

Original Research Article

Improving Learning in Robotics Teleoperation: Contribution of Eye-tracking in Digital Ethnography

ABSTRACT

Aims: Digital ethnography has shown its added-value for the analysis of work activity in order to improve the methods of developing the associated necessary competencies. In particular, tracing process methods based on first-person recording of the activity by first-person view camera combined with competencies-oriented and goal-oriented interviews have demonstrated their effectiveness in medicine, nuclear industry and education. However, the teleoperation of robots out of sight requires the use of a specific first-person view camera: an eye-tracking device. This is due to the fact that during the teleoperation, the pilot's head movements are almost non-existent while the eyes move enormously. Yet, the literature is completely void of this type of activity analysis using eye-tracking for teleoperation in robotics. The objective of the present article is to fill this gap by presenting a pilot study characterizing the potential contribution of such a method for this type of work activity.

Study design: The pilot study has involved two robot pilots individually performing a teleoperation task out of sight. The pilots have been chosen for their difference of experience in teleoperation. Whilst performing the activity, they have been equipped with an eye-tracking device, enabling the recording of the activity at the first-person perspective. An interview based on the Square of PERceived Action model (SPEAC model) has followed in order to access what makes their competencies.

Place and Duration of Study: The experiments were undertaken in the simulation training center of the Groupe INTRA-Intervention Robotique sur Accidents, in France, during 2023.

Methodology: Two pilots had to individually teleoperate a robot from a control console using the videos transmitted on several screens placed in front of them from the cameras on board the robot. The activity consisted of transporting the robot through a maze carrying a container in which a ring had to be put after being picked up from the ground, and then bring the whole out of the maze. The activity lasted about 10 to 20 minutes. Each pilot was equipped with an eye-tracking device that made it possible to record their activity for deferred access in order to identify the knowledge and know-how implemented. The interview was conducted using the SPEAC model. At the end of the interview, a matrix of competencies was built for each of the pilots. Software processing made it possible to access quantified data, in particular the vision fixation time for each of the pilots in order to take information from the screens and the control console.

Results: The comparison of the matrices of competencies made it possible to measure the gap in competence between an experienced pilot and a novice pilot, as well as to identify knowledge and know-how not yet taught in pilot training. The measurement of fixation times has also made it possible to identify a difference that appears interesting to be analyzed in more depth in a future study.

Conclusion: Results shows that the method applied is well suited for teleoperation of robots out of sight and provide relevant data to improve training.

Keywords: digital ethnography, eye-tracking, knowledge management, robotics, teleoperation, training

1. INTRODUCTION

Process tracing, a method for analyzing activities being part of the paradigm of Cognitive Task Analysis (CAT) described in detail in [1,2] has shown its performance to improve competencies (e.g. [3]).

Among the tracing process techniques, the use of video recordings of work activity for its analysis is identified in CAT as one of the most effective. Recording techniques can be based on the use of a camcorder for third person perspective, or a miniature on-board camera (most often mounted on a pair of spectacles) in order to obtain a first-person subjective recording of the work activity. For the latter case, it relates to Subjective Evidence Based Ethnography (SEBE; [4]). The method showed all its added-value compared to a third person recording [5-7]: the subjective video recording is analyzed during an interview between the analyst and the subject according to a self-confrontation technique [8,9]: the subject is led to watch and comment on the recording of the activity whilst being questioned by the analyst. Indeed, it has been shown that the interview analysis of a subjective recording of an activity (also called subjective re-situ interview or replay interview) presents an objective trace of the past activity to the subject, who can rely on this stable and reliable basis to make comments and a precise explanation of the actions reviewed in the situation [9,10]. According to Lahlou [4:623], in the framework of the SEBE, the subjective recording "provides material that is especially relevant for the reconstitution of the mental activity", a "recall effect [...] probably due to the nature of episodic memory [...], a multimodal association connected to an actual lived event (time, place, associated emotions, intentions, contextual knowledge and other associations), which come back as a bundle when the subject recalls the event."

The SEBE method has been used in various contexts and has proven its effectiveness in assisting in the analysis of human activity [11,12], especially for learning in education (e.g. [11,13,14] or for professional activity (e.g. [11,15-20]. However, in some cases, the equipment used for video capture appears sensitive to the type of work analyzed. Indeed, the use of a miniature camera mounted on a pair of spectacles is well suited to the capture of a subjective video when the person whose activity is analyzed evolves in a wide environment where head movements accompany the gaze: the analysis of the recorded video then makes it easy to understand what the person is looking at. This is fundamental when analyzing the recorded video. This is the case of a field agent working on an industrial installation [17-18] or an operator working in the control room of a complex industrial process [11]; for example, nuclear reactor pilots [15,17,18,21] or refinery operators. This can also be applied to activities in a more restricted environment such as doctors in hospitals [16]. On the other hand, this type of camera is inappropriate when the person whose activity is analyzed evolves in a narrow environment. This is the case, for example, of a robot pilot who teleoperates a robot out of sight whilst watching one or more screens and using a keyboard or a pair of joysticks. The miniature camera can no longer point where the gaze focuses: the main movements come from the eyes rather than from the head; so it is necessary to use an eye-tracking device. This type of device has recently been applied to the INTRA robotics group to analyze teleoperation activities. It is a first as the literature is void of such studies. This was undertaken in the framework of a pilot study aiming at assessing the potentiality of the SEBE method to contribute towards the improvement of the robotic pilots' performance.

The objective of this article is to present the pilot study, the results obtained and their analysis, as well as the discussion of their limitations and the associated perspectives, and

to conclude on the relevancy of the SEBE method based on eye-tracking recording for robotics teleoperation.

2. MATERIAL AND METHODS

2.1 Overall protocol to analyze pilots activity

The analysis of pilots activity was carried out according to a digital ethnography method applying the Square of PErceived Action protocol (or SPEAC protocol) described in [11]. It starts with capturing the individual's activity at the first person to get a subjective record of what has been done. Based on the subjective recording of the activity, the pilot is then interviewed by the analyst to determine what allows him/her to successfully put the competencies into action. The interview is based on the SPEAC model, which considers that an individual can successfully put his/her competencies into action provided that the four poles of the model are available and coherent. These four poles are:

- *Having to act*
- *Knowing to act*
- *Being able to act*
- *Wanting to act*

They are available when all the conditions are met to be effective in action. For example, the pole *Knowing to act* is available if the individual has all the necessary knowledge and has developed the know-how necessary to carry out the activity.

They are consistent when the interactions between poles are not antagonistic. For example, the *Having to act* and *Wanting to act* poles are coherent if what is requested by the procedure and by the order-giver does not oppose what the individual wants to do.

Previous studies have shown that the *Having to act* and *Knowing to act* poles are essentially determined before the activity [11]. Therefore, during the replay interview, they are questioned before viewing the recording of the activity. Their questioning then continues during the analysis of the video as for the other two poles, which are also questioned in detail after the video has been viewed. The type of questions asked is available in [11:117-120]. The replay interview is also goal-oriented: the analyst tries to make verbalized what the subject aimed to do and what the subject wanted to avoid.

The results of the interview are then analyzed by the analyst and transformed into a matrix of knowledge and know-how, individual and collective, tacit or explicit, identified as necessary to carry out the activity successfully. The resulting matrix is then presented to the individual who validates or adjusts the results.

The final matrix is used to improve training programs, and therefore the performance of the people involved in the analyzed activity. In the present case, the improvements will be applicable to the training of all pilots in the INTRA group. Improvements are identified when the matrices are validated by the pilots participating in the analysis: they are invited to identify the knowledge and know-how that are not explicitly taught during the training.

The knowledge and know-how identified as an improvement are finally categorized according to the classification of Suzuki et al. [22] who proposed a hierarchical classification of competencies in mechatronics teleoperation. This categorization makes it possible to identify the classes that might be improved in the training program. The hierarchical classification of competencies identifies 5 schematized classes: Fig. 1.



Fig. 1. Hierarchical classification of skills in mechatronics teleoperation by [22]

“The hierarchy consists of the following skills: social- (S1), planning- (S2), cognitive- (S3), motion- (S4) and sensorymotor (S5) skills. The S5, a sensory-motor skill, is the lowest and relates mainly to voluntary motion. This skill concerns cooperation between the neural system and the musculo-skeletal system inside the human body. S4 (motion skill) relates to the continuous execution of each body part motion. For example, the sequential execution of finger, hand and arm motions when using a hand tool. S3 is cognitive skill. This mainly concerns recognition of circumstance and includes a wide range of cognitive issues, such as the understanding of meanings. S2 (planning skill) relates to understanding a whole task process and recognising sub-tasks within the whole task. This skill also relates to the task management of subtasks, which requires consideration of events' causality and time scheduling. S1 is social skill, which relates to communication, conversation, negotiation and the estimation of other intentions” [22:2].

It is clear that classes S4 and S5 concern prerequisites associated with the physical and psychological abilities of pilots and do not concern training in robotics teleoperation. The analysis therefore covers classes S1, S2 and S3.

2.2 Equipment

The device used was the one provided by Pupils Lab. It consists of spectacles with adapted glasses correction to the user, with a wide angle camera mounted on the glasses stem, detectors on the glasses frame to capture and record the pupils movement, a wire between one spectacles stem and a smartphone for recording.

The equipment allows us to obtain a subjective video of what the person is looking at during the activity and a software posttreatment locates the points of fixation of the person's gaze on the video.

2.3 Data

In addition to the video recording, the software associated with Pupils Lab device makes it possible to extract from the video a series of quantified data. In this case, the data used was the time taken to fix particular sequences of activity as well as the number of these fixation points, and the number of switches between screens and joysticks.

As indicated in the “Material &Method” section, the SPEAC protocol produces a matrix of competencies needed to carry out the activity successfully. In addition to the data obtained through the replay interview to access what makes the pilot's competencies, demographic data (age, number of years of experience in the job, effect of past occupational experience on the job), and physiological data (vision problem)) were collected.

2.4Occupational activity analyzed

The activity chosen for the analysis was a simple activity of remote operation of a terrestrial robot out of sight: the pilot does not see the robot evolving; s/he teleoperates the robot from the videos sent back on the screens in front of her/him by the cameras embedded on the robot.

The robot was a tracked ground machine equipped with an articulated arm (Fig. 2 & 3). The robot had to move through a maze (Fig. 2) by carrying a cylindrical container which had been opened at the top in the gripper at the end of its arm (Fig. 3). Halfway through the maze, the container had to be put on the ground, and a ring had to be picked up with the gripper of the articulated arm, then dropped into the container, then the container had to be picked up to get out of the maze. This activity was carried out in a full-scale simulator in order to reduce the risks associated with the safety of people or the safety of equipment. For this reason, the recommended risk analysis prior to the use of subjective cameras [23,24] was not applied in this study.



Fig. 2. Overview of the maze

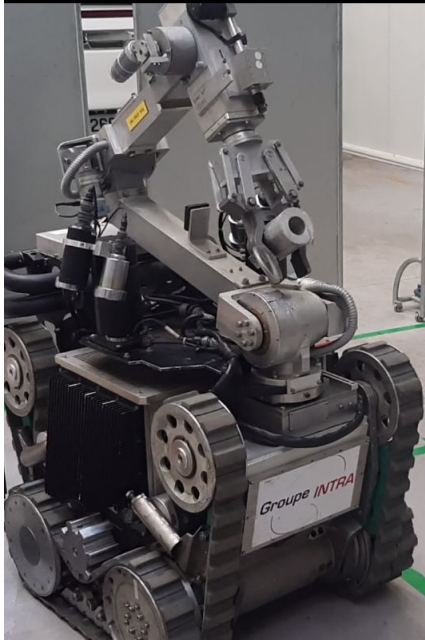


Fig. 3. Front view of the robot with the container in the gripper

The pilot was installed at the control desk consisting of a control console with buttons and joysticks, multiple screens associated with each of the on-board cameras, and a keyboard to adjust the zoom of certain cameras (Fig. 4).



Fig. 4. A pilot in front of the control console whilst teleoperating the robot

2.5 Participants

Two pilots have been selected from the INTRA group on a voluntary basis. The first pilot was chosen on the basis of his long professional experience and high level of competencies,

being a reference pilot within the group (Pil exp). The second pilot was chosen on the basis of his inexperience (Pilnov): the recording of the activity was carried out during his first training as a new pilot within the group. Both pilots' physiological and demographic data are presented in Table 1.

Table 1. Characteristics of the participants

	Experienced pilot	Novice pilot
Subject reference	Pil exp	Pilnov
Gender	male	male
Age (y)	53	48
Vision problem	myopia	none
Experience within the job (years)	10	0.08
Past occupational experience (years)	5	11
Past occupational experience (domain)	Teleoperation of robots in duct or pool on NPP(*)	Teleoperation of fuel loading machine on NPP(*)

(*)NPP: Nuclear Power Plant

3. RESULTS AND DISCUSSION

The data extracted by the operating software of the Pupil Lab device are given in Table 2. The data is associated with the performance metrics. The time it takes to carry out the handling part has been extracted from the overall time to complete the activity. The time taken to complete the handling part is measured from the viewing of the video recordings. The handling part starts from the moment the cylindrical container is placed on the ground and ends when it is taken up in the robot gripper. The proportion of the handling time in relation to the overall performance time of the activity is calculated. The number of switches of the pilot's gaze moving from the screens to the control console and vice versa was calculated by looking at the video recordings. Indeed, the Pupil Lab system provides measured information on the azimuthal position of the gaze, but this is not usable due to the bias induced by head movements, the reference for software measurement being the frame of the spectacles.

Table 2. Quantified data obtained from the eye-tracking software treatment

Participant Reference	Pil-nov	Pil-exp
activity duration (min)	17.70	12.26
handling duration (min)	10.68	7,17
proportion handling duration within total duration	60%	58%
number of fixation points	2189	1478
total fixation duration (s)	787,765	556,114
fixation proportion of activity duration	74%	76%
fixation peak range (ms)	[150; 200]	[250; 270]
nb switches screens and control console	77	62

The interviews were conducted individually with each of the pilots, based on first-person video recordings of the pilot's work activity (screenshot of a recording Fig. 5). This made it possible to establish matrices of competencies for each of the pilots. Table 3 shows the number of individual or collective knowledge and know-how, explicit or tacit, for the experienced pilot (Pil exp) and for the novice pilot (Pilnov). In the present study, the activity being individual and not collective, identified knowledge or know-how were mainly individual. The ratio between the number of tacit and explicit knowledge and know-how has been calculated. The number of knowledge and know-how commonly identified by the novice pilot

and the experienced pilot was calculated, and for each of the two pilots, the proportion of tacit knowledge and know-how compared to the explicit ones was calculated.

Table 3. Number of individual or collective knowledge and know-how, explicit or tacit, per pilot

	Pil-nov	Pil-exp
Individual K&KH(*)	20	15
collective K&KH	0	1
explicit K&KH	16	15
tacit K&KH	4	1
tacit/explicit K&KH ratio	25%	6.6%
common K&KH (number/%tacit/%explicit)		8/0/100
K&KH identified to improve training	3	1

(*)K&KH: Knowledge and Know-How

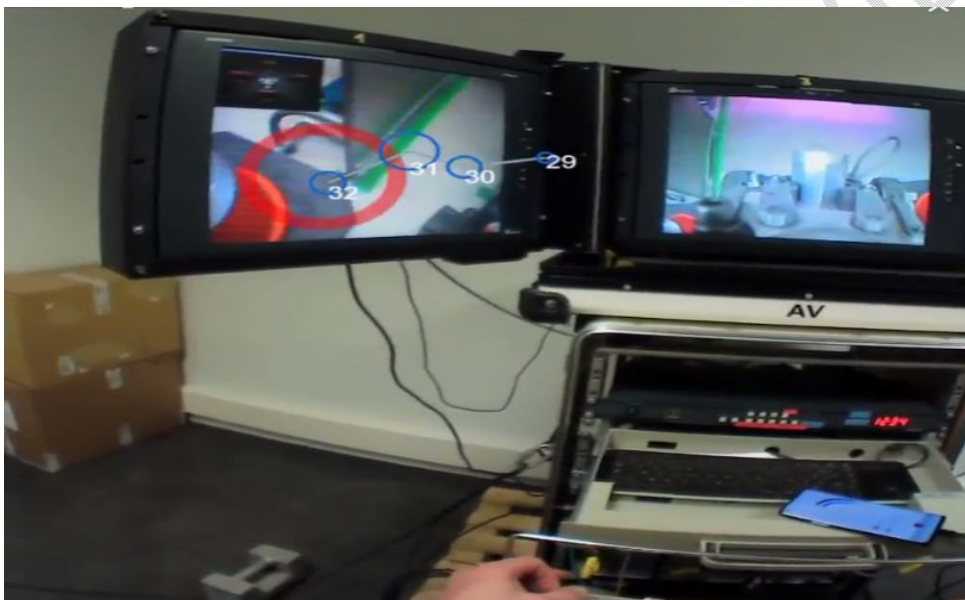


Fig. 5. Screenshot of the eye-tracking video recording and localization of the gaze (red circle) and fixation points (blue circles)

The validation of the skills matrices with the pilots made it possible to retain 4 knowledge and know-how not explicitly taught during the training, and therefore they were identified as areas for improvement for the training. The analysis according to the hierarchical classification of competencies in mechatronics by Suzuki et al. [22] leads to the following categorization:

- S1 Social skill
 - Positively welcome the advice of the Pilot Supervisor to adjust teleoperation
- S2 Planning skill
 - Understand the activity as a whole in order to break down the stages of teleoperation mentally before the action (action scheme strategy)
- *S3 Cognitive skill
 - Identify the way out (be attentive) to return to the starting point (end of the activity)

- Be aware of the times induced by teleoperation actions in order to optimize the placement of the robot or objects

The measurements indicate that the overall time for the novice pilot to complete the activity is 41% higher than for the experienced pilot. This is not surprising since an experienced pilot necessarily performs an activity more quickly than a novice pilot. The proportion of handling time is 60% for the novice pilot and 58% for the experienced pilot, which does not represent a significant difference. It may be assumed that the novice pilot experiences the same difficulties during the handling part as during the whole activity and that these difficulties take on the same proportion for the experienced pilot.

The gaze fixation time is 41% longer for the novice pilot compared to the experienced pilot. The same discrepancy is noted here as for the overall time of completion of the activity. The proportion of fixation time to the overall activity time is 74% for the novice pilot and 76% for the experienced pilot, which does not represent a significant difference. This may be understood as both pilots spend the same proportion of time setting informative cues during the handling activity, regardless of their level of experience. The average value of the fixation time differs by 4% between the two pilots, however, when ranking the values per time slots, the distribution of fixation times is different for each of them and the extremum of the curve is significantly different: in ms, interval [100-170] for the novice pilot and [170-240] for the experienced pilot. This difference can be attributed to at least three reasons that can be combined:

- the age difference between the pilots,
- the difference in vision ability between pilots,
- the influence of experience which leads the experienced pilot to spend more time on a fixation point in order to take in more information than the novice pilot.

In order to better understand this point, further experiments on a wide range of pilots are necessary.

The comparison of the matrices of competencies has highlighted 4 knowledge and know-how to add to the training program. This knowledge and know-how were identified as not being taught during the training of pilots in the INTRA group. As this method has never been applied to mechatronics teleoperation for activity analysis in order to improve pilot training, the results obtained here show the added-value of its application. This conclusion confirms the conclusions of similar studies carried out in other sectors of activity, in particular nuclear reactor operation [15,17,18,21], medicine [16] and school education [13,14].

The analysis of areas of improvement in terms of knowledge and know-how according to the hierarchical classification of competencies in mechatronics by Suzuki et al. [22] shows that the three upper levels of the classification identified as relevant are affected by these improvements. This shows that the SEBE approach applying the SPEAC protocol based on eye-tracking recordings in mechatronics teleoperation is robust enough not to forget any category of the classification.

The limitations identified for the present pilot study are as follows.

This pilot study showed the added-value of applying a SEBE method integrating eye-tracking in robotic teleoperation in order to improve pilot training. The areas for improvement identified in terms of competencies to be developed in training concerned all the upper classes of the hierarchical classification of competencies in mechatronics by Suzuki et al.

[22]. However, identification is not systematic. It depends on the activity analyzed, the construction of the associated training program, and the pilot participating in the analysis. Thus, to improve a training program for a given activity, it is necessary to conduct analyses with several pilots.

The present study was limited to demonstrating the possibility of identifying areas for improvement for training in robotics teleoperation. However, an area for improvement is only effective when the results it produces, in terms of real improvement on the activity analyzed, are measured and effective. However, the present study did not include the evaluation of the pilots' performance after adjusting their training. The study to be followed will have to integrate this aspect.

4. CONCLUSION

The purpose of the present pilot study was to evaluate the added-value of applying a SEBE analysis using eye-tracking in order to identify areas for improvement of training programs for robotics teleoperation. This objective was achieved successfully: 4 competencies have been identified for the improvement of the training program related to a simple work activity, and the three upper levels of hierarchical classification of competencies in mechatronics by Suzuki et al. [22] were concerned by these competencies, thus contributing to showing the robustness of the method applied. The method will now be applied to more complex robotics teleoperation activities.

The use of quantified data measured using the processing software of the eye-tracking device also made it possible to hypothesize that an experienced pilot spends more time acquiring information on a control screen than a novice pilot and thus gains in efficiency. As the present pilot study was carried out with only two pilots, the research perspective consists in multiplying these measurements on a wide spectrum of pilots in order to validate or not this hypothesis. This might yield another area of improvement for training.

CONSENT

Each participant has co-signed with the supervisor of the experiments an informed consent explaining the aim of the study, the participants profile expected, the procedure and duration of the experiment, the equipment used and how, the benefits and the risks for the participants and what was made to reduce the risks, the debriefing following the activity, and recalled the volunteer character of the participation as well as the participants' right to leave as they wanted. The informed consent was written according to the *British Psychological Society Guidelines for Minimal Standards of Ethical Approval in Psychological Research*.

ETHICAL APPROVAL

All authors hereby declare that the design of the study was examined and approved by the Committee for Ethics, Standards and Protection of the INTRA group, approval number #CESP/PHFA/100, and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

REFERENCES

1. Wei J, Salvendy G. The cognitive task analysis methods for job and task design: review and reappraisal. *Behaviour & Information Technology*. 2004;23(4):273-299.
2. Tofel-Grehl C, Feldon DF. Cognitive Task Analysis–Based Training: A Meta-Analysis of Studies. *Journal of Cognitive Engineering and Decision Making*. 2013;7(3):293-304.

3. Le Bellu S, Lahlou S, Nosulenko V. Capturing and transferring the knowledge incorporated into a professional gesture. *Social Science Information*. 2010;49:371-413. French.
4. Lahlou S. How can we capture the subject's perspective? An evidence-based approach for the social scientist. *Social science information*. 2011;50(3-4), 607-655.
5. Lahlou S. Activity from the actor's point of view and the question of inter-subjectivity: eight years of experiments with miniaturized cameras fixed to the actors' foreheads (subcam). *Communications*. 2006;(80), 209-234. French.
6. Lahlou S, Le Bellu S, Boesen-Mariani S. Subjective evidence-based ethnography: method and applications. *Integrative Psychological and Behavioral Science*. 2015;49(2), 216-238.
7. Le Bellu S. Capitalization of know-how and professional gestures in the industrial environment. PhD thesis, Un. of Bordeaux, reference 2011BOR21825; 2011. French.
8. Von Cranach M, Kalbermatten U. The Ordinary interactive action: theory, methods and some empirical findings. In: Von Cranach M, Harre R, editors. *The Analysis of Action*. London: Cambridge University Press; 1982.
9. Theureau J. The self-confrontation interview as a component of an empirical and technological research program. Paper presented at the 11^{ème} Journées Internationales des Sciences du Sport, Paris, 12-15 Nov 2002. French.
10. Rix G, Biache M. Recording in Situated Subjective Perspective and re-situ Subjective Interview: A Methodology for the Creation of Experience. *Intellectica*. 2004;1(38):363-396. French.
11. Fauquet-Alekhine Ph. Knowledge Management in High-Risk Industries – Coping with Skills Drain. London: Palgrave Macmillan; 2020.
12. Fauquet-Alekhine Ph, Lahlou S. Subjective Evidence-Based Ethnography and Psychology. Proceedings of the Vth All-Russian Scientific and Practical Conference with international participation in memory of M.Yu. Kondratiev, Moscow State Pedagogical University, Moscow, May 12 - 13 2020:535-538. Russian.
13. Fauquet-Alekhine Ph, Martinez P. Improving the competencies of students through the use of the square of perceived action model. Conference "Social Psychology and Society: History and Modernity", Moscow, 15-16 October 2019:475-479. Russian.
14. Fauquet-Alekhine Ph, Martinez P. Cognitive Task Analysis to Improve Student Competencies: a Pilot Study based on the Use of the Square of Perceived Action Model. *Journal Educational Verkenning*. 2021;2(1):20-28.
15. Fauquet-Alekhine Ph, Bauer MW, Lahlou S. Introspective interviewing for work activities: applying subjective digital ethnography in a nuclear industry case study. *Cognition, Technology & Work*. 2021;23(3):625-638.
16. Fauquet-Alekhine Ph, Bouhours G, Texier J, Loret A, Lahlou S, Granry JC. Training improvement through subjective work analysis: The example of radial puncture for arterial blood gas test. *Global Journal of Health Science*. 2021;13(3):59-69.

17. Fauquet-Alekhine Ph, Lahlou S. Performance in collaborative activity: contribution of intersubjectivity theory. *Journal of Psychology & Behavior Research*.2022 ;4(1):11-41.
18. Fauquet-Alekhine Ph, Lahlou S. Factors of efficiency in collaborative activity: the contribution of the theory of intersubjectivity and digital ethnography. *Proceedings of the VIIth All-Russian Scientific and Practical Conference with international participation in memory of M.Yu. Kondratiev, Moscow State Pedagogical University, Moscow, May 12 - 13, 2022*:499-502. Russian.
19. Lahlou S, Le Bellu S, Boesen-Mariani S. Subjective evidence-based ethnography: method and applications. *Integrative Psychological and Behavioral Science*.2015;49(2):216-238.
20. Le Bellu S. Learning the Secrets of the Craft Through the Real-Time Experience of Experts: Capturing and Transferring Experts' Tacit Knowledge to Novices. *PISTES*.2016:18-1. Available: <http://pistes.revues.org/4685>.
21. Fauquet-Alekhine Ph. Subjective Ethnographic Protocol for Work Activity Analysis and Occupational Training Improvement. *British Journal of Applied Science & Technology*. 2016;12(5):1-16. Article no.BJAST.21632.
22. Suzuki S, Igarashi H, Kobayashi H, Yasuda T, Harashima F. Human adaptive mechatronics and human-system modelling. *International Journal of Advanced Robotic Systems*.2013;10(3),152:1-14.
23. Fauquet-Alekhine Ph. Risk assessment for Subjective Evidence-Based Ethnography applied in high risk environment. *Advances in Research*. 2016;6(2):1-13, Article no.AIR.21597.
24. Fauquet-Alekhine Ph, Le Bellu S, Buchet M, Berton J, Bouhours G, Daviet Fr, Granry JC, Lahlou S. Risk Assessment for Subjective Evidence-based Ethnography Applied in High Risk environment: Improved Protocol. *Advances in Research*.2018;16(3):1-15; Article no.AIR.43259.