.1; P_2 = 0.9) = \frac{1}{2}\left[0.1^2\ln 0.1 - 0.9^2\ln 0.9 + \frac{1}{2}\left(0.9^2 - 0.1^2\right)\right]$$
$$= 0.23111$$

These data could be viewed as figures of merit for these three groups of people, as far as their ability to withstand DD is concerned. If one takes the performance of drivers in the 24–64 age group as 100%, the younger drivers are only 6.06% as good as the middle age people, and the older drivers are even worse – only 2.09%. These results should be attributed, first of all, to the belief that the level of the propensity to DD is lower, when the driver has more trust in himself/herself rather than in the system and because of that his/hers level of awareness is higher.

4 Conclusions

In any ergonomics effort of importance FOAT should be carried out beforehand to predict its outcome. Using several realistic examples, it was shown how significant should the ratio HCF/MWL be to make the probability of the human failure low enough. Flight simulator can be employed as an appropriate and ergonomically meaningful test vehicle that could be used to quantify, on the probabilistic basis, the required HCF/MWL for the successful fulfilment of a particular aerospace mission or to successfully cope with an extraordinary situation Human trust is an important HCF quality and HSI feature and should be included into the list of such qualities for a particular HITL task in aerospace ergonomics. The entropy of the double exponential probability distribution function (DEPDF) for the random HCF can be viewed as an appropriate quantitative characteristic of the propensity of a human to an under-trust or an over-trust judgment and, as the consequence of that, to an erroneous decision making or to a performance error.

References

1. Reason, J.T.: Human Error. Cambridge University Press, Cambridge, UK (1990)
2. Kern, A.T.: Controlling Pilot Error: Culture, Environment, and CRM (Crew Resource Management). McGraw-Hill, New York (2001)
3. Suhir, E.: Applied Probability for Engineers and Scientists. McGraw-Hill, New York (1997)
4. Suhir, E.: Helicopter-landing-ship: undercarriage strength and the role of the human factor. ASME Offshore Mech. Arctic Eng. J. **132**(1) (2009)
5. Suhir, E., Mogford, R.H.: “Two men in a cockpit”: probabilistic assessment of the likelihood of a casualty if one of the two navigators becomes incapacitated. J. Aircr. **48**(4) (2011)
6. Suhir, E.: “Miracle-on-the-Hudson”: quantified aftermath. Int. J. Hum. Factors Model. Simul. **4**(1), 35–62 (2012)
7. Suhir, E.: Likelihood of vehicular mission-success-and-safety. J. Aircr. **49**(1) (2012)
8. Suhir, E.: Human-in-the-loop (HITL): probabilistic predictive modeling of an aerospace mission/situation outcome. Aerospace **1**, 101–136 (2014)
9. Suhir, E.: Human-in-the-loop: probabilistic predictive modeling, its role, attributes, challenges and applications. Theor. Issues Ergon. Sci. (2014)
10. Suhir, E., Bey, C., Lini, S., Salotti, J.-M., Hourlier, S., Claverie B.: Anticipation in Aeronautics: Probabilistic Assessments. Theor. Issues Ergon. Sci. (2014)
11. Suhir, E.: Human-in-the-loop: application of the double exponential probability distribution function enables one to quantify the role of the human factor. Int. J. Hum. Factors Model. Simul. **5**(4), 354–377 (2017)
12. Suhir, E.: Aerospace electronics reliability prediction: application of two advanced probabilistic techniques. ZAMM **1**(16) (2017)
13. Suhir, E.: Aerospace mission outcome: predictive modeling, editorial, special issue “challenges in reliability analysis of aerospace electronics”. Aerospace **5**(2) (2018)

14. Suhir, E.: What could and should be done differently: failure-oriented-accelerated-testing (FOAT) and its role in making an aerospace electronics device into a product. J. Mater. Sci.: Mater. Electron. **29**(4), 2939–2948 (2018)
15. Suhir, E.: Human-in-the-Loop: Probabilistic Modeling of an Aerospace Mission Outcome. CRC Press (2018)
16. Suhir, E.: Failure-oriented-accelerated-testing and its possible application in ergonomics. Ergon. Int. J. **3**(2) (2019)
17. Suhir, E.: Assessment of the required human capacity factor (HCF) using flight simulator as an appropriate accelerated test vehicle. Int. J. Hum. Factor Model. Simul. **7**(1), 71–74 (2019)
18. Suhir, E.: Adequate trust, human-capacity-factor, probability-distribution-function of human non-failure and its entropy. Int. J. Hum. Factor Model. Simul. **7**(1), 75–83 (2019)
19. Kundinger, T., Riener, A., Sofra, N., Weigl, K., Driver drowsiness in automated and manual driving: insights from a test track study. In: Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI 2020). ACM (2020)