

A Method to Reveal Workload Weak-Resilience-Signals at a Rail Control Post

Aron W. Siegel¹ and Jan Maarten Schraagen^{1,2}

¹ University of Twente/GW, P.O. Box 217, 7500 AE Enschede, The Netherlands
{A.W.Siegel, J.M.C.Schraagen}@UTwente.nl

² TNO Earth, Life, and Social Sciences, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands
Jan_Maarten.Schraagen@TNO.nl

Abstract. Reorganization of a rail control post may affect its ability to cope with unexpected disruptions. The term ‘resilience’, the ability to manage spare adaptive capacity when unexpected events occur, encapsulates this situation. This paper focuses on the workload adaptive capacity through a method for revealing workload weak-resilience-signals (WRS). Three different workload measurements are adapted to identify structural changes in workload. The first, executed cognitive task load, targets system activities. The second, integrated workload scale, is a subjective measure. The last, heart rate variability, identifies physiological arousal because of workload. An experiment is designed to identify the workload change and distribution across group members during disruptions. A newly defined Stretch, the reaction of the system to an external cluster-event, is used to reveal a workload WRS. The method is suitable for real-time usage and provides the means for the rail signaler to influence the system through his subjective workload perception.

Keywords: Resilience, weak resilience signal, WRS, objective and subjective Stretch, workload, rail operations, rail control post.

1 Introduction

Organizations restructure to improve their work efficiency. This efficiency step can, however, affect their spare, and sometimes hidden, adaptive capacity needed when an unexpected disruption occurs. In addition, this efficiency step can also affect the organization’s ability to manage this capacity. A socio-technical organization needs this ability to cope with disruptions, commonly referred to as ‘resilience’ [1]. As improved work efficiency may conflict with an organization’s resilience due to common resource demands, we need methods to identify this potential conflict. This paper deals with such a method and concentrates on the restructuring of a rail control post. A rail control post is responsible for a large area containing railway stations, controlled by rail signalers managing the traffic on the rail infrastructure. The post is 24/7 active with between 10 to 20 rail professionals, depending on the number of railway stations covered. A rail control post is an example of a socio-technical system due to the critical human-system interaction. Siegel & Schraagen [2] argued

that resilience in rail operations influences the rail system's operating state in three main areas: safety, performance (capacity and punctuality), and workload. In each of these areas, a weak resilience signal (WRS) may occur indicating a possible change of the system resilience, which needs further investigation to draw a solid conclusion. In this study, we focus on one area – workload. Changes in workload due to an organizational restructuring imply a change of the workload capacity needed during disruptions. This change is a reflection of a workload weak resilience signal (WRS), assuming that a decrease in the workload capacity lowers the ability of a socio-technical system to cope with disruptions and calamities, thus decreasing its resilience. The method described in this paper provides a means for investigating this assumption. Specifically, it aims at answering the following research questions: 1) Can workload measurements identify the human consequences of an organizational change, and 2) does such a change imply a possible impact on the resilience of the system? The activities are performed in a real operational environment where improvement or degradation may occur. Although a WRS is intended to signal a degradation, an improvement will be sufficient as well for demonstrating the concept, which will be relevant for a reversed restructuring.

In the following sections, we describe the setting of a rail control post when it is restructuring tasks among the rail signalers and follow this description with the approach of the method used. Afterwards, we elaborate on the method ingredients and their specifics about the setting. We finish the paper describing an experiment design and analysis approach, and a discussion.

2 Setting

The setting is a rail control post with m_{Post} workstations and n_{Post} rail signalers evaluating a new organization form to increase their performance. Each workstation, WS_j , is allocated to a set of railway stations and operated by one rail signaller, RS_i , who is responsible for all the workstations' aspects. These aspects are roughly divided into logistics and safety. The workstations are split into two groups. The first group, G_T , is the target group that will reorganize to improve its performance. The second group, G_R , is the reference group that will not reorganize throughout the testing period. All the n_{Post} rail signalers of the control post may be allocated to each of the groups and to each of its workstations. In group G_T there are m_T workstations WS_{Tj} and in group G_R there are m_R workstations WS_{Rj} . In addition, there is a calamity workstation, WS_{cal} , which is added to give support to the workstation being at the core of a calamity. The calamity workstation can be added to each group, G_T or G_R . The setting is depicted in Fig. 1:

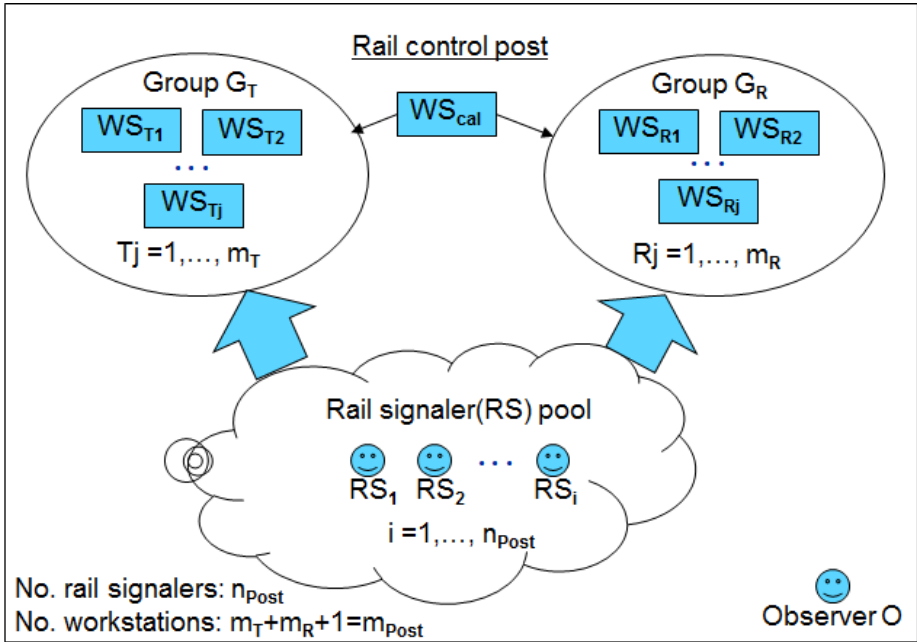


Fig. 1. Rail control-post setting with observer O

3 Approach

Pickup, Wilson, Nichols, & Smith [3] developed a conceptual framework of mental workload for railway signallers and differentiate among three types of mental load: (1) Imposed load, through the task characteristics; (2) Internal load and (3) perceived load, through the individual characteristics of the controller. We suggest using three different measurements to be able, to some extent, to differentiate among the influences of the mental load types: 1) external cognitive task load (XTL), 2) subjective workload, 3) physiological arousal created by workload. This is in line as well with Veltman [4] who argued that one needs performance data, subjective data and physiological data for a complete understanding of workload. Neerinx [5] modelled cognitive task load (CTL) in three dimensions: task complexity, task duration and task switching. We build upon this theory to compose XTL. We use the Integrated Workload Scale (IWS) [6] to measure subjective load and the extensively-researched heart rate variability (HRV) to identify physiological arousal due to workload change [7–12]. In the next sections, we elaborate these measures in more detail for the setting described above.

3.1 External Cognitive Task Load (XTL)

Rail signallers’ task execution can be divided into four main activities, which are measurable within the system: 1) monitoring (Mon), 2) plan mutations (Plan), 3)

manual actions (Man), and 4) communication (Com). We assume that monitoring is in proportion with automated activities executed by the system. This assumption refers to imposed task load, while in reality the rail controller can actually ignore it. Monitoring is measured by counting all the automated activities. These activities are counted in 5 minute base-slots, used throughout all the types of measurement for ease of comparison. We normalize these counts by dividing them by the maximum count (Mon_{max}) occurred throughout the test period. This causes the measurement to be normalized between 0 and 1. This same idea is applied to normalizing the plan mutations and the manual actions. Each of them are counted within the 5 minute base-slot and divided by the maximum count, $Plan_{max}$ and Man_{max} respectively, throughout the test period. The communication normalization is done differently. Communication is defined by the percentage of verbal exchanges over the phone, which is measurable, during the 5 minute base-slot. If the XTL is concerning a group, then 100% communication is defined by all members talking the whole 5 minutes.

The combination of these four normalized activities refers to task complexity as stated by Neerinx [5]. However, Neerinx used the Skill-Rule-Knowledge (SRK) model [13] to express task complexity by rating each task on its SRK cognition load level. We have chosen to describe the cognitive load of each of the four activities and track their identity throughout the whole process. Monitoring is about following the automated system. Updating the planning is a logistics task and coordination task with external parties, such as the train operators. Manual activities include direct operations on the infrastructure instead of the automated system. This demands a logistic understanding as well and needs good perception and insight of the infrastructure in the field. Telephone conversation has a large cognitive task load. In most cases, the signaller needs to understand the logistic and infrastructure situation outside while talking with the person on the phone, such as a train driver, and visualize the issues in the field. It is challenging to perform another task during demanding telephone conversations. In addition to these activities, task switching and task duration are two extra dimensions amplifying the workload. To estimate the number of task switches, we look at the task activations and count them in each time slot as long as they are activated, to reflect the task duration. In figure 2, we list the task activations imposed on a particular workstation or on the group R or T. These activations result in the activities discussed above and result in workload we are measuring by XTL, IWS and HRV. We divide the number of activations, occurring in the 5 minute base-slot, by the maximum activations occurring throughout the test period to achieve a normalized switching factor between 0 and 1. Task switching and duration are a cognitive add-on to the activity load. With the same activity load, 0 to n parallel task switches can occur, behaving like a cognitive amplifier to the activity load. We add one to the normalized switching factor to act as a cognitive amplifier by becoming a growth multiplier of the activity load. Graphically, the multiplication will show jumps attracting the attention needed for interpretation. Thus, the switching factor becomes:

$$K_{switch} = \frac{\text{number of activations in 5 min base-slot}}{\text{maximum number of activations in 5 min base-slot}} + 1 \quad (1)$$

We multiply the task switching factor with the added four normalized tasks to achieve a combined XTL number. This approach will create a number between 0 and 8 to be used as an overall graphical indication on the XTL magnitude and change. Maximum load due to task execution is: $4 \times 1 = 4$, multiplied by a maximum switching factor: $2 \times 4 = 8$. However, it is important to present all the components and their relationships separately, to understand the situation. We will discuss this further in the “Experiment Design and Analysis Approach” section.

The XTL calculations can be performed for the following units:

- Workstation WS_j ; XTL_{WS-j} , $j=1, \dots, m_{Post}$

$$XTL_{WS-j} = K_{switch-WS-j} \times \left(\frac{Mon_{WS-j}}{Mon_{max}} + \frac{Plan_{WS-j}}{Plan_{max}} + \frac{Man_{WS-j}}{Man_{max}} + Com_{WS-j} \right) \quad (2)$$

- Target group; XTL_{G-T} for WS_{Tj} , $Tj=1, \dots, m_T + 1$ (WS_{CAL})

$$XTL_{G-T} = \frac{1}{m_T + 1} \times \sum_{Tj=1}^{m_T+1} XTL_{WS-Tj} \quad (3)$$

- Reference group; XTL_{G-R} for WS_{Rj} , $Rj=1, \dots, m_R + 1$ (WS_{CAL})

$$XTL_{G-R} = \frac{1}{m_R + 1} \times \sum_{Rj=1}^{m_R+1} XTL_{WS-Rj} \quad (4)$$

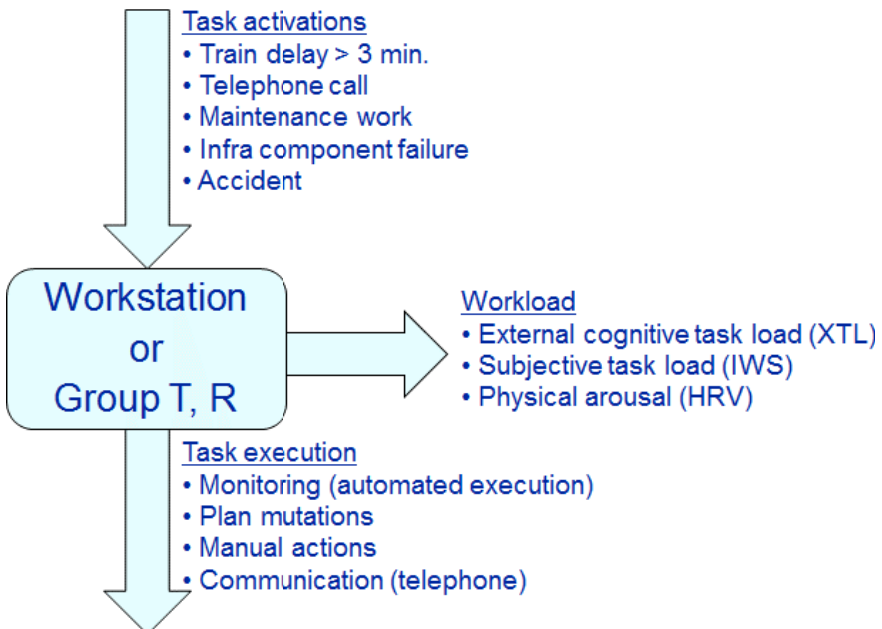


Fig. 2. Task flow per workstation or group

3.2 Integrated Workload Scale (IWS)

The Integrated Workload Scale [6] for a computer program runs on a laptop near each work station. The rail signaller RS_i , working at work station WS_j , is alerted every 5 minutes by a peripheral blinking rectangle, to rate his or her subjective workload. He or she is presented a 9 scale figure with the following text (in Dutch) (see figure 3):

1. Not demanding
2. Minimal effort
3. Some spare time
4. Moderate effort
5. Moderate pressure
6. Very busy
7. Extreme effort
8. Struggling to keep up
9. Work too demanding

The rail signaller has the possibility to add a comment to his or her rating and gets a graphic overview of his or her scoring (see Fig. 3).

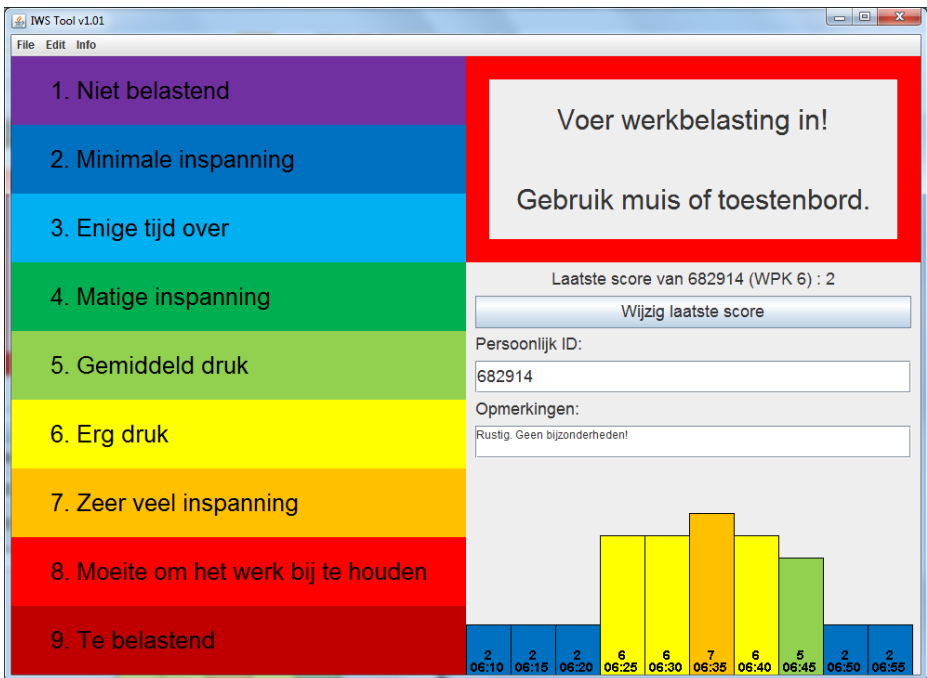


Fig. 3. IWS application screenshot (upper-right red rectangle blinks to draw attention)

The IWS_{WS_j,RS_i} is initially rated personally by rail signaller RS_i . We calibrate their scoring in order to combine it with the scoring of other signalers. Wilms & Zeilstra [14] have calibrated only with a quiet period rather than a rush hour situation. We

propose to extend the calibration to two situations: the quiet period and rush hour. Both situations are well defined by the normal planning and therefore suitable for calibration. Each signaler is asked about their rating when nothing special is happening between 10-11 AM, defining a quiet period, and between 5-6 PM, defining rush hour. We perform a linear transformation to the IWS calibrated values 2 and 5, matching the two situations, while maximizing it to 9. This results in a calibrated IWS^C_{ws-j} for every 5 minute base-slot (see figure 4).

The IWS calculated for each group is the average of all the workstations within the group. For the target group GT: $IWS^C_{GT} = \text{average}(IWS_{ws-j}, j=1, \dots, mT)$, and for the reference group GR: $IWS^C_{GR} = \text{average}(IWS_{ws-j}, j=1, \dots, mR)$.

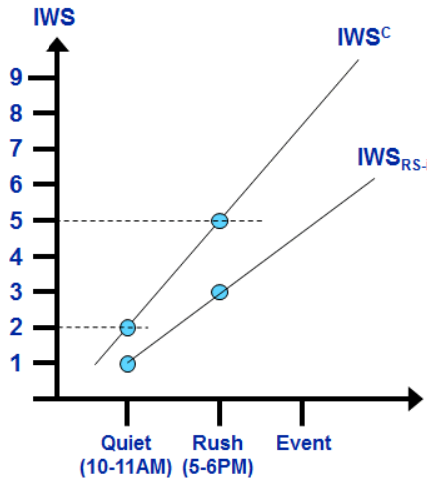


Fig. 4. IWS calibration

After calibration, it is possible to plot IWS together with XTL against the time for each workstation and group, since IWS has the same 5 minute base-slot as the XTL. We will use this relation in the “Experiment Design and Analysis Approach” section.

During the experiment an independent observer uses the IWS tool to rate $IWS_{O,G-T}$ for the whole target group and $IWS_{O,G-R}$ for the whole reference group. The observer uses the comment box to record relevant events.

3.3 Heart Rate Variability (HRV)

The heart rate variability is measured with a heartbeat device (Zephyr HxM BT), which is positioned on a breast strap and transfers its data to a laptop near each workstation. Every signaler wears the heartbeat device at the start of their work. The device sends continuous strings with recorded R-R beats in msec. The R-R interval is the time between subsequent R peaks of an ECG waveform. The information is stored on the laptop and post-processed the following day. HRV can be calculated in various ways roughly divided in time domain methods and frequency domain methods [11].

We use the most common occupational health method [12]: SDNN, the standard deviation (SD) of all normal-to-normal (NN) intervals, from the time domain and the low-high frequency (LF/HF) ratio in the frequency domain. We calculate both measures in the same 5 min base-slot used for the calculation of XTL and IWS. The SDNN is calculated by the standard deviation of R-R, the peak to peak interval, which is a very close measure of N-N, normal to normal interval. The LF/HF ratio is calculated through a discrete Fourier transform (DFT) of the first 256 measures, imposed by the DFT methodology using 2^n samples, in the 5 min base-slot. For a heartbeat rate of 80 bpm there are about 400 R-R samples in 5 minutes implying 256 measures to be the maximum integer power of two. The LF is the spectral integral of frequencies between 0.04 Hz and 0.15 Hz. The HF uses frequencies between 0.15 Hz and 0.4 Hz. We use these two HRV measures in the analysis.

4 Experiment Design and Analysis Approach

The control post, searching to optimize its processes, is restructuring around corridors. Until recently, all rail signalers were allocated individually to a few railway stations, being responsible for safety and the logistics through planning. This way of working is typical for the reference group. The target group, around one corridor, will divide safety and logistics responsibilities differently. One rail signaller will be responsible for all the planning activities within the corridor and the other rail signalers will only deal with safety. The experiment is designed to have two measurement periods of one week (Monday-Friday). The first period is a baseline measurement when no organizational changes have yet taken place. The second period is at least one month after the target group has reorganized and settled into the new setting. The plan is to record XTL, IWS and HRV 24 hours a day, but can be less due to practical reasons. Phenomena occurring in the target group before and after the change are likely to be caused by the organizational change but may also be caused by the measurements themselves [15]. We ignore the last possibility, since we are, for practical reasons, not able to perform an extra measurement set without a reorganization to show the measurement influence. Under these conditions we assume that phenomena, which do not occur in the reference group are due solely to the reorganization of the target group.

The analysis for each measurement period focuses at first separately on each of the workload methods: XTL, IWS and HRV. The external cognitive task load (XTL) with its 5 rail components – monitoring, planning, manual actions, communication and parallel tasking - is the main basis of estimating the workload. The XTL is the main basis since it is objectively measurable and represents facts derived from the system, while IWS and HRV have a more subjective character. The XTL information will be organized for each workstation in the 5 minute base-slots $t5$: $XTL_{ws-j}(t5)$, $j=1, \dots, m_{post}$. Afterwards, XTL is clustered for the groups T and R in the same base-slots: $XTL_{G-T}(t5)$, $XTL_{G-R}(t5)$. These XTL values are plotted against the time (see Fig. 5). The IWS is calibrated for each person and combined for each work station as well and plotted together with XTL, but with its own y-axis (see Fig. 5). IWS and

XTL behave differently. The XTL is more steady due to human estimation which changes gradually. The XTL is derived from system parameters, which causes a more wavering character. In order to relate the IWS and XTL measurements, a new term is introduced – Stretch. A Stretch is the accumulative workload effort during a period initially defined by IWS rising from a baseline until it returns to the baseline. The IWS-baseline is defined as the steady state IWS rating before and after a disruption. However, the activity in the system may have started earlier and ended later. Therefore, the starting moment of a Stretch is adjusted to the first XTL-minimum moment before the IWS rising. Similarly, the ending moment of a Stretch is adjusted to the first XTL-minimum moment after the IWS return. In other words, a Stretch is the reaction of the system to an external cluster-event. We use the term cluster-event, since more than one event may occur during a stretch. An Objective Stretch is the name of the area under XTL, since it is objectively measured. We name the area under IWS a Subjective Stretch due to its subjective IWS rating. The division of the Subjective Stretch by the Objective Stretch is called the Stretch ratio and is used to identify changes in the workload revealing a workload WRS. An average larger Stretch ratio during the period after the reorganization compared to the baseline period, indicates more subjective workload on similar external events. The Objective Stretch is used to identify an absolute workload growth, throughout a specific period like a day or a workweek. In the example of Fig. 6 we have plotted the Stretch ratio during a week for groups T & R. A significant change in the Stretch ratio, comparing two weeks, indicates a change in the relative workload and may be considered as a workload WRS.

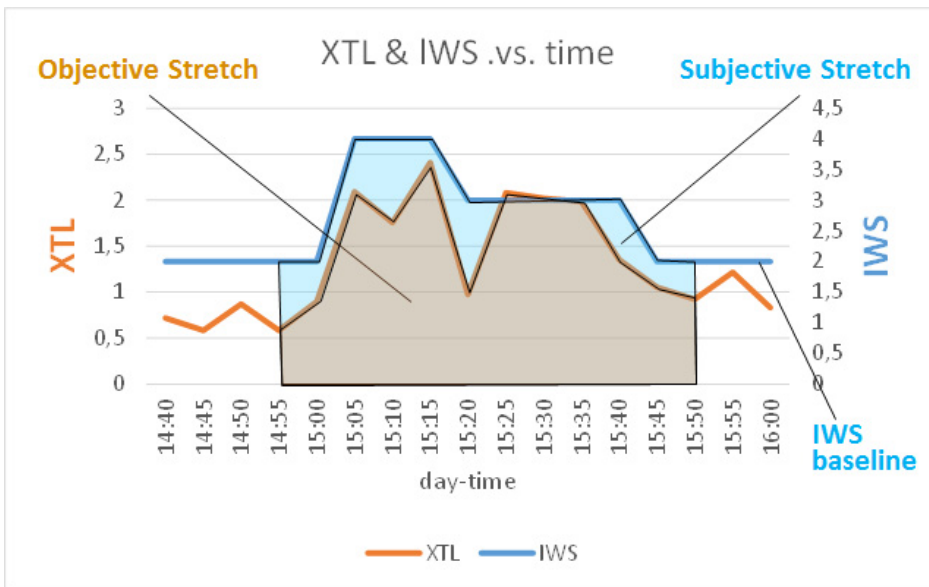


Fig. 5. Objective and Subjective Stretch

We use the HRV to validate the subjective IWS ratings. We expect the HRV to decline during a stretch as evidence that a growing IWS is an expression of a growing mental load.[12]

The XTL components are used to show the workload distribution among the members of each group R and T. The standard deviation (SD) presents the work distribution for each of the activities (for an example, see Fig. 6). The work distribution can be calculated for each stretch or for longer periods like a day or week and is calculated for each Group. The work distribution among group members, can help to explain the reasoning behind structural changes of Stretches and amplify the presence of a workload WRS.

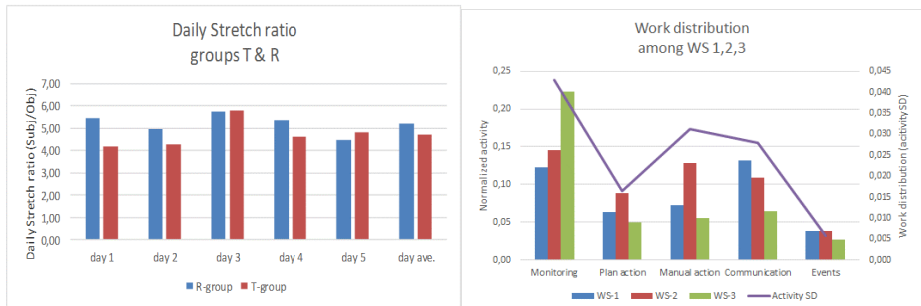


Fig. 6. Analysis examples: Daily Stretch ratio and work distribution

5 Discussion

The method described in this paper will be used for the experiment design worked out in the previous section. It provides the means to analyze the relation between workload and resilience, in particular, the method enables research on weak resilience signals (WRS), as explained in the introduction. The underlying assumption is that a decrease in the workload capacity lowers the ability of a socio-technical system to cope with disruptions and calamities. The decrease in workload is identified through Stretches, the reaction of the system to an external cluster-event, which has three types: objective, subjective and their ratio. The work distribution analysis ability provides deeper reasoning on Stretch changes. In future research, we plan to compare these results with other methods of identifying the resilience state of a team, such as questionnaires. This research should underpin in more depth the relation between workload and resilience.

Our method is applicable beyond the designed experiment and opens new dimensions for incorporating the human factor in decisions and control throughout disruption handling. The method has resulted in a real-time software tool, providing the possibility for the rail signaller to influence operations, through his or her subjective workload perception combined with objective measurements. We choose measurement variables that are available in real-time. This allows the possibility to provide feedback on the resilience state during operations and stimulates corrections

on the spot. In addition, we have designed the workload measurement balancing the objective and subjective. The rail signaller can express his or her personal opinion, show that to his or her environment and directly influence decisions and actions. We present the personal opinion in relation to the objectively measured workload and physiological arousal, providing a balanced view. Leveling the human state with the technical one is a change from today's situation in the rail organization, where the responsibility for human well-being is hidden in the lower level management on the work floor. By using this software tooling, it provides real-time insight of the human status to all levels in the organization. The impact of this tooling and methodology on the system resilience needs further research as well.

Acknowledgement. We thank Jaldert van der Werf for his development of the IWS and analysis software tooling, and his contribution to the experiment. We appreciate the constructive comments of the reviewers and thank Alfons Schaafsma guidance. This research was conducted within the RAILROAD project and is supported by ProRail and the Netherlands organization for scientific research (NWO) (under grant 438-12-306).

References

1. Hollnagel, E., Woods, D.D., Leveson, N. (eds.): Resilience engineering: concepts and percepts. Ashgate Publishing Limited, Hampshire (2006)
2. Siegel, A.W., Schraagen, J.M.: Developing resilience signals for the Dutch railway system. In: 5th Resilience Engineering Symposium (in press)
3. Pickup, L., Wilson, J.R., Nichols, S., Smith, S.: A conceptual framework of mental workload and the development of a self-supporting integrated workload scale for railway signallers. In: Wilson, J., Norris, B.J., Clarke, T., Mills, A. (eds.) Rail Human Factors, pp. 319–329. Ashgate, Surrey (2005)
4. Veltman, J.A., Gaillard, A.W.K.: Pilot workload evaluated with subjective and physiological measures. In: Brookhuis, K., Weikert, C., Moraal, J., de Waard, D. (eds.) Aging and Human Factors, pp. 107–128. University of Groningen, Haren (1996)
5. Neerinx, M.A.: Cognitive task load analysis: allocating tasks and designing support. In: Hollnagel, E. (ed.) Handbook of Cognitive Task Design, pp. 283–305. Lawrence Erlbaum Associates, Mahwah (2003)
6. Pickup, L., Wilson, J.R., Norris, B.J., Mitchell, L., Morrisroe, G.: The integrated workload scale (IWS): a new self-report tool to assess railway signaller workload. Appl. Ergon. 36, 681–693 (2005)
7. Billman, G.E.: Heart rate variability - a historical perspective. Front. Physiol. 2, 86 (2011)
8. Goedhart, A.D., van der Sluis, S., Houtveen, J.H., Willemsen, G., de Geus, E.J.C.: Comparison of time and frequency domain measures of RSA in ambulatory recordings. Psychophysiology 44, 203–215 (2007)
9. Hoover, A., Singh, A., Fishel-Brown, S., Muth, E.: Real-time detection of workload changes using heart rate variability. Biomed. Signal Process. Control 7, 333–341 (2012)
10. Jorna, P.G.A.M.: Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. Biol. Psychol. 34, 237–257 (1992)

11. Malik, M.: Heart Rate Variability. *Ann. Noninvasive Electrocardiol.* 1, 151–181 (1996)
12. Togo, F., Takahashi, M.: Heart rate variability in occupational health-a systematic review. *Ind. Health* 47, 589–602 (2009)
13. Rasmussen, J.: Risk management in a dynamic society: a modelling problem. *Saf. Sci.* 27, 183–213 (1997)
14. Wilms, M.S., Zeilstra, M.P.: Subjective mental workload of Dutch train dispatchers: Validation of IWS in a practical setting. In: 4th International Conference on Rail Human Factor, pp. 641–650 (2013)
15. Shadish, W.R., Cook, T.D., Campbell, D.T.: *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company, Boston (2002)