Investigating the Use of Wearable Technology to Support Safety in the Workplace

A thesis

submitted in fulfilment

of the requirements for the degree

of

Master of Philosophy in Computer Science

at

The University of Waikato

by

CHRISTOPHER JOHN GILDER GRIFFITHS

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2018

Abstract

Many workplaces in New Zealand can be described as hazardous. That is, the nature of the work and/or workplace, or the combination of the two, can lead to situations where workers may be at risk of workplace accidents. One contributor to such accidents is worker fatigue, which is the result of the nature and intensity level of the work they are undertaking. This can be exacerbated by factors such as the length of the working day, shift work and roles that require high levels of concentration. Most existing risk minimization processes rely on self-reporting methodologies and health and safety procedures; neither of these are necessarily the most effective methods for dealing with workers in hazardous jobs and work environments.

Wearable technology which collects physiological data, such as step and heart rates, as an individual performs workplace tasks has been proposed as a possible solution. While wearable devices are minimally intrusive to the individual and so can be used throughout the working day it is unclear how suitable they are for in-situ measurements in real-world work scenarios. In this work, we describe a series of studies conducted with New Zealand forestry workers and present an analysis of the data gathered to consider the suitability of the collection methods as well as the suitability of the data itself as a method to identify fatigue and reduce risk in the workplace.

Introduction

Judy Bowen, March 2018

This collection of work is based on research and writing undertaken by Christopher Griffiths during his enrolment as a PhD candidate in the Department of Computer Science from February 1st 2016 to September 2017. It begins with introductory material written by Chris as part of his PhD research proposal, and as such describes the domain of this studies and his intended research plan.

Following this introduction are three publications arising from his work. The first of these is "Investigating Wearable Technology for Fatigue Identification in the Workplace" which was accepted for publication and presentation by IFIP TC13 Conference on Human-Computer Interaction (INTERACT) 2017 [1]. This is an 'A' ranked peer-reviewed conference (CORE rankings). Chris was the primary author on this paper which described his preliminary investigations into the use of wearable technology to capture physiological and psychological data under differing (physical and mental) loading types. It summarised the results of studies conducted during the first 6 months of his enrolment.

The second publication is a journal paper "Investigating Real-Time Monitoring of Fatigue Indicators of New Zealand Forestry Workers" accepted by Accident Analysis and Prevention for their special issue: Managing Fatigue to Improve Safety, Wellness, and Effectiveness [2]. Most of the writing for this paper was done by myself and Annike Hinze, but all of the data comes from Chris's first in-situ study with forestry workers with insights drawn from his analysis.

The third publication, "Personal Data Collection in the Workplace: Ethical and Technical Challenges" [3] was accepted for publication and presentation by British HCI Conference 2017. This peer-reviewed conference is ranked as a National conference by CORE and as an A ranked conference by ERA. The work in this paper describes some of the wider research project that Chris's PhD work contributes to and as such draws on findings from the first 8 months of his PhD research.

We are currently working through several data-sets from studies conducted by Chris between February 2017 and September 2017. The largest of these include another in-situ study with forestry workers which investigate correlations between heart-rate variability (HRV), reaction time and task load measurements. The others are smaller ad-hoc studies and include an investigation into HRV and reaction time of night-shift workers and experiments into differing driving conditions over long periods and how these affect both HRV and reaction time and how they are mitigated by recovery time. We are in the process of writing a journal paper which will include the forestry study data, and anticipate that the rest of the data will contribute to a conference publication.

Research Introduction by Christopher Griffiths

The New Zealand workplace is fraught with potential harm, high injury rates have seen large numbers of claims made to ACC, with some 226,100 reported in 2014 alone [4]. More importantly,44 workplace fatalities wer reported for the same year. Recent reforms to health and safety legislation have been implemented by way of the Health and Safety Reform Bill [5] designed to address the high levels of accident and injury rates. Amendments to the Bill have seen the shifting of focus to a more employee-centred approach; stating the need for effective worker engagement and participation, with the anticipation of cutting the current injury rate by twenty five percent by 2020.

Risk within the workplace is defined as anything that has the potential to cause harm, and places emphasis on both employer and employee participation in addressing workplace risk. Current risk assessment practices require the identification of risk, impact of identified risk, and procedures to eliminate, isolate or minimise the risk. This process is usually undertaken on a task by task basis by a company nominated Health and Safety representative and is generally based on the practical aspects of the nature of the task, in conjunction with legislative guidelines. This method has many pitfalls as it is subjective by nature, the assessor of a task may have a holistic view of the energy requirements of a worker required for task completion rather than the true energy requirements (amount of energy used for task completion). Furthermore, the process can result in a generic view of a task and its physical requirements, whereas in reality this is unique to the individual performing the task. It must also be remembered that the effects of energy expenditure are cumulative; frequent task repetition adds to the daily total of energy expenditure and can lead to excessive levels of fatigue. Such high levels of fatigue can negatively impact employee well-being, increase the risk of accident, and negatively impact performance [6, 7, 8].

The Department of Labour recognise fatigue as a contributor to workplace safety and reference their 2003 document 'Healthy work managing stress and fatigue in the workplace' when providing identification and preventative measures in dealing with fatigue. Recommended measures for identification are predominantly self reported; employees complete questionnaire's designed to identify fatigue by way of how they feel. Results of these questionnaires are used by health and safety representatives to both measure and minimise the impact of fatigue within the workplace.

Current technologies exist that facilitate the automatic collection of activity levels. Targeted as fitness trackers these devices are minimally intrusive to the user and operate autonomously. The current generation of fitness trackers [9] are typically wrist worn devices connecting via Bluetooth to a smart phone device to view the collected data. Devices such as those manufactured by Fitbit, Jawbone, Polar etc. provide additional data analysis which can be viewed through a web portal. It is my intention to use the physiological data supplied by these devices, and other similar non-intrusive data gathering methods, in conjunction with targeted subjective workload assessments to assess the impact of task and role based activity on performance. Quantitative physiological data collected

3

automatically as individuals perform their daily duties will facilitate a better understanding of task/role energy requirements. Health and Safety professionals provided with this information will be able to better identify workplace risk and therefore better address risk mitigation techniques. It is hoped that a more accurate risk assessment that identifies task loading can both minimise risk and increase productivity. Furthermore, by identifying early signs of impairment, timely intervention can minimise the risk of harm within the workplace.

Study Domain

This work builds on the studies and relationships from work undertaken during my MSc, in which non-intrusive monitoring technologies were investigated for use with forestry employees. The forestry industry is one of the most dangerous occupations in New Zealand, employees being 15 times more likely to suffer a workplace injury compared to other NZ based industries [5].The nature of the work is physically demanding, employees can expend as much energy throughout the course of their day as an athlete would in running a marathon [10]. Such a high level of energy expenditure combined with hazardous working environments can substantially contribute to the levels of risk employees are exposed to. Worksites are generally remote, requiring the working day be lengthened to incorporate travel time. This extension to the working day combined with the physically demanding workload can cause issues with fatigue in workers and adversely impact safe working practices [11]and be a contributor to the high numbers of incidents within the industry.

Background

There is limited information on physiological monitoring of forestry employees and most information to date focuses on heart rate and aerobic capacity [12, 13, 14, 15]. Similar limited information is present for fatigue investigation in forestry workers and existing literature is predominantly based on subjective reporting [16, 17]. This method of data collection has many pitfalls as it is subjective by nature and susceptible to response bias; individuals tend to agree with questions [18].

As part of my Masters degree I investigated tools and methods for data capture of forestry workers. This involved collecting data on physiological variables (activity) and investigating their relationship with performance (reaction time). Roles investigated were varied, tree felling and log making roles required large amounts of physical activity, conversely machine operator roles were of a more sedentary nature. This highlighted a problem for determination of activity in roles of a sedentary nature. These roles are generally more biased to cognitive rather than physical activity as they are using machinery for loading trucks, hauling logs, stacking logs within confined work areas, which requires high levels of attention.

Measurement of activity in cognitive biased roles requires additional monitoring technologies and techniques to those originally used. This proposal is

4

designed to investigate real-time data capture of both physical and mental activity. It is hoped that data collected in conjunction with reaction time testing can be used as indicators of increasing fatigue levels within forestry workers. In my work I will consider the nature of fatigue, both generally and within a workplace context and suggest physiological indicators negatively impacted by fatigue. Furthermore I will build on previous works undertaken on the identification of fatigue by use of autonomic activity indicators.

Fatigue

Fatigue, which is often misrepresented as a feeling of sleepiness, can be classified into two general types; mental fatigue that affects an individual's cognitive processes and physical fatigue that affects an individual's ability to maintain physical actions. Both types of fatigue are cumulative by nature in that the more a task is performed the level of fatigue generated by performing the task increases. Both types of fatigue lessen after a period of rest. Although quantifiable by such indicators as reduction of muscle strength for physical fatigue, and slower response times for mental fatigue, it is a subjective physiological state experienced by individuals as a result of either physical or mental exertion. The levels of fatigue experienced are determined by the individual's perception of these physiological changes. The nature of this physiological state can be best explained as:
"The taking possession of the mind by a sense of lassitude. It is a reduction in the capacity and desire to react. It is characterised by tiredness and an aversion to the continuation of goal-directed work. It is accompanied by a strong desire for rest through the cessation of ongoing activity." [19]
This desire for rest is an evolutionary physiological process designed to protect an individual from harm. Fatigue provides a barrier designed to leave enough energy available to an individual should unexpected threats occur and rapid reaction be required. This barrier is able to be overcome by an individual's determination or desire to complete a goal. This can be thought of as a 'return on investment', with energy expenditure being the invested variable. For example, does the reward from the goal completion justify the energy expenditure?

Fatigue is classified into two separate types, physical and mental, each contributing to the overall fatigue levels experienced by an individual. When examining these fatigue types within a workplace context we can, in most cases, identify the major contributor to the overall fatigue levels experienced. Roles requiring high levels of task repetition, or forceful extensions of muscle groups are inclined to see individuals at risk from physical fatigue, conversely roles requiring large amounts of sustained attention are more inclined to see individuals at risk from mental fatigue.

The impact of high fatigue levels

There have been many investigations into the impact of high levels of fatigue, both physical and mental. Typical findings are that there is a decline in mus-

cle performance on repeated exertions and a decline in performance as load duration increases [20, 7]. Furthermore, Gregory et al. investigating muscle activation patterns during highly physically demanding tasks identified significant differences (p ¡ 0.05) between muscle groups in pre- and post-test conditions[21]. Similar traits exist for mental fatigue, the higher the loading the more impacting the fatigue. In a study Linden et al. found:

"Willingness to exert effort on the experimental tasks and to do ones best on these tasks, as measured directly after the manipulation, was significantly lower for the fatigued participants than for the non-fatigued participants." [22]

As such, a higher the level of mental fatigue experienced by an individual can result in lower productivity levels. This conclusion was reinforced in [23], where the authors found "fatigue is usually related to a loss of efficiency and disinclination to work."

subsubsection*Fatigue in the workplace Work is seen as a 'must do' rather than a 'desire to do' activity. One does not encounter fatigue to the same extent as and when goal completion is not imposed. Hockey states:

"The widespread interpretation of fatigue as a negative consequence of work may be true for externally imposed goals; meaningful or self-initiated work is rarely tiring and often invigorating."[24]

Work related goals, by their nature, impose stressors on an individual. Work is a reward for activity based concept being reliant on imposed expectations for goal completion (performance). These stressors are multi-faceted, consisting of both internal (workplace) and external (social) elements. Identifying the causal nature of workplace fatigue can lessen the risks encountered in the workplace, however it must be remembered that external stressors are also contributors. Limited or poor quality sleep and insufficient recovery time also impact performance. While investigating sleep quality Akerstedt, et al. found that disturbed sleep is an important predictor of fatigue. They also state that fatigue is closely related to high work demands, immersion in work and disturbed sleep [8].

Recovery time is also a contributor to fatigue, in a study into cumulative fatigue Pichot et al. used measurements of the autonomic nervous system index using ECG readings taken nocturnally from six male French rubbish collectors. They found a progressively increasing resting heart rate throughout the working week with a reduction occurring as the recovery period between work periods extended [25].

Monitoring technologies

Fitness monitoring existing previously within the realms of athletes and the military has now become commonplace, with individuals collecting physiological data for their own personal use. Developed from the simple pedometer, the later generations of these devices use combinations of three axis accelerometers to determine cadence. Proprietary algorithms, based on height and weight are used to calculate calorific burn and distance travelled. Further developments have seen the incorporation of technology similar to pulse oximetry where LEDs illuminate an area of the skin (usually the wrist) and reflected light is measured

by a photo electric cell. This facilitates the collection of heart-rate data, however this methodology shows variation when compared to measurements taken using chest strap monitoring systems. Sleep quantity and quality can also be determined based on time stamp and movement data. Interruptions in sleep are recorded when the device registers movement data at times during the night, conversely long periods of inactivity are defined as sleep periods.

The use of these devices is becoming more commonplace in experimental studies with automated physiological data collection by these devices facilitating research into areas that lie beyond providing goal-centred information to users. Recent research has seen the use of commercially available devices for investigation into areas beyond that of activity monitoring. Zambotti et al. used the FitBit Charge HR to investigate cardiac functioning of adolescents during sleep. They found that the device showed good agreement with polysomnography and electrocardiography in measuring sleep and heart rate whilst sleeping [26].

Current iterations of commercially available activity monitors (e.g. Microsoft Band 2) facilitate the collection of additional metrics such as galvanic skin response and body temperature. Updates to sensor technologies also allow for improved heart rate determination using pulse oximetry. Cormack et al. investigated the suitability of these trackers in monitoring cognition, mood and behaviour, suggesting
"these devices may be used for daily cognitive testing, complementing periodic in-person assessment in clinical research or interventions."[9]
Driven by sales, refinement of these devices is constant with additional metrics, improved accuracy, and data granularity all being added to improve market share. Perpetual development cycles add to the value of commercially available activity monitors in experimental research, facilitating data collection in domains where traditional methodologies may not be practical.

Measurable Indicators of activity

When examining fatigue indicators we are faced with a choice, do we use measured activity as a determinant or do we focus on the body's response to activity? I suggest that regulatory change in response to activity is a preferred measurement technique when collecting data on mental loading. Physical loading by way of movement, heart rate and calorific burn rates have been successfully employed in the determination of activity in our proposed domain [27] as such I suggest these as indicators of physical loading.

However, it must be remembered that everyone is unique, although the response to activity is universal; faster heart rates, lower beat to beat intervals of heart beat and differing suppression levels in brainwave activity. All these changes occur automatically in response to external stressors, and are governed by the autonomic nervous system, a brief overview of which is presented next.

The autonomic nervous system

Part of the peripheral nervous system, the autonomic system controls the involuntary processes in the body, automatically regulated without conscious thought. Heart beat, digestion perspiration etc. are all functions regulated by this system, activation rates being controlled in response to physiological stressors. Split into two parts the sympathetic nervous system is responsible for the fight or flight grouping of processes such as increasing respiration and heart rate. Complementary to this is the parasympathetic nervous system responsible for the rest or digest processes such as decreasing respiration and heart rate. Changes in the balance between these two systems occur when the body experiences the addition or removal of physiological stressors. Imbalance between these two systems can be used to determine the impact of external stressors, and as such can be used as indicators of fatigue [28, 29].

Human performance measures

Seen as the response variable, human performance will be impacted by the variation in activity one performs. Reaction time has long been a stalwart for the measurement of human performance. Developed by Galton in 1889, the simple reaction time test is a measure of the length of time it takes an individual to respond to either a visual or auditory stimulus. Response is usually the press of a button on presentation of the stimulus. The test itself is usually comprised of a predetermined number of sub trials with the result being presented as an average of these sub trials. It is used in the testing of general alertness and motor speed of an individual with uncertainty being introduced by a variable time period between individual trials.

Being somewhat similar to simple reaction time, choice reaction times measures the response time of an individual when presented with a visual stimulus, although this test requires the subject to make a choice. These choices are usually presented visually and require a certain response to a stimulus. A prime example being that a user is presented with a visual representation of an arrow pointing either left or right and they must select the corresponding button on a keypad. The act of having to choose the correct response requires more cognitive processing than simple reaction time and can indicate delays that may be caused by impairment such as fatigue. Tests are usually comprised of a predetermined number of sub trials which require selection of the correct response to a stimulus. Uncertainty is added by the requirement to choose, and a variable time between presentations of the stimulus.

Both of these methodologies have been used many times in the investigation of activity in human performance, with most of these studies reporting increasing reaction times as a result of cognitive or physical loading [30, 31, 32]. Similar investigations have been undertaken within workplace contexts, Baulk et al. investigated fatigue levels experienced by individuals with varying workloads. Using a similar methodology to that proposed for this research they identified both slower reaction times and an increasing trend of subjective fatigue across

work periods [33].

Research Questions

Based on the identified problem and the background literature, my aim is to address the following research questions:

- What are the physiological variables that we should measure as determinants of activity?

- What are the technological solutions we can use to measure them?

- How do we infer meaning from the data?

- What do we do with the data?

I will answer these through in-situ studies with forestry workers as well as additional experiments and studies as required. Ultimately we aim to discover what is measurable and what is meaningful so that we can incorporate the measured data into real-time solutions for worker safety.

References

[1] Griffiths, C., Bowen, J., Hinze, A.: Investigating wearable technology for fatigue identification in the workplace. In Bernhaupt, R., Dalvi, G., Joshi, A., K. Balkrishan, D., O'Neill, J., Winckler, M., eds.: Human-Computer Interaction - INTERACT 2017, Cham, Springer International Publishing (2017) 370–380

[2] Bowen, J., Hinze, A., Griffiths, C.: Investigating real-time monitoring of fatigue indicators of new zealand forestry workers. Accident Analysis and Prevention (2017)

[3] Bowen, J., Hinze, A., Griffiths, C., Kumar, V., Bainbridge, D.: Personal data collection in the workplace: ethical and technical challenges. In: British Human Computer Interaction Conference, (BHCI). (2017) 10 pages

[4] Statistics New Zealand: Business demography data 2003-2013 (2014)

[5] Ministry for Business and Innovation: Independent forestry safety review an agenda for change in the forestry sector (2014)

[6] Graham, T., Hibbert, E., Sathasivam, P.: Metabolic and exercise endurance effects of coffee and caffeine ingestion. Journal of Applied Physiology **85**(3) (1998) 883–889

[7] Kumar, S., Amell, T.and Narayan, Y., Prasad, N.: Indices of muscle fatigue. In: Proceeding of Human Factors and Ergonomics Society Annual Meeting. Volume 13. (2002)

[8] Åkerstedt, T., Fredlund, P., Gillberg, M., Jansson, B.: Work load and work hours in relation to disturbed sleep and fatigue in a large representative sample. Journal of psychosomatic research **53**(1) (2002) 585–588

[9] Cormack, F.and Taptiklis, N., Barnett, J., King, J., Fehnert, B.: High frequency monitoring of cognition, mood and behaviour using commercially available wearable devices. The journal of the Alzheimer's Association **12**(7) (2016) 159

[10] Kirk, P.M.: Reducing the impacts of fatigue on forest workers. **21**(3) (1996) 1 – 6

[11] Spurgeon, A., Harrington, J.M., Cooper, C.L.: Health and safety problems associated with long working hours: a review of the current position. Occupational and Environmental Medicine **54**(6) (1997) 367–375

[12] CKirk, P.M., Sullman, M.J.: Heart rate strain in cable hauler choker setters in new zealand logging operations. Applied Ergonomics **32**(4) (2001) 389–298

[13] Scott, P., Christie, C.: An indirect method to assess the energy expenditure of manual labourers in situ. South African Journal of Science **100**(11-12) (2004) 694–698

[14] Toupin, D., LeBel, L., Dubeau, D., Imbeau, D., Bouthillier, L.: Measuring the productivity and physical workload of brushcutters within the context of a production-based pay system. Forest Policy and Economics **9**(8) (2007) 1046 – 1055

[15] Erhan Çaliskan, S.c.: An assessment of physiological workload of forest workers in felling operations. African Journal of Biotechnology **9**(35) (2010) 5651–5658

[16] Lilley, R., Feyer, A., Kirk, P., Gander, P.: A survey of forest workers in new zealand. do hours of work, rest, and recovery play a role in accidents and injury? Journal of Safety Research **33**(1) (2002) 53–71

[17] Yoshimura, T., Acar, H.H.: Occupational safety and health conditions of forestry workers in turkey. Journal of Forest Research **9**(3) (Aug 2004) 225–232

[18] Morrel-Samuels, P.: Getting the truth into workplace surveys. Harvard Business Review **80**(2) (2002) 111–118

[19] Matthews, G., Hancock, P., Desmond, P.: The Handbook of Operator Fatigue. CRC Press, London (2012)

[20] Allen, D.G., Westerblad, H.: Role of phosphate and calcium stores in muscle fatigue. The Journal of Physiology **536**(3) (2001) 657–665

[21] Gregory, D.E., Narula, S., Howarth, S.J., Russell, C., Callaghan, J.P.: The effect of fatigue on trunk muscle activation patterns and spine postures during simulated firefighting tasks. Ergonomics **51**(7) (2008) 1032–1041

[22] van der Linden, D., Frese, M., Meijman, T.F.: Mental fatigue and the control of cognitive processes: effects on perseveration and planning. Acta Psychologica **113**(1) (2003) 45 – 65

[23] Murata, A., Uetake, A., Takasawa, Y.: Evaluation of mental fatigue using feature parameter extracted from event-related potential. International journal of industrial ergonomics **35**(8) (2005) 761–770

[24] Hockey, R.: The Psychology of Fatigue, Work, Effort and Control. University of Sheffield (2013)

[25] Pichot, V., Bourin, E., Roche, F., Garet, M., Gaspoz, J.M., Duverney, D., Antoniadis, A., Lacour, J.R., Barthélémy, J.C.: Quantification of cumulated physical fatigue at the workplace. Pflügers Archiv **445**(2) (2002) 267–272

[26] de Zambotti, M., Baker, F.C., Willoughby, A.R., Godino, J.G., Wing, D., Patrick, K., Colrain, I.M.: Measures of sleep and cardiac functioning during sleep using a multi-sensory commercially-available wristband in adolescents. Physiology Behavior **158** (2016) 143 – 149

[27] Bowen, J., Hinze, A., Cunningham, S.J., Parker, R.: Evaluating low-cost activity trackers for use in large-scale data gathering of forestry workers. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. OzCHI '15, New York, NY, USA, ACM (2015) 474–482

[28] Mascord, D.J., Heath, R.A.: Behavioral and physiological indices of fatigue in a visual tracking task. Journal of safety research **23**(1) (1992) 19–25

[29] Collet, C., Averty, P., Dittmar, A.: Autonomic nervous system and subjective ratings of strain in air-traffic control. Applied ergonomics **40**(1) (2009) 23–32

[30] Brisswalter, J., Arcelin, R., Audiffren, M., Delignières, D.: Influence of physical exercise on simple reaction time: effect of physical fitness. Perceptual and Motor Skills Journal **85** (1997) 1019–1027

[31] Langner, R., Steinborn, M.B., Chatterjee, A., Sturm, W., Willmes, K.: Mental fatigue and temporal preparation in simple reaction-time performance. Acta Psychologica **133**(1) (2010) 64 – 72

[32] Sabzi, A.H.: The effect of different fatigue protocols on choice reaction time. Middle-East Journal of Scientific Research **12** (2012)

[33] Baulk, S.D., Kandelaars, K.J., Lamond, N., Roach, G.D., Dawson, D.,
Fletcher, A.: Does variation in workload affect fatigue in a regular 12 hour
shift system? Sleep and Biological Rhythms **5**(1) (2007) 74–77

Investigating Wearable Technology for Fatigue Identification in the Workplace

Christopher Griffiths, Judy Bowen[✉], and Annika Hinze

The University of Waikato, Hamilton, New Zealand
cjgg1@students.waikato.ac.nz, {jbowen,hinze}@waikato.ac.nz

Abstract. Fatigue has been identified as a significant contributor to workplace accident rates. However, risk minimisation is a process largely based on self-reporting methodologies, which are not suitable for fatigue identification in high risk industries. Wearable technology which is capable of collecting physiological data such as step and heart rates as an individual performs workplace tasks has been proposed as a possible solution. Such devices are minimally intrusive to the individual and so can be used throughout the working day. Much is promised by the providers of such technology, but it is unclear how suitable it is for in-situ measurements in real-world work scenarios. To investigate this, we performed a series of studies designed to capture physiological and psychological data under differing (physical and mental) loading types with the intention of finding out how suitable such equipment is. Using reaction time (simple and choice) as a measure of performance we found similar correlations exist between loading duration and our measured indicators as those found in large-scale laboratory studies using state of the art equipment. Our results suggest that commercially available activity monitors are capable of collecting meaningful data in workplaces and are, therefore, worth investigating further for this purpose.

1 Introduction

Fatigue is, by nature, cumulative, and is influenced by many variables such as activity, time of day, sleep levels and social pressures. The impact of high fatigue levels, especially in high-risk workplaces, can lead to increased risk for employees, and is seen as a large contributor to workplace accident rates [12,19,29]. Risk assessment processes include attempts to assess the impact of workplace activities on fatigue levels. Typically these use qualitative self-reporting methodologies rather than quantitative data measurements. This type of data collection in workplace contexts is susceptible to response bias [24] and can result in a generalised view of risk. However the actual impact of activity upon fatigue is individualised. For example, tasks performed by a young person impacts fatigue levels to a lesser extent than tasks performed by an older person [2].

Many studies investigating the impact of activity upon performance have been undertaken. Most identify that an individual's performance is negatively

impacted by activity, both physical and mental [9,16,23,27]. Each of these studies collected data using different methods, but the majority are conducted in a laboratory setting with large numbers of participants. In real-world work environments, individuals encounter naturally occurring stressors that may not be observed in laboratory experiments. Similarly the longer-term nature of fatiguing activities in the workplace and the limited number of participants who can be measured in in-situ studies may make it hard to reproduce known results.

The rise in popularity of wearable devices has provided additional tools for quantification of individualised activity levels. These devices are designed to operate autonomously and enable the real-time collection of quantifiable data in the field throughout the day. Physiological markers such as step and heart rates can be used to determine workload intensities, especially in physically-biased roles. Conversely, changes in the balance of the autonomic nervous system can be used to quantify cognitive loading. However, it is unclear how accurate these measures are when using commercial products designed primarily for the personal user. As a first investigation into this, we performed a series of single-person studies designed to see if we could reproduce known correlations between activity (both mental and physical) and response times. If we find similar correlations can be identified by these devices then it is worth further investigating their use in real-world work environments.

In this paper, we present the results of these initial studies investigating the use of low cost commercially available devices as a means of collecting data that has sufficient accuracy and granularity to identify physiological changes associated with mental and physical activity types which may indicate fatigue. We examine the suitability of such devices for use in the field, specifically for gathering data and monitoring of forestry workers throughout their working day.

1.1   Motivation

The forestry industry in New Zealand has a poor Health and Safety record with some 12,921 active Accident Compensation Corporation (ACC) claims between 2008 and 2013. More importantly the number of reported fatalities for the same period is 32. Such high levels of fatalities are concerning to the industry with reforms being planned to address safety of employees. Suggestions based on experimental data are limited, and difficult to source. Currently, the only practicable solution is seen as increasing the level of mechanisation resulting in removal of the employee from the worksite.

Recently (2014), the large numbers of incidents reported prompted the Independent Forestry Safety Review [1], designed to investigate factors that impact on health and safety within the forestry industry and to provide guidelines designed to minimize the number of incidents. It was identified that the forestry industry is one of the most dangerous occupations in New Zealand. Employees are 15 times more likely to suffer a workplace injury compared to other NZ based industries.

The physical nature of the work, long working days and tasks requiring high levels of concentration can all exacerbate the impact of fatigue [13,28] with high

demand tasks requiring more energy to complete. In workplaces with high task demands employees may experience the effects of fatigue earlier. There is also a recognition that the lack of welfare facilities in the forestry environment may be an additional contributor to fatigue.

1.2   Measurements

There are numerous physiological indicators that can be used to quantify activity. Step and heart rates have been successfully used to quantify physical activity [8], whereas changes in the balance of the autonomic nervous system have been used as an indicator of cognitive activity [25].

Currently seen as the gold standard for measurement of these variables are step counting for ambulatory activity and, electrocardiography for heart rate data collection. However, we must remember that in our proposed domain the devices we choose need to be minimally intrusive and capture data autonomously. Before we can assess suitable devices for the data collection we must first consider what the appropriate data to collect is. We discuss proposed identifiers of activity next, these will then inform subsequent choices of appropriate apparatus for data capture.

Step Rate: The use of pedometers to measure an individual's daily activity has been used many times in differing domains. Designed to capture accelerations of the hip during gait cycles they count the number of steps taken by an individual over a given time period. Using the number of steps taken we can gain an insight into how role-based activity may contribute to excessive fatigue levels. This method has been used successfully to differentiate activity levels by role types in forestry harvesting operations [17] with large differences between manual and mechanised activity types being identified.

Heart Rate: Heart rate data captured throughout the course of a work period can be an used as an indicator of task demand. Higher heart rates typically accompany higher periods of physical activity as higher levels of oxygenated blood are required due to increased physiological demand. Increases in heart rate are individualized with demand being dependent on such criteria as age and fitness levels. However, maximal heart rate can be calculated from the general formula equation Maximal Heart Rate = 220 minus Age [20]. The resultant figure can then be used in conjunction with resting heart rate to determine periods of high and low activity where activity is defined as deviation from the resting heart rate.

Heart Rate Variability (HRV): Heart rate variability is the time interval between successive heart beats. Shorter periods are indicative of active loading while longer time periods are indicative of rest. There are a large number of variables that can be extracted from collected data each providing an insight into autonomic nervous system activity. For our studies we use the low frequency/high frequency ratio of the power spectrum density of the heart. This variable has been found to correlate well with mental activity across differing levels of cognitive load [7, 10, 26].

Workload Intensity: Workload intensity, or how hard an individual perceives their workload can be an indicator of increasing fatigue levels. Monotonous or repetitive tasks have been identified as a contributor to motivational levels [22] with individual performance slowing as motivational levels decrease [34]. The Task Load Intensity Tool (TLX) developed at the Ames Research Center is designed as a self-reporting estimator of how difficult an individual perceives their workload. Using a six point scale it provides a workload score for workplace activities [18]. Increasing workload scores for tasks where the only change is duration can indicate increasing workplace fatigue [3].

Performance: Reaction time (both simple and choice) has been used many times to measure individual performance. It is a measure of the elapsed time between the presentation of a given stimulus and the participant's response. In simple reaction time the user responds to one stimulus whereas in choice reaction time the user must identify the correct response from a set of choices. Reaction time has been found to deteriorate with increasing fatigue [5,13,21] furthermore the time difference between choice and simple reaction time can give insights into the speed of mental processing.

2   State of the Art Methods and Tools

When considering our methodologies for collecting data we must also consider the accuracy of our proposed devices. The most accurate measure for step counting is manual counting. However, this is impractical for many purposes and in particular for our chosen domain - one cannot manually count steps in the workplace. The Yamax range of pedometers are widely regarded as the most accurate equipment for automatic step counting [32] and have been used in many studies investigating step rates [6,8,30]. When comparing the accuracy of our chosen device (Fitbit Charge HR) many studies have found good agreement between actual and recorded steps taken [14,15].

Real-time data collection of cardiovascular activity in the workplace presents additional challenges. This type of data collection is typically done using an electrocardiograph and is undertaken at hospitals or dedicated research laboratories. The equipment can be bulky and cumbersome with an individual having wired sensors placed on the body. As previously discussed, our proposed domain is forestry operations in which traditional ECG measurement is not practical.

Wearable devices exist that are designed to capture data in free living activities. These devices range from chest strap based monitoring through to smart clothing containing electrodes to capture the electrical signals produced during a heart beat cycle. Devices are designed to be minimally intrusive to the individual and collect data autonomously facilitating use in the field.

The Polar range of products have been used extensively in studies investigating the cardiovascular system. For example, Paritala's work investigating the effects of physical and mental activity used the Polar RS800 monitor to capture the heart rate of 24 participants during laboratory testing [25]. In a similar study the relationship between markers of work related fatigue and HRV of 28 participants used Polar devices for the data capture [33].

In the above section we have discussed the gold standards for data collection and identified limitations for use in workplace domains. Tools for performance and perceived workload intensity are numerous however, given our proposed domain our choice is limited. Any devices used to measure our required metrics will be worn throughout the working day by forestry workers and so should not cause discomfort during their physical activities. In order to collect data we therefore selected to use Fitbit Charge HR activity trackers to collect ambulatory data and the Polar RS8020CX fitness watch paired with a Polar chest strap to collect heart-rate data. In addition, field-based testing must be quick to conduct to minimise the impact on both productivity and the individual and as such we implemented an electronic version of the unweighted NASA TLX to collect data on perceived workload intensity as this will be quicker than a paper-based survey. For performance data we selected the Deary-Liewald Reaction Time application developed by the Centre for Cognitive Ageing and Cognitive Epidemiology [11]. It is an application specifically designed for conducting reaction time testing using portable computing devices.

3   Studies

The aim of these initial studies was to investigate whether the tracking devices and tools listed above could be used to replicate the results of large-scale, laboratory-based tests which investigate the effects of activity on fatigue. Performing monitoring studies of forestry workers is time-consuming and requires considerable buy-in from a number of different entities (forestry owners, contracting companies, health and safety organisations and the workers themselves). Before undertaking the field studies with workers, we therefore wanted to be certain that our equipment choices would be suitable and we could gather meaningful data. As such we focussed on evaluating combinations of different measurements across different activity types. Each of the studies described below were conducted with a single participant over short periods of time (the course of a day or a focussed activity) as a means to conduct such an evaluation.

3.1   Activity and Recovery

To assess the impact of activity on the physiological and psychological systems our first study was designed to assess the impact of differing loading types (physical and mental) on performance and psychological indicators. During physical loading the measurements were undertaken in an ad-hoc manner as and when opportunities presented themselves. Workload intensity was measured using an electronic implementation of the NASA TLX running on a dedicated server and measured at the same time as reaction time testing. Apparatus was worn for the duration of the monitoring exercise.

3.2   Physical Loading

Monitoring sessions were undertaken over the course of a working day. The participant worked as a floor team member at a large retail outlet, a position

requiring large amounts of ambulatory activity. Both ambulatory activity and heart rate data collection commenced at the start of the work day and concluded at the end of the work period. Workload intensity and performance measurement was undertaken at the start and end of the work period and at the participant's designated break times (2×15min and 1×30 min).

When extracting data for heart rate variability, the raw data is put through an analysis program that computes the various indicators of cardiographic activity. For this task we used the gHRV software developed by Milegroup based at the University of Vigo in Spain. The application is specifically designed for the analysis of heart rate variability.[1]

As mentioned in 1.2, we use the ratio of low frequency to high frequency of the power spectrum density of the heart as an indicator of autonomic nervous system activity, inferring increasing fatigue from the increase of this ratio. The large datasets we create during monitoring can be used to provide point data, however, for our estimation of increasing fatigue we use the cumulative mean of the data over the duration of the monitoring period.

When assessing performance as the speed of mental processing (choice reaction time minus simple reaction time) we see an initial period of improvement followed by a period of slowing as the workday lengthens. The psychological impact was found to increase across duration with increasing perceived task intensity being reported (Fig. 1).
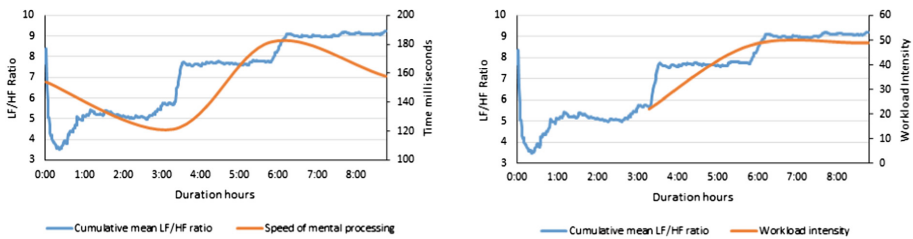
Fig. 1. Speed of mental processing (left) and workload intensity (right) vs cumulative mean LF/HF ratio

Activity during this period was measured as 22,322 steps (18.8 km) with the majority occurring within the first 6 h of the work session. The graphs indicate that speed of mental processing and perceived workload intensity are impacted by this high physical loading.

3.3   Cognitive Loading

To investigate the impact of cognitive load we conducted experiments where driving was used as the mental activity and stressor. Driving is a task requiring constant vigilance and high levels of spatial awareness. The experimenter

---

[1] Available from https://milegroup.github.io/ghrv/doc.html.

performed a 3 h driving exercise on real roads encountering typical driving conditions. Reaction time testing in conjunction with workload intensity monitoring was undertaken at 30 min intervals for the duration of the experiment. On completion of the driving exercise monitoring was continued to assess performance through a recovery period.

We found a similar increasing trend as that in physical loading however, the increase in LF/HF ratio was found to be higher. Workload intensity and performance were also negatively impacted over the duration of the experiment (Fig. 2).
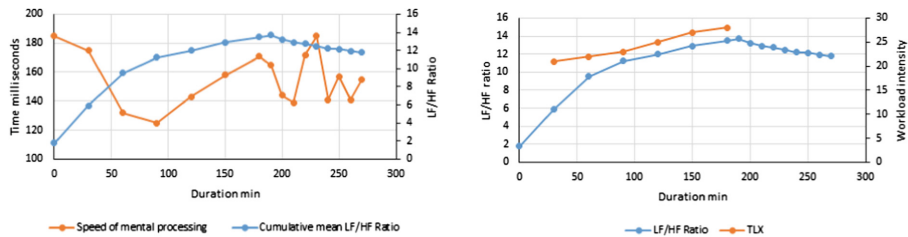
Fig. 2. Speed of mental processing (left) and workload intensity (right) v cumulative mean LF/HF ratio for driving experiment

The higher number of data points collected during the driving experiments facilitates further analysis of existing relationships between our variables. We identified a good correlation for the majority of our measured variables. Table 1 presents the results of the correlation between our measured variables.

Table 1. Correlation of physiological and psychological variables

Condition	Correlation R	Comments
Driving v speed of mental processing	0.73	Speed of mental processing increases with loading duration
Driving v workload intensity	0.98	Perceived workload increases with loading duration
Driving v LF/HF ratio	0.89	LF/HF ratio increases with loading duration
Recovery v speed of mental processing	−0.69	Speed of mental processing decreases on loading cessation
Recovery v LF/HF ratio	−0.75	LF/HF ratio decreases on cessation of loading
LF/HF ratio v workload intensity	0.92	LF/HF ratio increases with increasing workload perception
LF/HF ratio v speed of mental processing driving	0.72	Speed of mental processing increases with increasing LF/HF ratio
Speed of mental processing v workload intensity	0.56	Speed of mental processing increases with increasing perceived workload

4   Discussion

When examining the results of our studies, we can see that each of our measured variables indicate an impact resulting from activity. In both physical and cognitive loading types, we see a reduction in performance across loading duration indicated by slowing of reaction times. We also observe that this reduction in performance occurs as the cumulative mean of low frequency/high frequency ratio increases (Fig. 1). Furthermore, we note that cognitive load impacts the cumulative mean of low frequency/high frequency ratio to a greater extent than physical loading with higher values being recorded (Fig. 2) for our driving study. These results are similar to the findings of other studies investigating the impact mental and physical loading on the autonomic nervous system [10, 23, 27].

In our driving study, we found a good correlation between our measured variables (Table 1) indicating relationships between performance and increasing levels of the high frequency/low frequency ratio. We also identified that performance continues to degrade on cessation of activity prior to improvement facilitated through a recovery period. These findings agree with those of previous studies investigating the impact of loading on the autonomic nervous system [9, 10].

Similar to other studies [9], we found perceived workload intensity increases as loading duration lengthens, individuals report higher workload scores later in the day for the same tasks. The perception of workload intensity was also found to relate to our measured physiological indicators. We identified higher workload scores were reported with increasing cumulative mean low frequency/high frequency ratios.

What we are seeking to identify here is the suitability of wearable devices for capturing meaningful data. As such the results described are not intended to prove relationships between activity, fatigue and response times per se. But rather act as a proof of concept that such methods and tools can replicate known correlations in such data. Given that they indicate this is possible, we can then move on to study their use in the field within our larger-scale studies with forestry workers.

5   Conclusions and Future Work

Our research question asked if low cost lightweight methods can be used for meaningful data collection of data pertaining to physiological indicators related to the impact of activity on performance. As found in other studies [4, 31] the Fitbit Charge HR successfully collected data on ambulatory activity. Having a similar footprint to a wrist watch we found the device minimally intrusive, and caused no discomfort to the individual.

The Polar chest strap used for the collection of heart rate variability proved to be capable of collecting meaningful data. However, the practicality of this device can be called into question. We found the fastening clasp on the chest strap can cause discomfort to the individual, resulting in reluctance to wear the

device over longer periods. We also note a period of spiking may be present due to insufficient moisture between skin and electrode to accurately record heart rate data. This can be overcome by the application of electrode gel prior to the commencement of data capture.

As in other studies [3, 16, 29], our results indicate workplace activity impacts both physiological and psychological states. We identify increasing values of the cumulative mean low frequency/high frequency ratio as loading duration extends. Furthermore, we identify that cognitive loading has a greater impact on the individual than physical loading. We have also identified that on cessation of activity performance continues to degrade prior to improvement. We acknowledge that monitoring undertaken during our study was brief and extended data collection over longer time frames is required to further identify any trends that maybe present. Furthermore, we acknowledge that the data collected represents a single individual as such future studies should incorporate a larger participant base to gain a better understanding of the physiological and psychological impact of workplace activity.

In conclusion, we assessed commercially available fitness monitoring devices as tools for physiological data capture under differing loading types. We found that the selected activity trackers are capable of collecting meaningful data providing researchers with additional tools for monitoring activity in free living. The use of the cumulative mean of the power spectrum density of the heart may be a useful indicator of the impact of activity on the autonomic nervous system and as such may be useful for the determination of workplace role/task intensity.

References

1. Adams, G., Armstrong, H., Cosman, M.: Independent forestry safety review. Final report (2014)
2. Avin, K.G., Law, L.A.F.: Age-related differences in muscle fatigue vary by contraction type: a meta-analysis. Phys. Ther. **91**(8), 1153 (2011)
3. Baulk, S.D., Kandelaars, K.J., Lamond, N., Roach, G.D., Dawson, D., Fletcher, A.: Does variation in workload affect fatigue in a regular 12-hour shift system? Sleep Biol. Rhythms **5**(1), 74–77 (2007)
4. Bowen, J., Hinze, A., Cunningham, S.J., Parker, R.: Investigating the use of activity trackers to observe high-risk work environments (2015)
5. Brisswalter, J., Arcelin, R., Audiffren, M., Delignieres, D.: Influence of physical exercise on simple reaction time: effect of physical fitness. Percept. Mot. Skills **85**(3), 1019–1027 (1997)
6. Castillo-Retamal, M., Hinckson, E.A.: Measuring physical activity and sedentary behaviour at work: a review. Work **40**(4), 345–357 (2011)
7. Chandola, T., Britton, A., Brunner, E., Hemingway, H., Malik, M., Kumari, M., Badrick, E., Kivimaki, M., Marmot, M.: Work stress and coronary heart disease: what are the mechanisms? Eur. Heart J. **29**(5), 640–648 (2008)
8. Choi, S.W., Lee, J.H., Jang, Y.K., Kim, J.R.: Assessment of ambulatory activity in the Republic of Korea Navy submarine crew. Undersea Hyperb. Med. **37**(6), 413 (2010)

9. Cinaz, B., La Marca, R., Arnrich, B., Tröster, G.: Monitoring of mental workload levels. In: Proceedings of IADIS eHealth conference, pp. 189–193 (2010)
10. Collet, C., Averty, P., Dittmar, A.: Autonomic nervous system and subjective ratings of strain in air-traffic control. Appl. Ergon. **40**(1), 23–32 (2009)
11. Deary, I.J., Liewald, D., Nissan, J.: A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. Behav. Res. Methods **43**(1), 258–268 (2011)
12. Dobbie, K.: Fatigue-Related Crashes: An Analysis of Fatigue-Related Crashes on Australian Roads Using an Operational Definition of Fatigue. No. OR 23. Australian Government, Department of Infrastructure and Regional Development, Road Safety Publications, Australia (2002)
13. Ellis, H.D.: The effects of cold on the performance of serial choice reaction time and various discrete tasks. Hum. Factors **24**(5), 589–598 (1982)
14. Evenson, K.R., Goto, M.M., Furberg, R.D.: Systematic review of the validity and reliability of consumer-wearable activity trackers. Int. J. Behav. Nutr. Phys. Act. **12**(1), 159 (2015)
15. Ferguson, T., Rowlands, A.V., Olds, T., Maher, C.: The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int. J. Behav. Nutr. Phys. Act. **12**(1), 42 (2015)
16. Galy, E., Cariou, M., Mélan, C.: What is the relationship between mental workload factors and cognitive load types? Int. J. Psychophysiol. **83**(3), 269–275 (2012)
17. Griffiths, C.J.G.: Investigating tools and methods for data capture of forestry workers. Ph.D. thesis, University of Waikato (2016)
18. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Adv. Psychol. **52**, 139–183 (1988)
19. Haslam, R.A., Hide, S.A., Gibb, A.G., Gyi, D.E., Pavitt, T., Atkinson, S., Duff, A.: Contributing factors in construction accidents. Appl. Ergon. **36**(4), 401–415 (2005)
20. Karvonen, M.J., Kentala, E., Mustala, O.: The effects of training on heart rate; a longitudinal study. In: Annales Medicinae Experimentalis et Biologiae Fenniae, vol. 35, p. 307 (1957)
21. Lim, J., Dinges, D.F.: Sleep deprivation and vigilant attention. Ann. N. Y. Acad. Sci. **1129**(1), 305–322 (2008)
22. MacDonald, W.: The impact of job demands and workload on stress and fatigue. Aust. Psychol. **38**(2), 102–117 (2003)
23. McDuff, D., Gontarek, S., Picard, R.: Remote measurement of cognitive stress via heart rate variability. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2957–2960. IEEE (2014)
24. Morrel-Samuels, P.: Getting the truth into workplace surveys. Harv. Bus. Rev. **80**(2), 111–118 (2002)
25. Paritala, S.A.: Effects of physical and mental tasks on heart rate variability. Ph.D. thesis, Kakatiya University, India (2009)
26. Patel, M., Lal, S., Kavanagh, D., Rossiter, P.: Applying neural network analysis on heart rate variability data to assess driver fatigue. Expert Syst. Appl. **38**(6), 7235–7242 (2011)
27. Pichot, V., Bourin, E., Roche, F., Garet, M., Gaspoz, J.M., Duverney, D., Antoniadis, A., Lacour, J.R., Barthélémy, J.C.: Quantification of cumulated physical fatigue at the workplace. Pflügers Arch. **445**(2), 267–272 (2002)
28. Pilcher, J.J., Nadler, E., Busch, C.: Effects of hot and cold temperature exposure on performance: a meta-analytic review. Ergonomics **45**(10), 682–698 (2002)

29. Rosa, R.R.: Extended workshifts and excessive fatigue. J. Sleep Res. **4**(s2), 51–56 (1995)
30. Soroush, A.: A 6 month physical activity intervention in university staff: effectiveness and health outcomes-the ASUKI step study. Inst för biovetenskaper och näringslära/Dept of Biosciences and Nutrition (2013)
31. Takacs, J., Pollock, C.L., Guenther, J.R., Bahar, M., Napier, C., Hunt, M.A.: Validation of the fitbit one activity monitor device during treadmill walking. J. Sci. Med. Sport **17**(5), 496–500 (2014)
32. Tully, M.A., McBride, C., Heron, L., Hunter, R.F.: The validation of Fitbit Zip physical activity monitor as a measure of free-living physical activity. BMC Res. Notes **7**(1), 952 (2014)
33. Völker, I., Kirchner, C., Bock, O.L.: Relation between multiple markers of work-related fatigue. Saf. Health Work **7**(2), 124–129 (2016)
34. Wright, R.A., Pantaleo, G.: Effort processes in achieving performance outcomes: interrelations among and roles of core constructs. Behav. Brain Sci. **36**(06), 705–706 (2013)

Investigating real-time monitoring of fatigue indicators of New Zealand forestry workers

Judy Bowen*, Annika Hinze, Christopher Griffiths

Department of Computer Science, University of Waikato, New Zealand

ARTICLE INFO

ABSTRACT

The New Zealand forestry industry has one of the highest fatality and injury rates of any industrial sector in the country. Worker fatigue has been identified as one of the main contributing factors. Currently no independent and objective large data source is available that might support an analysis of this, or provide the basis for ongoing monitoring to further investigate. In order to successfully manage fatigue in the forestry workplace, we must identify suitable ways of detecting it. Industry partners are increasingly looking at monitoring solutions (particularly lightweight, wearable technology) that aim to measure worker activities and physiological metrics in order to determine if they are fatigued.

In this article we present the results of studies which investigate whether or not such technology can capture meaningful data in a reliable way that is both practical and usable within the forestry domain. Two series of studies were undertaken with in-situ forestry workers using reaction and decision-making times as a measure of potential impairment, while considering activity levels (via step count and heart rate) and job-roles. We present the results of these studies and further provide a comparison of results across different ambient temperatures (winter vs. summer periods). The results of our studies suggest that it may not be possible to identify correlations between workloads (based on both physical and cognitive stresses) and fatigue measures using in-situ measurements as results are highly personalised to individual workers and can be misleading if the wider context is not also taken into consideration.

1. Introduction

Forestry has one of the highest fatality and injury rates of any industrial sector in New Zealand.[1] Since 2008 there have been 32 fatalities, with claims to the NZ accident compensation scheme (ACC) in excess of two million NZ dollars each year. There were 12,921 active ACC claims between 2008 and 2013.[2] As a result of these poor safety statistics an independent review was conducted by Adams et al. (2014) for the Forestry Industry Contractors Association (FICAday). Data was gathered through interviews and self-reporting of all involved in the sector, from forestry owners and managers to the workers themselves, although the actual number of respondents was relatively small in comparison to the size of the industry. The initial report identified a number of factors contributing to the high accident rate (such as fatigue, lack of training, poor health and safety cultures) and included recommendations primarily aimed at increased participation in training and certification for workers and contractors. However, the report did not consider how the potential underlying causes of accidents might be identified or monitored, nor did it address potentially unsafe work practices or question why these continue to exist.

Aside from the FICA report there has been no large-scale data collection on fatigue in the forestry industry in NZ; information to date has been based on questionnaires and self-reporting of selected groups of workers. This method of data collection has many pitfalls, being subjective by nature and therefore susceptible to response bias, for example it is known that the structure of a questionnaire can have an impact on results as individuals tend to agree with questions (Morrel-Samuels, 2002).

The poor safety statistics and lack of large-scale data collection or analysis of accident causes beyond worker compliance provided the initial motivation for our work. In order to collect a meaningful amount of data from workers going about their everyday tasks, we needed to identify appropriate types of data that could be collected unobtrusively as well as suitable ways to collect such data from forestry workers. In conjunction with this, the industry health and safety organisations began to take an interest in such an approach, and began looking at
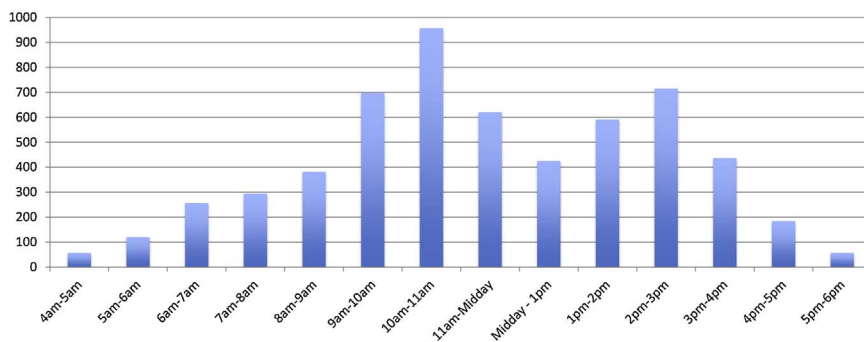
Fig. 1. Reported accidents in forestry by time of day over an 8 year period.

technological solutions used in industries such as the military, mining, haulage driving, etc. Before committing to large-scale purchases of either bespoke or off-the-shelf solutions they also wanted more information about the feasibility of monitoring workers in this way.

We initially considered the use of light-weight mechanisms, such as off-the-shelf activity trackers (e.g., Fitbit, Jawbone, etc.), to gather data on levels of activity (via steps and hill-climbing measurements) and quality and quantity of sleep of forestry workers (Bowen et al., 2015). The goal of this work was two-fold: firstly we wanted to develop an actual data set (rather then self-reported data) from which to understand the working environment and identify worker fatigue (a known cause of accidents and contributor to risk). Secondly we aimed at identifying data that may be used by real-time technological interventions to identify potentially hazardous situations. We began with a series of experiments conducted within the research team and with a small group of forestry workers to validate equipment, methods and data types before moving on to the larger studies described in this article.

From these initial experiments we found that the steps and distances walked by workers did not seem to be significant factors leading to fatigue; relevant aspects seemed to lie in the stress caused by the necessity to pay close and ongoing attention to tasks performed in a potentially hazardous environment. Furthermore, although sleep data was relevant, there were issues of privacy and ethics that needed to be considered when collecting such data, particularly if we were to retain worker buy-in to our studies. We also found that the off-the-shelf solutions for sleep tracking were in some cases unreliable or impractical (levels of discomfort made it unlikely study participants would engage with longer studies involving their use). In these initial investigations we also encountered some resistance to this out-of-work tracking (forestry workers were concerned about privacy and how the data may be interpreted) and as such we decided that we would not include sleep data in our larger studies. We discuss this in more detail in Sections 4.3 and 7.

Some observations from the initial investigations (Bowen et al., 2015), relating to cumulative fatigue were confirmed by analysis of accident statistics made available by one of the independent forestry management companies. The data, collected over 8.75 years, indicates an increasing incident rate throughout the working day, mitigated by breaks taken for lunch (see Fig. 1). As well as the date and time of day, the accident data contains information about which activity and task was involved (we discuss roles and activities of workers later in Section 4) as well as a self-reported reason for the accident. We do not discuss the details of these reports further in this article other than to consider the time-of-day for accident occurrence in Section 7.

Contributions. In this article, we present the results of two studies designed to investigate methods for identifying and measuring contributors to fatigue in the workplace. Each study consisted of a series of visits undertaken with groups of in-situ forestry workers. We use reaction and decision-making times as a measure of potential impairment and compare this with activity levels (measured by step count and heart rate) and job-roles, to try and identify contributors to fatigue. We also

introduce a comparison of results across different ambient temperatures (taken during summer and winter periods) to consider any amplification effect this may have on the same measures. We present the results of our studies and discuss the limiting factors of capturing and using such data in a meaningful way.

Structure of the article. The article is structured as follows: Section 2 provides background information on NZ forestry and the concepts of fatigue and reaction time. In Section 3 we discuss related work in the context of both forestry and workplace monitoring more generally. Section 4 describes the methodology used for our studies and the results are given in Section 5 (Study 1) and Section 6 (Study 2). We also provide a comparison across different working temperatures (seen in winter and summer). A discussion of the results is given in Section 7 with a final summary, conclusions and future work presented in Section 8

2. Background

This section briefly discusses aspects of managerial structures and work organisation in NZ forestry, and introduces concepts and measures of fatigue that are relevant to this article.

2.1. NZ forestry

According to the New Zealand Treasury (2016), the forestry industry contributes about 1% of New Zealand's GDP and provides about 10% of New Zealand's total merchandise exports. There are roughly 1.8 million hectares of plantation forests split between state and private ownership. The industry is governed by four primary associations. The New Zealand Institute of Forestry is the main professional body and there are three separate associations beneath it: Forest Owners Association (FOA); Forest Industry Contractors Association (FICA); the New Zealand Farm Forest Association (NZFFA). Management and consultancy companies exist to source and manage contractors for operations such as harvesting, planting and forest maintenance. In addition these bodies monitor contractors, ensuring on-site operations comply with both operational standards and relevant legislation such as the Health and Safety at Work Act (The Parliament of New Zealand, 2015).

Work is distributed by way of tender for which smaller companies compete for contracts covering the operational aspects of harvesting, transport and siviculture. Each of these smaller companies generally has a workforce of less than 20 employees, with roles split between manual and mechanical operators.

An industry census conducted in 2012 by FITEC (the independent trade organisation now merged with Competenz) found the workforce is biased towards both younger and Māori[3] employees: 21.5% are aged 15–24 (5.6% higher than in the total New Zealand workforce) with 38.5% identifying as Māori (27.2% higher than the total New Zealand workforce) (Competenz, 2012). Low levels of educational achievement

---

[3] The indigenous population of New Zealand.

are prevalent within the industry with over 60% of workers having no formal post-school qualifications. However, current competency requirements are seeing more and more individuals completing industry-specific training.

The work is both physically and mentally demanding, with operations being performed irrespective of the weather. This combined with a poor pay rate of $17.50 per hour median (Payscale, 2014), which is only slightly higher than the New Zealand minimum wage of $15.25 per hour (Legislation, New Zealand) has contributed to a high rate of worker attrition, 45% of forestry workers change jobs within the first 12 months leading to a start rate of around 3000 individuals annually (Ministry for Business and Innovation, 2014). The structure and workforce demographics are relevant to the wider consideration of safety as they provide the context of the workforce and working conditions. For example, one of the contractors involved in our early studies stated that he would like his workers to do fewer hours and to pay them more, but if he did "the contracting company down the road would get all my business as his guys would do more for less and so they would win all the tenders for work".

2.2. Fatigue

Fatigue is a subjective physiological state experienced by individuals as a result of either physical or mental exertion (Hockey, 2013). *Mental fatigue* affects an individual's cognitive processes and *physical fatigue* affects an individual's ability to maintain physical actions. Both types are cumulative by nature, i.e., the level of fatigue generated by a task increases the more a task is performed. Both types of fatigue lessen after a period of rest.

Fatigue is quantifiable by indicators such as reduction of muscle strength (physical fatigue) and slower response times (mental fatigue). Studies into physical fatigue typically seek to replicate high levels of exertion by taxing muscle groups and measure how much force can be exerted over time. For example, Kumar et al. (2002) found that as the duration of muscle contractions increased the level of fatigue also increased.

Mental fatigue can impact the execution of complex thoughts and behaviour and reduce their efficiency (Alvarez and Emory, 2006). It can be induced by long periods of cognitive processing and the effects are particularly pertinent in high-risk work environments. Norman and Shallice (1986) identified five situations that are potentially impacted when an individual is suffering from mental fatigue: (1) planning and decision making, (2) error correction and troubleshooting, (3) un-rehearsed or novel sequences of actions, (4) dangerous or technically difficult situations, and (5) situations that require the overcoming of a strong habitual response or resisting temptation.

The levels of fatigue experienced are determined by an individual's perception of these physiological changes. The energy required for a task (physical or cognitive) are personal to the individual concerned; they are also dependent upon factors such as age, physical fitness and mental ability.

2.3. Measuring fatigue

Physiological changes occurring as one enters a fatigued state can be used as indicators of reduced performance. For example, increased heart rate and core temperature may indicate physical activity, while a reduced reaction time may indicate mental fatigue.

In physically demanding roles, physiological monitoring may provide a better understanding of the impact of the roles and tasks upon an individual. Cumulative measurements taken throughout the workday can provide an understanding of both role- and task-based activities allowing for causes of fatigue to be identified and the associated risk to be managed.

We consider reaction time measurements as a potential indicator of fatigue within an individual over the course of a work period, where a

reduction in reaction time indicates the effects of fatigue. The two 'gold-standard' tests for reaction time are simple reaction time and choice reaction time. We describe these next.

The *Simple Reaction Time (SRT)* test by Galton (1889) is a measure of the (averaged) length of time it takes an individual to respond to either a visual or auditory stimulus. The reaction is typically the pressing of a button. This test is appropriate for gaging general alertness and motor speed of an individual, with an uncertainty factor being introduced by the variable time period between the individual trials.

Brisswalter et al. (1997) examined the influence of physical exercise on SRT with participants exercising at 20%, 40%, 60% and 80% of their maximal aerobic power. During the exercise period they identified a decrease in cognitive performance measured by SRT, although on completion of the exercise period no significant difference in simple reaction time was apparent. This suggests that while we might identify workers as being impaired during the task, post-task measurement may not be able to identify this.

Choice reaction time (CRT) is similar to SRT, but measures the (averaged) response time of an individual by requiring them to make a choice when presented with a visual stimulus. For example, a visual representation of an arrow pointing either left or right is presented and the participant is required to select the corresponding button on a keypad. Having to choose the correct response for CRT requires more cognitive processing than SRT.

Tests are typically run as a number of trials, with uncertainty being added by the choice and a variable time between the stimuli.

3. Related work

This section discusses related work on safety in forestry, as well as fatigue and worker monitoring in general.

3.1. NZ forestry safety

NZ forestry work-sites are generally remote, requiring the working day be lengthened to incorporate travel time (occasionally up to several hours). These long working days combined with a physically demanding workload can contribute to fatigue in workers. This may in turn adversely impact safe working practices and be a contributor to the high numbers of incidents within the industry (Spurgeon et al., 1997).

Lilley et al. (2002) undertook an investigation into the role that rest and recovery play in accidents and injury of workers. This study relied on self-reporting and involved 367 workers responding to a self-administered questionnaire. 78% of participants reported experiencing fatigue at work at least some of the time and the study concluded that the combination of slim margin for error and impairment due to fatigue constituted a significant risk factor within the industry.

In an attempt to gain more detailed data, Parker (2010) conducted a study using wearable video cameras to capture forestry worker behaviours. This work was limited by the small number of participants (due to equipment costs) and the time and expertise required to analyse the footage to understand what was being observed.

As a result of the forestry sector's poor safety statistics, FICA conducted the aforementioned independent review and interim results were published in Adams et al. (2014). The focus of the review was on identifying and analysing the factors that impact on health and safety within the forestry industry and to produce guidelines designed to minimise the amount of incidents. The interim report identified a number of factors which may be contributing to the high accident rate, for example, worker fatigue, lack of training, poor health and safety cultures in the workplace. The report also included a number of recommendations which were primarily based around the creation of new processes, action groups and codes of practice to support and increase participation in training and certification for workers and contractors. However the report did not consider how the potential underlying causes might be identified or monitored.

3.2. Fatigue and reaction time

Levels of activity, both physical and mental, have been shown to increase the level of fatigue experienced by an individual during the course of their day (Pichot et al., 2002; Murata et al., 2005b). In a workplace context, especially where employees are working in remote locations such as those found within the forestry industry, detrimental effects from mental fatigue can invariably lead to situations where individuals can come to harm. From the situations identified by Norman and Shallice (1986) (described in Section 2.2) we consider Points 1, 3 and 4 as being particularly pertinent for forestry workers.

Increasing levels of mental fatigue have also been shown to adversely affect task motivation and mood. A study by van der Linden et al. (2003) using the Wisconsin Card Sorting Test, and the Tower of London Test designed to induce mental fatigue, found that the willingness to apply oneself to a task and do one's best was "significantly lower" for fatigued participants. As such, a higher level of mental fatigue can result in lower productivity levels. This conclusion is reinforced by Murata et al. (2005a), who observe that "fatigue is usually related to a loss of efficiency and disinclination to work."

Williamson et al. (2011) conducted a survey of research into different categories of fatigue and their effect on safety. They concluded that there was strong evidence to link task-related fatigue and performance to safety outcomes. They also noted that there was only limited data available and concluded that more research was needed to understand which roles and activities in the workplace may be more affected by fatigue. We seek to address this by considering different roles within the forestry environment to see if there are differences in the fatigue measures and effects we measure.

Sabzi (2012) investigated the effect of exercise-induced fatigue on choice reaction time. Using a mix of exercise types (aerobic, anaerobic, mixed, prolonged-intermittent and super-maximal-intermittent), they measured the choice reaction time of 15 participants both before and after exercise periods. An increased reaction time was identified across all exercise types with anaerobic, mixed and super maximal intermittent producing the largest differences. They concluded that "exercise-induced fatigue could reduce choice reaction time".

The related work discussed here, as well as many similar studies, (Saito, 1999; Williamson et al., 2001; Lin et al., 2008) all suggest a relationship between work-induced fatigue (either physical or mental) and reaction times. This relationship informs our understanding of the forestry accident rate and its causes.

3.3. Observing challenging workplaces

There are many challenges inherent in collecting observational data in workplace environments. The HumanWork Interaction Workshop (2015) specifically focused on design for challenging work environments. Discussions around data gathering in such environments identified common themes from a variety of different work domains studied, such as safe access to industrial sites, ethical considerations of monitoring employees (including use of, and access to, data) and finding suitable and unobtrusive study methods.

As an example, although there is a reasonable amount of evidence to suggest that activities outside of work (particularly sleep quality and quantity) have an effect on workers, collecting such data raises many ethical issues (as it involves tracking workers during their personal time) and may also lead to resistance on the part of workers to take part in such studies. We discuss this further in Section 7.

4. Methodology

We conducted two studies with forestry workers in their working environment, one in the winter and one in the summer (for details see Table 1). We describe next the methodology of both the winter study, which we refer to as Study 1 (see Section 5) and the summer study (see

Table 1
Days per crew per Visit for Studies 1 (winter) and 2 (summer).

Study	Crew 1	Jul-20	Jul-21	Jul-22	Jul-23		4 days
		Aug-24	Aug-25	Aug-25	Aug-27	Aug-28	5 days
						TOTAL	**9 days**
	Crew 2	Jul-27	Jul-28	Jul-29			3 days
	Crew 3	Aug-10	Aug-11	Aug-12			3 days
Study 2	Crew 3	Feb-22	Feb-23	Feb-24	Feb-25	Feb-26	5 days

Table 2
Participant demographic for Studies 1 (winter) and 2 (summer).

Crew	Participant	Role	Age group	Winter	Summer
1	1A	Loader driver	40–50	×	
1	1B	Loader driver	50–60	×	
1	1C	Quality control	< 20	×	
1	1D	Manual feller	> 60	×	
1	1E	Process operator	20–30	×	
2	2A	Loader driver	40–50	×	
2	2B	Quality control	20–30	×	
2	2C	Quality control	20–30	×	
2	2D	Manual feller	30–40	×	
2	2E	Process operator	20–30	×	
3	3A	Quality control	50–60	×	×
3	3B	Quality control	20–30	×	×
3	3C	Manual feller	30–40	×	×
3	3D	Loader driver	30–40	×	×
3	3E	Loader driver	40–50	×	×

Section 6), which is Study 2. The aim of both studies was to investigate:

(A) if we could identify a measurable correlation between levels of physical activity and cognitive response times, using lightweight and unobtrusive measurement equipment, and

(B) if there were measurable differences between roles with a high physical load (the manual workers) and those with a high cognitive load (machine operators).

This would then enable us to further consider the most suitable measurements and corresponding equipment for an in-situ solution designed to detect fatigue in forestry workers.

4.1. Participants, location and timing

Participants were sourced from three separate forestry industry subcontractors, all performing harvesting operations (for details on individual roles see Section 4.2). All participants were male. The studies were performed with forestry workers at their place of work in the forest during their normal working hours. Study 1 was conducted with three crews (see upper part of Table 1), one at each of one of three operational sites which were all located in the North Island of New Zealand. The participants ranged in ages from 17 to 62 years; participant demographics are presented in Table 2. Study 1 took place in July and August 2015 (during the NZ winter) and was conducted over 3, 4 and 5 day visits (as operational conditions allowed). Table 1 provides a summary of days for each visit per crew in each study. Monitoring commenced at the start of the working day, generally 06.45 am, and concluded at the end of the working day, generally around 3.45 pm.

Study 2 was conducted during the NZ summer months (see lower part of Table 1). It largely replicated the methodology of the winter study. This study was intended to re-investigate (A) and (B) above to see if results were consistent and reproducible, and also provide a comparison of reaction times between different ambient temperatures. Anecdotally we had heard from the workers during the first study that they felt more impaired when working in the heat of the summer months than in the cold temperatures over winter. The participants of

Fig. 2. Examples of role and terrain types.

A) Example of loader driver role

B) Example of quality control role - manual

C) Example of quality control role - mechanised

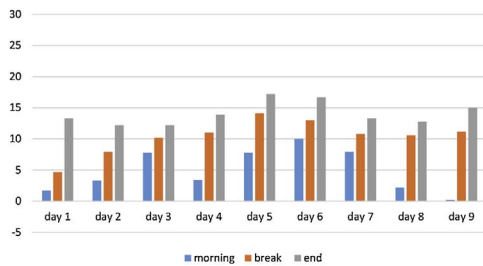D) Example of typical terrain type

Fig. 3. Temperature during study days (Crew 1).

this second study were sourced from one of the crews participating in the winter study (Crew 3), see Table 2 last column. The study took place in February 2016. The workload and activities for the different roles of workers in our study is typically the same irrespective of the season. Some roles (not included in our studies) are more seasonal (for example, planting) but this is not a factor in our studies. The only variation seen was for one of the workers who adopted a more multi-functional role during the summer due to nature of the work required on the site, we discuss this further in Section 6.

The number of days per visit, as well as access times to workers to

conduct reaction time measurements were out of the control of the research team. One of the challenges of collecting data from hazardous workplaces is site access, and restrictions in place on different sites. In addition, the needs of the workers with respect to work performance targets and mandatory break times meant that we could not structure the studies around our preferred requirements, but rather to fit in with the crews. One of the consequences of this was that we were restricted as to when we could interact with the workers to perform reaction time testing. Ideally this would occur at the start of day, before the lunch break (to measure after a prolonged work period), after the lunch break (to measure after rest has occurred) and at the end of day. However, the limited time the workers have for their break and their need to have lunch and rest meant that we could only perform these tests three times a day (morning, after the lunch break and at the end of day). We discuss this further in Section 7.

4.2. Participant roles

Participants performed a variety of tasks on site ranging from manual roles through to mechanised operations. There are two main types of machinery used in forestry: those which remain in a single location while tasks are performed (typically referred to as mechanised
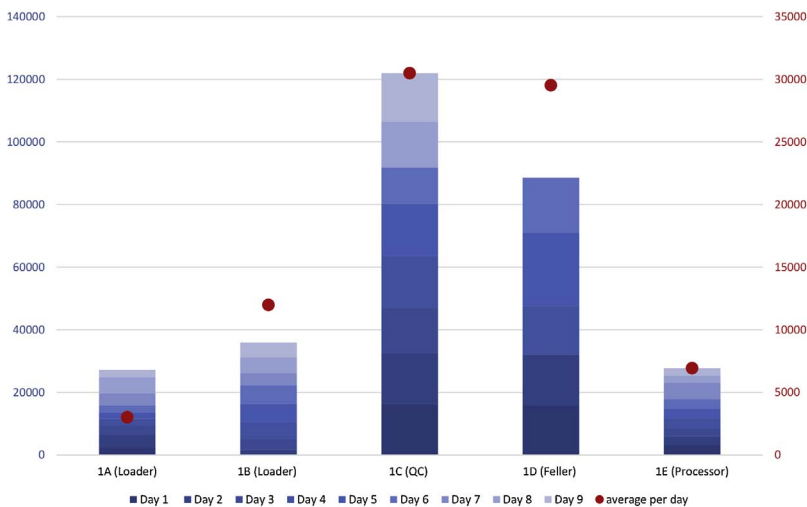

Fig. 4. Step count by participant and role (Crew 1); step aggregation in blue (left axis) and average steps in red (right axis). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
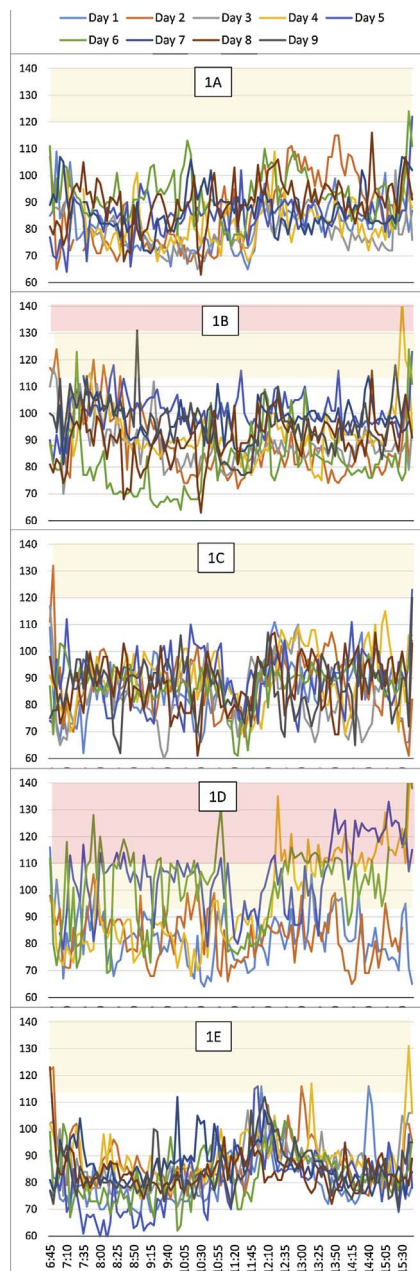
Fig. 5. Heart rate (Crew 1). (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

plant), for example the machines used by loader drivers to lift logs onto trucks; and those which are driven around the site during operations (machines and trucks) such as diggers, log trucks, transport vehicles, etc. Forestry roles are generally dictated by the operational conditions. In smaller work areas (determined by the terrain and planted areas), less machinery may be used due to limitations with getting larger machines into these sites. The size and terrain of the work areas often dictates the production methodology. We have included an overview of the roles monitored during this study and include images to further demonstrate harvesting methods.

Loader driver. Using machinery these operators are responsible for skid site maintenance; stacking of logs for transport, laying out logs for quality control and loading processed logs for transport. Generally, loader drivers commence site operations earlier than other crew members to facilitate early transport of processed logs and prepare the operational area for the day's production. Fig. 2A provides an example of product loading for removal from the worksite.

Quality control. Quality control operations consist of log grading (sizing by diameter) and removal of any remaining branch stems prior to shipping. Quality controllers generally work in close proximity to machinery requiring high levels of spatial awareness. In operations where log making is performed manually the quality control operations include length cutting. Fig. 2B illustrates manual quality control; Fig. 2C illustrates mechanised quality control methods. Quality control roles can be the most varied in terms of the physical activity levels. The amount of walking required is highly dependent on the site, and the role of quality control is often combined with other roles (such as safety observer) as required which also effects the physical nature of the role.

Manual felling. Manual felling generally occurs in remote locations where environmental conditions prevent the use of mechanised felling (steep slopes, inaccessible areas). These locations are generally remote from other site operations and as such the feller typically works alone. On completion of felling operations trees are transported to the skid site for processing using mechanised methods. Fig. 2D illustrates the terrain type requiring manual felling operations be undertaken.

Process operator. The process operators are machinery operators who de-branch and trim the harvested trees. In areas where machinery cannot operate the trees are hauled to a central location for this processing.

4.3. Measurements

We captured three types of data: (1) physiological data as a measure for the level of *activity*, (2) reaction times as a measure for *fatigue*, and (3) ambient temperature at the workplace as a measure of *environmental factors*.

As discussed earlier, in Section 1, we had decided not to include automatic monitoring of sleep by way of wrist-worn sleep trackers. However, given the relevance of sleep data we initially tried to include self-reporting of this by the participants. A questionnaire was created which asked the participants to provide information about what time they went to bed, how long they felt it took them to get to sleep, what time they woke up, and how they rated the quality of their sleep (excellent, very good, average, below average, poor). The questionnaire was available online and on paper so that the participants could choose how to complete it and the onsite researcher could fill it in for them if they wanted to just provide the data directly to him. However, the response-rate for the questionnaire was very poor (less than 20%); therefore, we did not include it as part of Study 2 and do not include the results here.

Activity: physiological data measurement. The level of activity of the participants was measured using two types of physiological data: number of steps taken and heart rate. The measuring was executed using FitBit Charge HR activity trackers (Fitbit Inc., San Francisco, CA, USA). This is a wrist worn device utilising a triaxial accelerometer, vibration monitor, and altimeter to determine activity. Heart rate is collected at the wrist using 'Pure Pulse', a proprietary technology based on photoplethysmography. LED lights are used to illuminate the skin, and an electro-optical cell monitors the change in intensity of reflected light which, in turn, is interpreted as pulse.

These devices are capable of monitoring steps, heart rate, distance travelled, calories expended and sleep. Furthermore, the manufacturers allow third party access to stored data via an API, allowing developers to access stored data for incorporation into other applications.

Each participant was assigned an alpha-numeric identifier (also shown in column 'participant' in Table 2) in order to protect their identities and an account was created for them on the Fitbit.com web application. Accounts were created using the same alpha-numeric identifier which provided the link between the reaction time component of the study and the physiological component.

A Fitbit Charge HR was given to participants at the start of each work day and they were instructed to wear it on the wrist of their non-dominant hand. Throughout the day, the Fitbit Charge HR collected
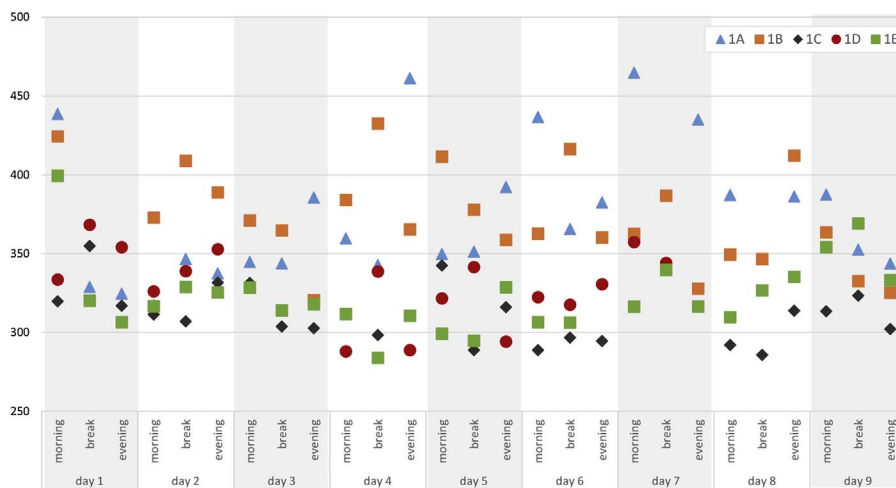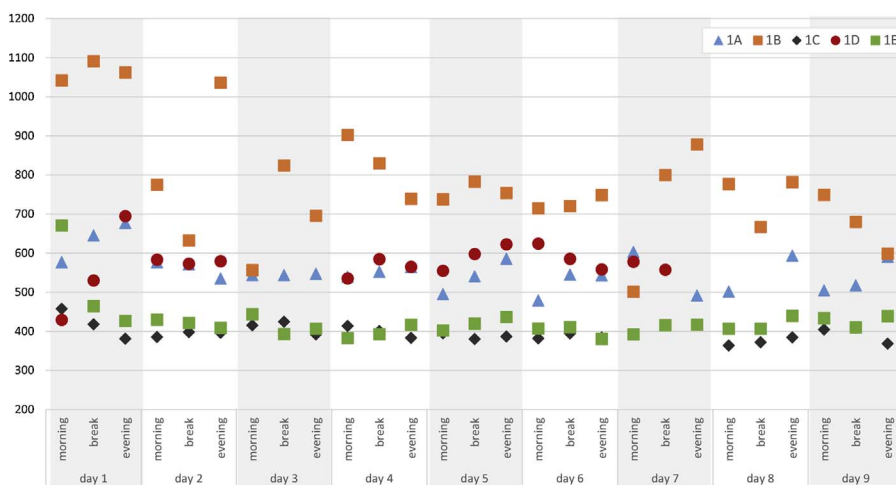
Fig. 6. Simple reaction times (Crew 1).

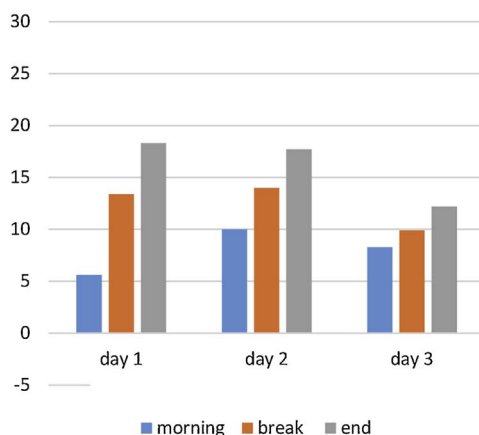**Fig. 7.** Choice reaction times (Crew 1).

Fig. 8. Temperature during study days (Crew 2).

physiological data from the participant. At the end of each working day, the Fitbit Charge HR devices were collected for synchronizing with the Fitbit.com web application and for re-charging.

Fatigue: reaction time measurement. Reaction times were measured using the Deary-Liewald Reaction time test application[4] that was developed by the Centre for Cognitive Ageing and Cognitive

Epidemiology based at the University of Edinburgh (Deary et al., 2011). For the SRT, the software shows a white square against a blue backdrop; the stimulus is the appearance of a diagonal cross in the square to which the participants respond with pressing any key quickly. For the CRT, four white squares are shown next to each other, each corresponding to one of four keys. The stimulus is the appearance of a cross in one of the squares, to which the participant has to respond by pressing the appropriate key.

The software records response times, the inter-stimulus interval for each trial, the keys that were pressed and if the response was correct.

We used an HP Laptop with Windows 10 for the testing using the Deary-Liewald Reaction time software. Simple and choice reaction time measurements for each participant were undertaken three times a day: prior to the commencement of the employees' work period, on completion of their break, and at the end of their work day. Donders (1969), proposed a simple subtraction method for determining decision making times. Splitting the decision-making process into four stages, detection, discrimination, response and motor, Donders used the difference in timings between tests where no choice is required (simple reaction time) and tests where a choice is required (choice reaction time) to infer the speed of mental processing, or decision making time. Using this method in conjunction with our testing data we may gain an insight into the speed of mental processing or, how long it takes an individual to make a decision. We include this calculation as part of our aggregated reaction time results.

Environment: ambient temperature measurements. Temperature readings were undertaken using a McGregors digital thermometer at the same time as the reaction time measurements (described above). We

---

[4] The software is freely available at http://www.ccace.ed.ac.uk/research/software-resources/software.
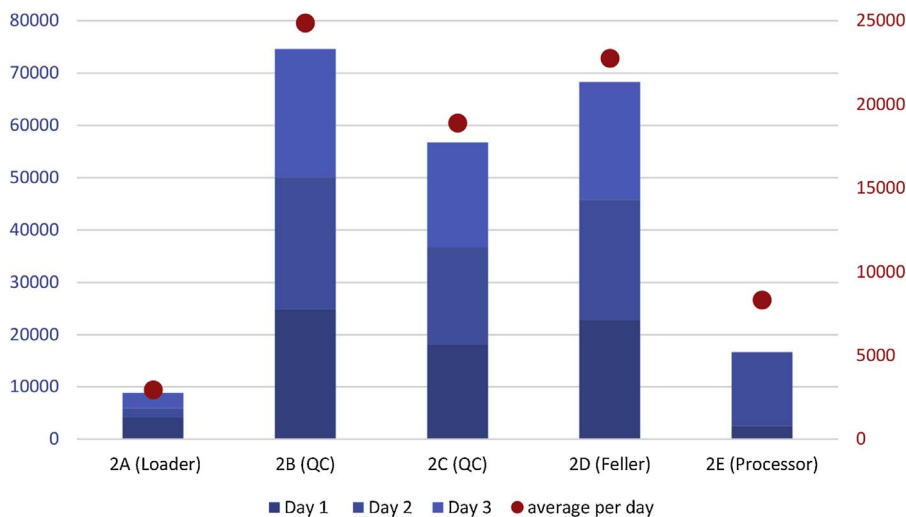
Fig. 9. Step count by participant (Crew 2): aggregated (blue) and average (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

aimed to capture the range of temperatures that the employees were exposed to throughout the course of a work period.

5. Study 1 (winter): results

This study was performed with three crews over varying periods of time (see Table 1). We first present the results for each of the crews, and then discuss the aggregated data.

5.1. Results for Crew 1

Crew 1 is primarily a mechanised crew, i.e., the majority of tasks were performed using machinery (participants 1A, 1B, 1E) with the exception of manual felling (1D) and quality control (1C), see Table 2.

Monitoring took place over two separate sessions; 20th July to 23rd July 2015 and 24th to 28th August 2015. Both studies commenced at 06.45 and finished at 15.45 each day. Where participants were absent from work no data was recorded.

Two of the participants were loader drivers (1A and 1B) who started work on site at 4.00 am to load the trucks for early departure. The other participants from this crew commenced work at 6.45 am. For practical reasons, the first data collection for all participants of this crew was performed at 6:45 am. As such the data collected and reported as *start of the day* for these two participants does not represent the actual starting time, but rather are the start of daily data collection.

Travel time to get to their place of work for a 6:45 am start was not considered for any participants.

Ambient temperature. The temperatures across the study period varied between 0 °C and 17.1 °C with the weather being mainly mild with no rain during the monitoring periods. Fig. 3 shows the temperatures in the morning, during break time and at the end of the shift, measured at the time the participants took the reaction time tests reported below.

Physiological data. Fig. 4 shows both the cumulative steps (in blue) and average steps per day (in red) for each participant across the study duration. Note that not all participants were present each day; for example, participant 1D shows a similar average step number to 1C although their cumulative steps for five days are lower than those for 1C's eight days. Overall, we observed a large variation in the number of steps taken due to the respective tasks performed by the participants. For Crew 1, the member with the Quality Control role (crew member 1C) generated the largest number of steps across both study periods with a weekly mean of 77,072 (equivalent to approximately 58 km). The crew member with a Tree Feller role (1D) also presents a high step rate, 52,020 (equivalent to approximately 40 km). These roles are

predominantly ground-based, with operational layout and terrain type dictating the required levels of manual work.

Fig. 5 shows the mean heart rates recorded during the study. The intensity of activity can be modelled using heart rate (Robergs and Landwehr, 2002), with periods of high and low activity throughout the workday represented by higher and lower heart rates. A maximal heart rate of 220 beats per minute minus the participant's age was used to determine intensity of activity. These were colour-coded in the figure in the following way: a pale green background indicates areas of low intensity, pale yellow indicates medium and red is high intensity. We observe that the heart rates of participants 1A, 1C and 1E indicate low intensity throughout the day, while 1B had peaks of medium and high activity, and participant 1D showed predominantly medium and high activity. While both participants 1C and 1D (QC and feller) walked on average 3–4 times as many steps as their three colleagues, the work intensity of 1D, the feller, is much higher than that of 1C. This is an example of where context is important, but missing, from individual measures. Although both the quality controller and feller walk similar distances, the other activity levels within their roles have very different intensities. The QC role is mainly walking and visually checking, whereas the feller is operating a chainsaw, cutting branches, moving logs, etc. This more physically demanding role can be indicated by the heart rate profile but not the step count.

Participants 1A and 1B perform the same role, however, 1B has a higher average step count and his heart rate profile indicates higher levels of exertion. It may be that the location of his loading tasks requires more frequent movements in and out of his vehicle, or that he is in an area where higher levels of concentration are required (more people and machines around), but there is no way to determine this from the data alone. All of the participants show some increase in heart rate during the afternoon, although 1E (processor) tails off towards the end of day, perhaps suggesting a reduction in workload as the day comes to an end, or self-pacing towards the end of day.

It should also be noted that there are some gaps in the heart rate data. Participant 1B has no data for day 1, and also has some smaller gaps across other days which appear to be where the monitoring device was removed for short periods of time (although we have no confirmation of this). We are missing data for 4 days for 1D, and again there are gaps in some of the days where we do have data. In previous studies we had already noticed that there was a propensity for unexplained gaps in data as well as equipment being lost temporarily which affected continuity of data, we discuss this further in Section 7.

Reaction time. Both simple reaction time (SRT) and choice reaction time (CRT) were measured at three periods throughout the working day (at the start of the shift, at the end of break time and end of the shift).
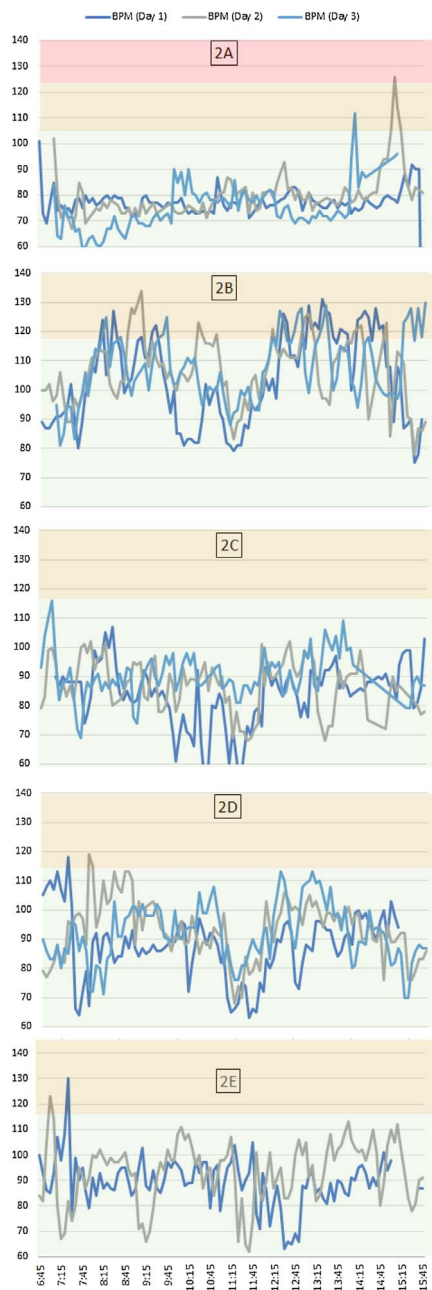
J. Bowen et al.

Fig. 10. Crew 2 – heart rate. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

SRT was measured at each instance 15 times; CRT was measured 20 times. For CRT, only those instances with the correct choice selection were considered. For both SRT and CRT, outliers were removed. We considered as outliers those data points below the first quartile minus 1.5 times the inner quartile range and those data points above the third quartile plus 1.5 times the inner quartile range. For the final CRT or SRT measurement, the mean across the remaining test instances was taken. Figs. 6 and 7 show those values for SRT and CRT, respectively, for Crew 1. Note that the values for participant 1B for SRT on day 1 (break and evening) are omitted here as they are much higher (763 and 1048) and out of the range shown. Higher values represent longer response times, i.e., slower reactions.

Considering the SRT and CRT performance of each participant across the duration of the study we can better identify variation across work periods. All our performance measurements indicated variation across the workday, although some participants had flatter graph

profiles indicating less variation than others.

Individual results for SRT show that there are no defined patterns across each working day for any of the workers. For example, although participant 1A has the slowest reaction times across all days compared with the other participants, on some days he gets faster over the course of the day (day 1), whilst on others he gets slower (day 5). Similarly, we cannot see any direct correlation between heart rate and reaction times. On day 5, participant 1D has a high heart rate consistently in the high intensity zone all afternoon, but his SRT at the end of the day is one of his fastest sets of results. Two of the participants who are the least physically active (by way of step count) are 1A and 1B and these participants also have the slowest SRT results in general, 1B also has the slowest CRT results. Their heart rate intensities reflect their low activity levels (although 1B does spike almost into intense activity at day 4), but they are performing tasks that may be more mentally tiring as it requires them to be alert and aware of their surroundings at all times. However, 1E who has a similar role to the loader 1B (being mostly a machine operator) has the second lowest step count does not exhibit similarly slow SRT and has one of the fastest and most consistent set of results for CRT.

5.2. Results for Crew 2

Monitoring took place over a three day period; 27th July to 29th July 2015 commencing at 06.45 and finishing at 15.45 each day. Crew 2 has a higher level of manual operation than Crew 1 (see Table 2). Participant 2E was absent on the 29th July (day 3).

Participant 2A was a loader driver starting at 4.00 am. As such, data collected and reported as start of the day does not represent the actual starting time for the same reasons as for Crew 1. Remaining participants commence their duties at 6.45 am and therefore data collected encompasses the full working day. Participant 2E has a multi-functional role, mostly operating machinery, but occasionally providing assistance to the feller.

Ambient temperature. Temperature across the course of this study varied between 6 °C and 18 °C with the weather being mainly mild apart from occasional showers on the 28th July (see Fig. 8). Ambient temperatures were warmer during monitoring than those encountered with Crew 1, as such comparisons between the two crews at lower ambient temperatures could not be performed.

Physiological data. Fig. 9 shows the cumulative recorded steps for participants in Crew 2 over the study period. Similar patterns to the roles of participants in Crew 1 (Fig. 4) can be observed with higher step rates in roles such as felling and quality control.

Heart rate monitoring results are presented in Fig. 10; as in Crew 1 activity intensity is indicated using colour bands. Heart rate intensity for all participants in Crew 2 was generally lower than those of Crew 1. In particular the participant with the most physically demanding role, the feller (2D), did not exceed a low level of heart rate intensity apart from a couple of occasions. It may be that he was somehow 'pacing' himself so as to not exert too much physical effort, or it could be that he has a high-level of physical fitness which can affect heart rate intensity. Once again the lack of contextual information means that we cannot determine which, if either, is the case.

Reaction time. Fig. 11 presents the results of the mean reaction time by period for participants in Crew 2 across the course of the work period. Similar to the results for Crew 1 there is no discernible pattern in the SRT results for any participant. In the CRT results, however, participant 2C (feller) shows improvement in speed during the day for each of the three days monitored. Also, the slowest participants were again those with the least physically demanding roles (2A and 2E, loader and processor). The feller (2D) has the second fastest set of SRT and CRT results, both quality controllers (2B and 2C) have fast SRT and CRT results, with 2C showing the fastest CRT response times (end of day 3) across all crews.
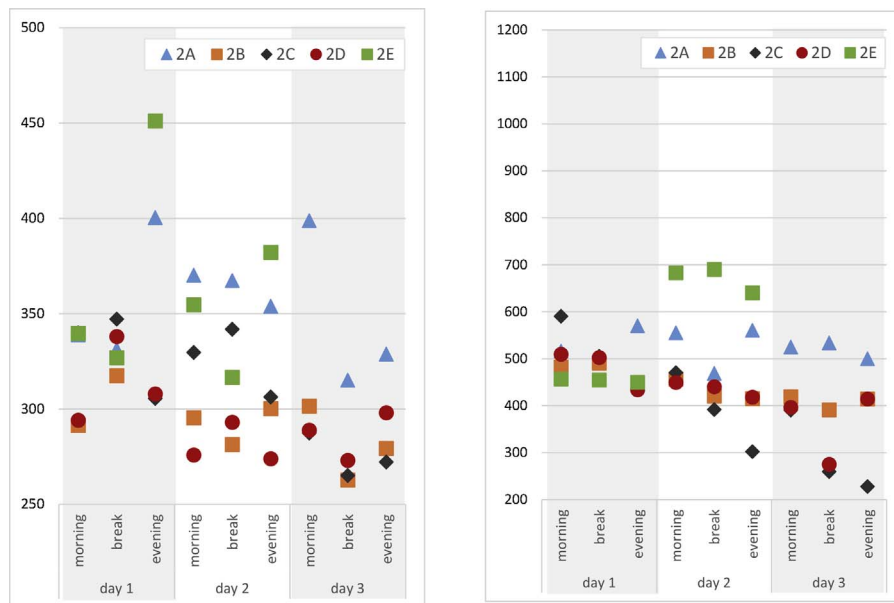
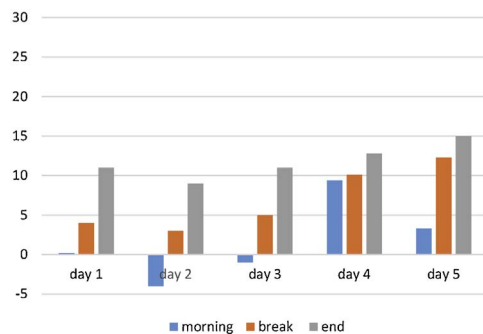**Fig. 11.** Reaction times: SRT – left and CRT – right, (Crew 2).

Fig. 12. Temperature during study days (Crew 3).

5.3. Results for Crew 3

Monitoring took place over a three day period, 10th August to 12th August 2015, commencing at 6.45 am and finishing at 3.45 pm each day, secondary visits were made on the 29th August and 4th September to expand on the data collected during the initial visit. Participant 3E was the loader driver whose duties commenced at 2.00 am, data collection for this participant did not begin until 6.45 am when the remaining participants commenced their work duties.

Ambient temperature. Temperature across the course of the study with Crew 3 varied between −4 °C and 11 °C with the weather being mild throughout the monitoring period (Fig. 12). Days 1, 2 and 3 showed the lowest morning temperatures seen in the study, 0 °C, −4 °C and −1 °C respectively. There is no evidence that this had any effect on reaction time measurements for any of the participants.

Physiological data. Activity was measured in the same way as for Crews 1 and 2 via step counts and heart rate monitoring. This crew performs harvesting operations using a cable hauling system to deliver felled trees to an elevated processing platform; as such the majority of operations are performed using mechanised techniques.

Fig. 13 show the levels of activity (measured as steps) for the differing roles.

The feller role has the highest step rates (some 109,000 equivalent to 83 km) due to the remote locations of stock for harvesting.

Heart rate data is presented in Fig. 14. As with the other crews activity intensity is represented by colour. This crew showed similar heart rate levels across roles to Crew 1. Participant 3E exhibits some large data spikes around the middle of day 1, and also has missing data for the afternoons of days 1–3 which suggests he may have removed the monitor for some reason. This is more likely than an intermittent failure
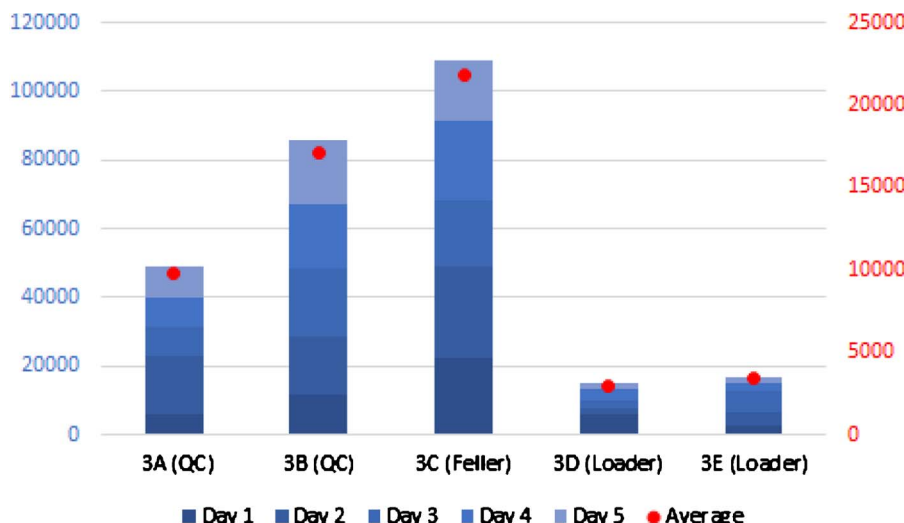

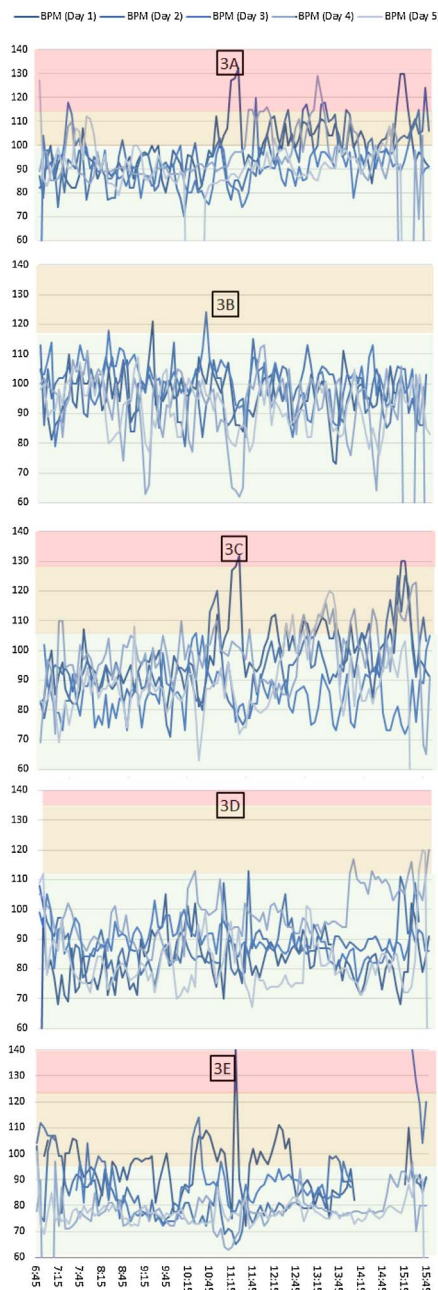**Fig. 13.** Step count by participant (Crew 3).

Fig. 14. Crew 3 – heart rate. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

of the monitoring device at the same time on 3 occasions, although this cannot be entirely ruled out. The feller, 3C, has greater variance between days than the fellers from the first two crews who exhibit a more regular pattern across days. However, unlike the feller from Crew 1, he barely hits the maximum exertion levels on most days.

Reaction time. There are no discernible patterns in the simple reaction times for any of the crew members (Fig. 15). 3A and 3E were always slowest first thing in the morning but sometimes got slower again at the end of the day and sometimes not, they also had slower simple reaction times generally than the other participants. 3D had the fastest, and most consistent simple reaction times, but again there was no fixed pattern. The CRT graph, Fig. 16 is flatter, showing more consistent scores for participants across each day. The biggest change is seen in 3A who starts slow but improves gradually as the week progresses. Of interest is the fact that both 3A and 3E were at their slowest on the morning of day 2, which was the coldest morning seen across all

studies. However, there is no direct evidence that these are related as we do not see the same effect with the other participants.

5.4. Aggregated analysis across Study 1 (winter)

Our analysis of results is comparative by nature. For example, do high levels of activity result in a decrease in performance? (Indicated by increasing reaction times). Or, does temperature impact an individual's ability to perform? We further examine the differences in performance by role.

Steps counted. When considering activity, our investigation shows large differences for physical activity levels (measured as step counts), see Fig. 17. Mechanised operations generate the lowest step rates (i.e., average step counts per day for all loaders is under the average step count of 14,459 across the 15 participants) as operators perform tasks from inside machines in seated positions preventing high levels of activity. The average per day of all fellers is above the average step count across all three crews.

Reaction times. Figs. 18 and 19 show the mean reaction times of each participant.

Note that the aggregation is across 9 days for Crew 1, across 3 days for Crew 2 and across 5 days for Crew 3. The colour schemes indicate crews (blue – Crew 1, orange – Crew 2 and green – Crew 3). The marker shapes indicate roles (square – loader, circle – quality control, triangle – feller, and diamond – process operator).

For selected participants, overall patterns can be detected (e.g., participant 1B shows a slowing down of both simple reaction time and choice reaction time in the evening while participant 3E drastically improves over the course of the day). 1C and 2C have similar trends for both SRT and CRT, but for 3B and 3D they are reversed. As the participants are potentially affected by a number of factors (such as physical exhaustion, cognitive load or even temperature), it is difficult to observe specific patterns by role or across crews.

Fig. 20 shows the decision making times calculated as difference in reaction time. While this averaged data is interesting in identifying if there are trends that can be more obviously detected, it does not necessarily help if the goal is ascribing meaning to data during real-time collection. We discuss this further in Section 7.

Table 3 shows the difference between decision making time in the evening and in the morning. We can see which participants have (on average) faster decision making times at the end of the day (indicated in green in column 1), as opposed to those who get slower (yellow, red). This is not consistent across roles, but rather appears as a personalised measurement. However, Table 4 which orders the decision-making time just on speed in the evening, does show some correlation between role types and decision-making time. We can see that the role-types are clustered based on this ordering, apart from the fellers. Given that actual speed of mental processing is an individual measure (irrespective of alterations over the course of a day due to fatigue, etc.) it is interesting to note this clustering. It could be an artefact of the types of roles people choose to do based on their perceived abilities. There has been some interest from the forestry industry in whether or not specific traits related to both decision making and visual acuity may be useful in assigning workers to different roles, but there is no evidence that this is possible or useful at this stage.

6. Study 2 (summer): results

The purpose of the summer study was comparative along two axes. Firstly to see if we could identify similar trends and patterns in the data to those seen in Study 1 (as a means of identifying if results were generally reproducible). Secondly to see if we could discern differences in the data that could be attributed to the differences in temperature, i.e., to find out if extremes of heat or cold had any amplification effect. The study was conducted with Crew 3 from the winter study so that participants (and roles) were the same and could be directly compared.
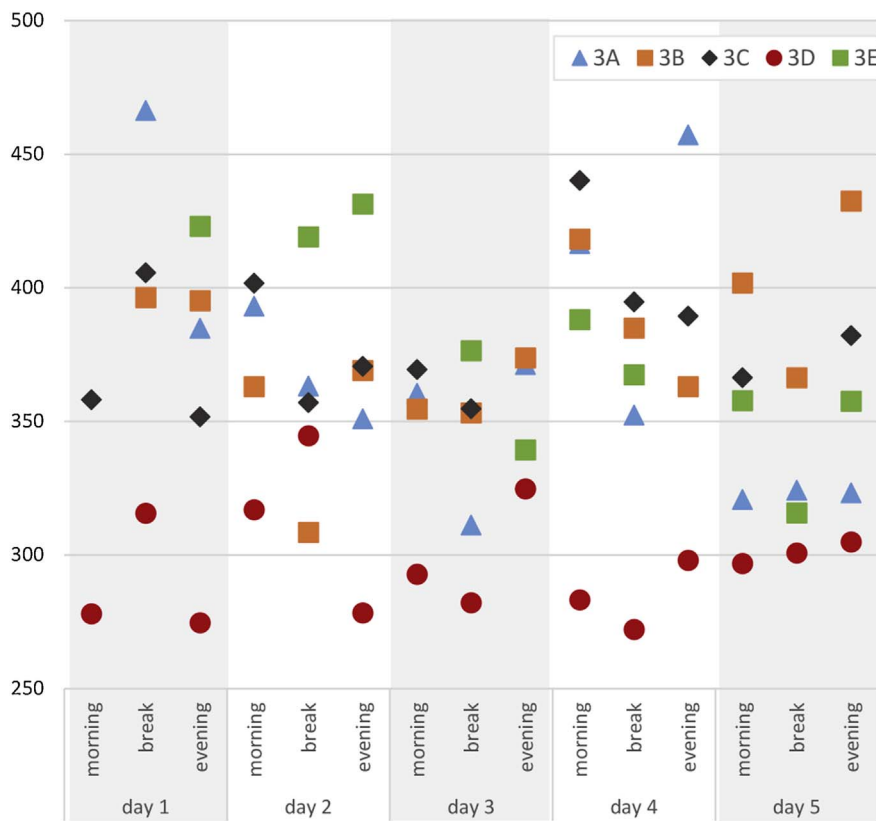
Fig. 15. Simple reaction times (Crew 3).
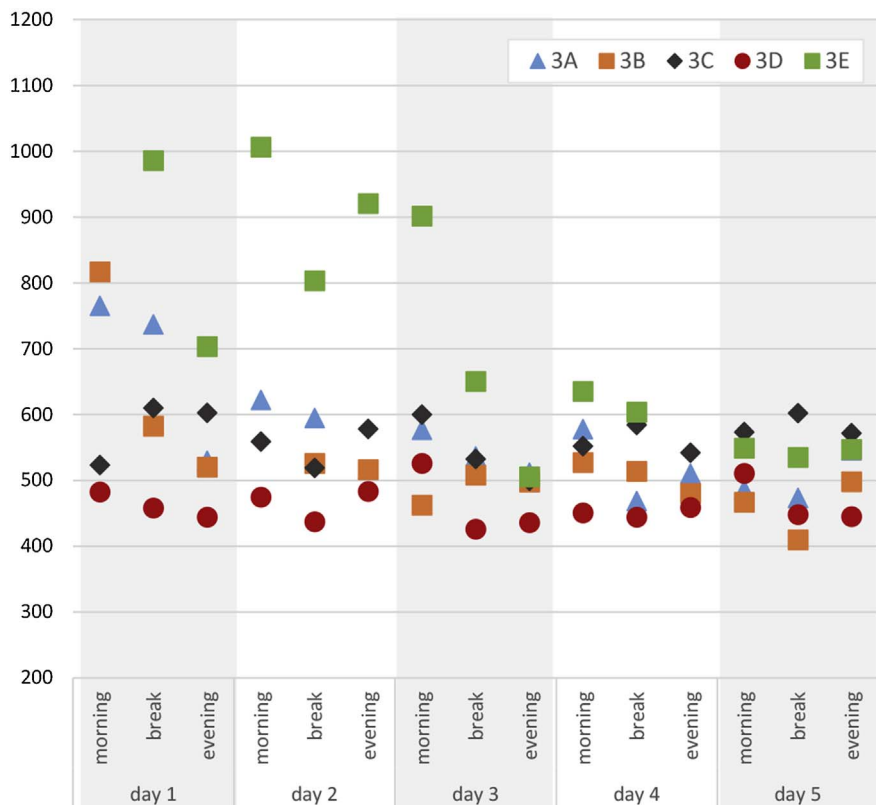
Fig. 16. Choice reaction times (Crew 3).

However, there were some differences in the working environment that were outside of our control and which directly impact comparison of results. The crew was working in a different location, as such the terrain and distances between parts of the work site were different from the first study. In addition, one of the quality controllers (3B) performed other duties during the summer and acted in a more multi-functional role (assisting the feller and log makers primarily on day 3) which affected his physiological data. The location itself also had a direct impact

Participant	Average steps	Role
1C	30493.0	quality control
1D	29517.7	feller
2B	24874.7	quality control
2D	22777.0	feller
3C	21821.6	feller
2C	18906.7	quality control
3B	17118.6	quality control
1B	11977.3	loader
3A	11766.8	quality control
2E	8333.0	processor
1E	6923.3	processor
3E	3395.6	loader
1A	3019.1	loader
3D	3002.6	loader
2A	2964.7	loader

Fig. 17. Comparison mean steps per day.

on our reaction time monitoring as we were only able to gain access to the participants at the start and end of day and not at the lunch break. We first present the results for Crew 3 during the summer, then we compare the results of the two studies in terms of patterns that can be observed and the influence of different temperatures.

6.1. Results for Crew 3

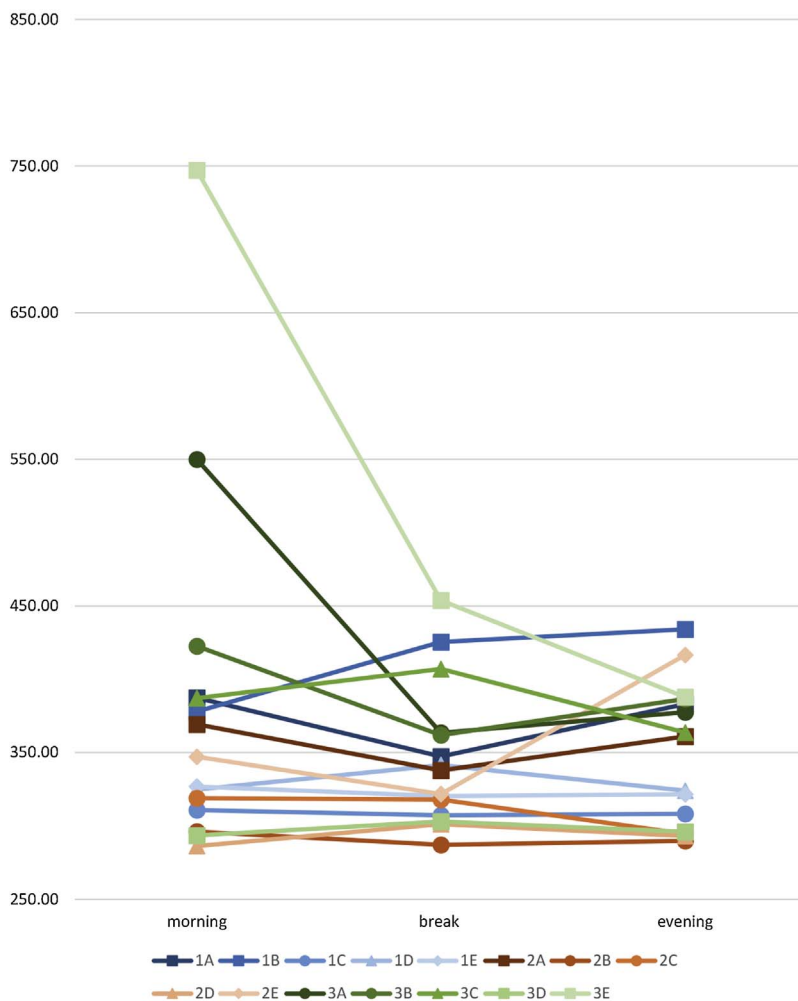*Ambient temperature.* Temperature across the course of this study varied between 12 °C and 27 °C with the weather being dry and mostly sunny throughout the monitoring period. Fig. 21 shows the temperature at the start and end of each day for the 5 day study period.

Physiological data. Activity levels were determined using the same techniques employed during the winter study with data being collected on step and heart rates. The longer duration of this summer monitoring period provides data for a full working week (40 h) providing a better indication of weekly activity by role. Fig. 22 details the cumulative step rates encountered by role. It should be noted that the unusually high step rates performed by one of the quality control operators (3B) (> 100,000) is likely to be a result of a more multi-functional role during the study period.

As in our winter studies, we monitored heart rates throughout the workday to assess if any significant difference was present between summer and winter months. Fig. 23 presents the results of the heart rate measurements across the duration of the study. Participant 3A (Quality Control), who has the second highest step count, shows the highest levels of heart rate exertion. Participant 3B (Quality Control) however, with the highest step rate, has a more consistently low level of heart rate exertion, although also high variability (which is also seen with 3C). The feller, 3C is also at higher exertion levels which is to be expected given the physical nature of his role. The two loader drivers, 3D and 3E have heart rates that generally remain in the low exertion range, although both are typically higher in the afternoon, and 3E in particular has increased levels across two afternoons and several spikes during Days 4 and 5.

Reaction time. Due to site access limitations reaction time measurements were only taken at the start and end of the working day.

Fig. 18. Average SRT per person. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

Fig. 19. Average CRT per person. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)
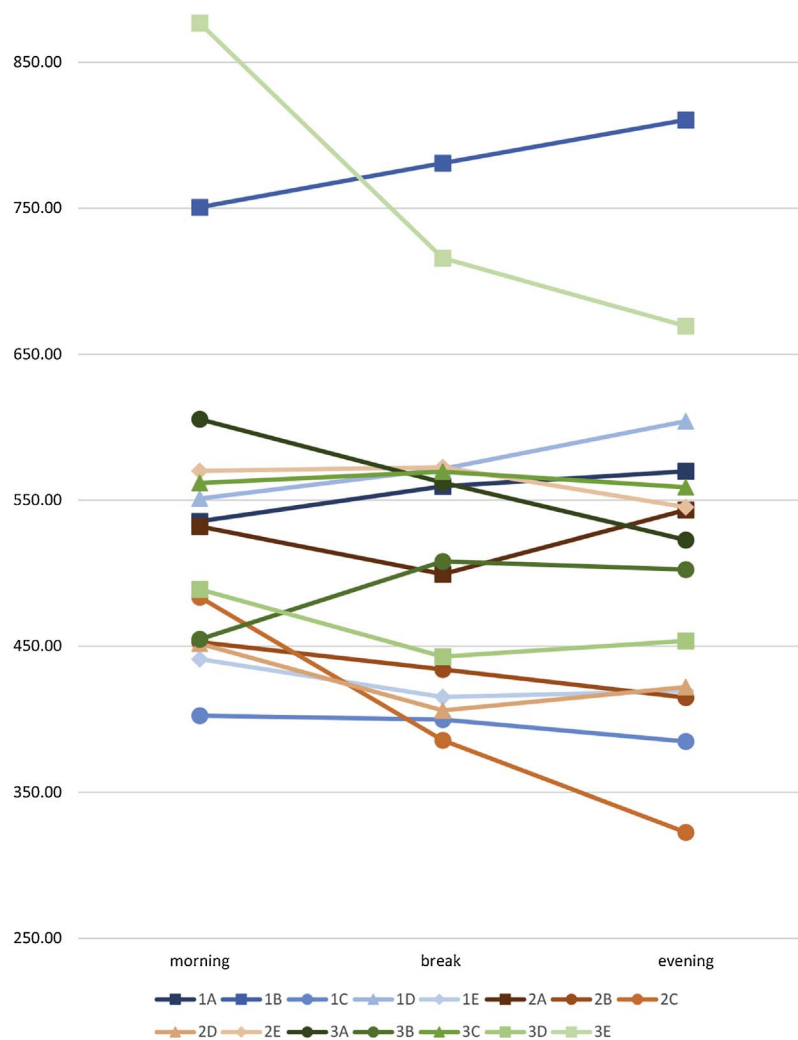
Figs. 24 and 25 present the results of the reaction time testing for the summer study. Of interest are the results for Loader 3D. He has the fastest SRT times for days 1, 4 and 5, but his corresponding CRT times are the slowest. The other Loader, 3E has very erratic (and slow) SRT times (on the morning of day 1 his speed is 804ms which is too slow to show on the graph). However his CRT results are consistent (and among the fastest). No data was recorded for the evening on day 5 for both 3D and 3E. Given the early start the loader drivers have they are keen to leave site at the end of the working day and it is not always possible to persuade them to stay and undertake the reaction time tests.

6.2. Analysis of summer vs. winter

We now present a comparison of the winter and summer data to see if there are similar patterns that can be identified, or obvious indications of an amplification effect due to temperature.

Steps. Fig. 26 shows a comparison of average steps taken by the participants of Crew 3 in summer and winter. We found comparison of step rates across studies can be problematic, especially for Quality Control roles, the distance to the operational area can be significantly different, requiring higher step rates between the safe zone and skid where operations are performed. This is similar for the fellers who are walking between operational areas which may differ greatly across different sites. There was a decrease in mean daily steps for the feller of 7500 during the summer session. While this may indicate the feller was self pacing (due to higher temperature and humidity levels encountered in their operational area), it could equally be an artefact of the different

working location where perhaps the site area being covered was smaller.

There is some evidence that heart rate patterns are consistent across summer and winter, suggesting that the data capture methods are providing reproducible results. Fig. 27 presents a comparison for each participant's heart rate between the two studies. Participants 3B, 3C, 3D and 3E show similar trends across the daily average, while 3A has a larger variation between summer and winter averages. Winter heart rates are on average slightly higher, and for four of the participants are higher at the start of day in the coldest temperatures, which may suggest that the colder temperatures do have an amplification effect on heart rate exertion measures.

When we compare the results of simple and choice reaction times between the two studies (see Fig. 28), we can see similar results for both CRT and SRT, apart from 3E (Loader) who is noticeably slower in winter, particularly at the start of day. 3A (Quality control) is also slower (SRT) at the start of day in winter.

The heart rate and reaction time comparisons across winter and summer suggest that the data is reproducible (similar patterns seen) which gives some confidence in the measurements and collection methods. However, there are no discernible differences between different ambient temperatures (or the extremes at either end) that show this is having any effect generally. Table 5 provides a comparison of decision making times at start and end of day in winter and summer. The lighter colours indicate faster speeds so we can see individual patterns for each worker in the two different seasons. 3A, 3D and 3E exhibit the same patterns (fast-slow, slow-fast, fast-slow respectively) in

Fig. 20. Decision making time per person.
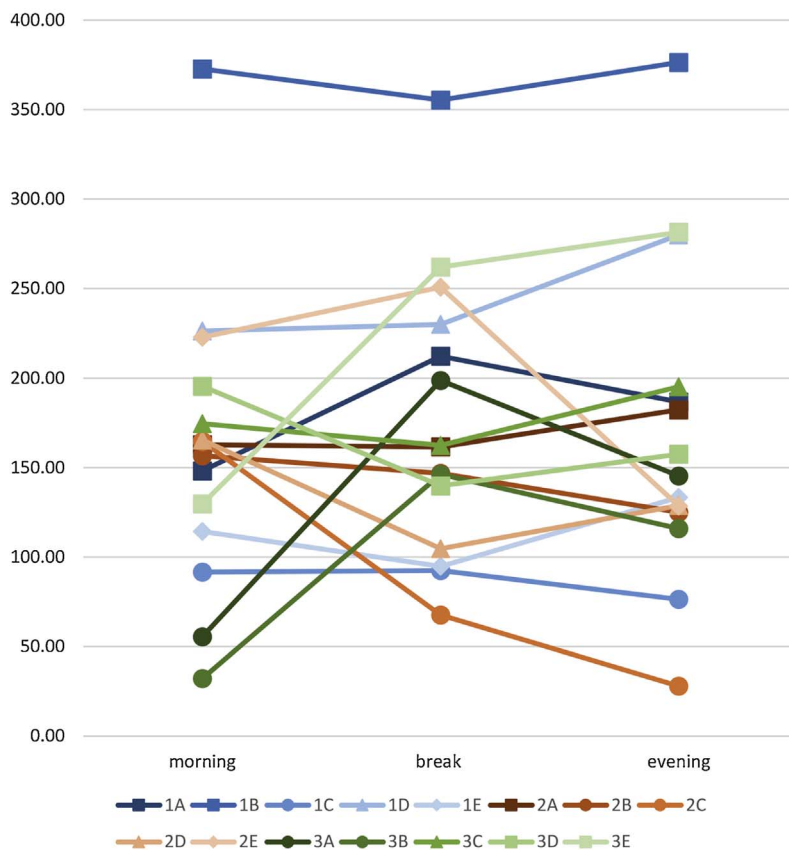
Table 3
Difference Decision Making Time (DMT): morning and evening.

Participant	Diff Morning / Evening	Role	Morning	Break	Evening
2C	-136.77	quality	164.70	67.56	27.93
2E	-94.22	processer	222.81	250.81	128.60
3D	-37.89	loader	195.44	139.88	157.55
2D	-36.62	feller	165.47	104.71	128.85
2B	-31.65	quality	156.69	146.90	125.03
1C	-15.20	quality	91.57	92.49	76.37
1B	3.76	loader	372.65	355.33	376.40
1E	19.04	processer	114.32	94.81	133.37
2A	19.57	loader	162.73	161.55	182.30
3C	20.59	feller	174.57	162.42	195.16
1A	38.39	loader	148.21	212.22	186.60
1D	53.77	feller	226.28	229.99	280.05
3B	83.72	quality	32.18	146.24	115.89
3A	89.75	quality	55.55	198.67	145.30
3E	151.73	loader	129.75	262.06	281.48

both winter and summer, but 3B and 3C reverse their winter fast-slow pattern to slow-fast in summer. We discuss this further in the next section.

7. Discussion

We here discuss the results of our two studies in the light of related work and the context of workplace monitoring ethics.

Steps taken. When examining the amount of steps taken across the working day we find that manual roles result in high step rates that exceed those found in other employment types. Porcari and Ekhwan (2007) conducted a study for the American Council on Exercise into the amount of steps taken by employees in 10 common occupations. They found a wide variation in the step rates of these professions ranging

from 4300 steps for secretaries through to 15,251 for mail carriers. As we can see, harvesting roles generate levels of activity (measured as steps) that far outweigh those encountered in other professions. However, there is no clear indication that these high step counts equate to higher levels of fatigue.

Heart rate. We consider whether the heart rate measurements enable us to differentiate between roles encountered in our study. We find mechanised roles do not produce overly high heart rates. We hypothesise that high points seen around the start and end of the lunch break only occur as a result of individuals climbing in and out of machinery cabs. We would need to confirm this with visual analysis at the same times to be sure of this.

For the manual roles, the variation in heart rate across the workday is more pronounced (e.g., 1D, 3C). In addition, the higher heart rates of

Table 4
Decision Making Time (DMT): ordered by evening speed.

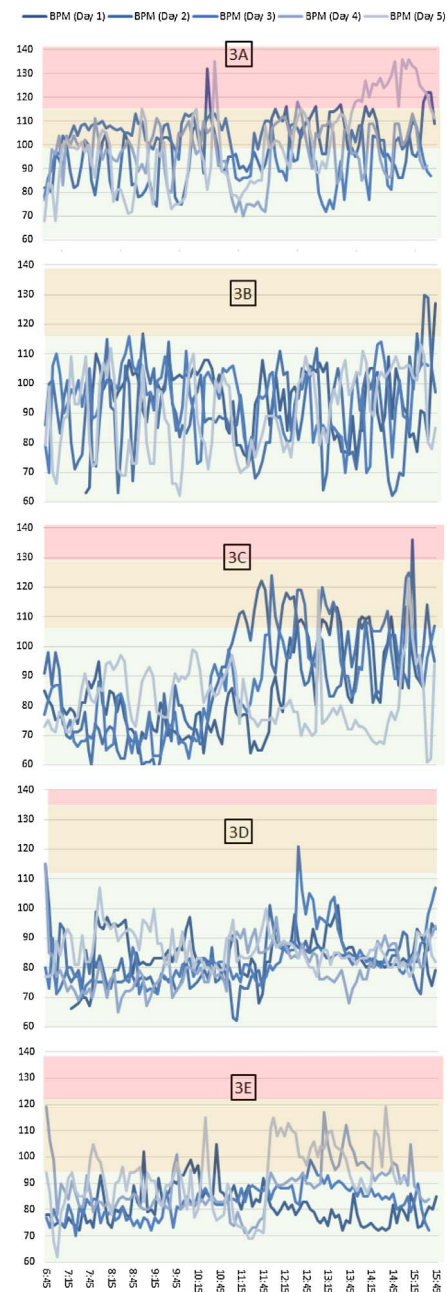| Participant	Role	Morning	Break	Evening
2C	quality	164.70	67.56	27.93
1C	quality	91.57	92.49	76.37
3B	quality	32.18	146.24	115.89
2B	quality	156.69	146.90	125.03
2E	processer	222.81	250.81	128.60
2D	feller	165.47	104.71	128.85
1E	processer	114.32	94.81	133.37
3A	quality	55.55	198.67	145.30
3D	loader	195.44	139.88	157.55
2A	loader	162.73	161.55	182.30
1A	loader	148.21	212.22	186.60
3C	feller	174.57	162.42	195.16
1D	feller	226.28	229.99	280.05
3E	loader	129.75	262.06	281.48
1B	loader	372.65	355.33	376.40
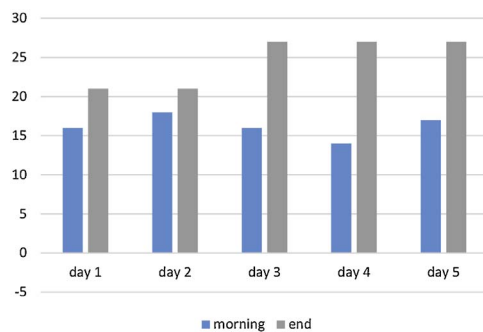
Fig. 21. Temperature during study days, only morning and evening available (Crew 3 summer).
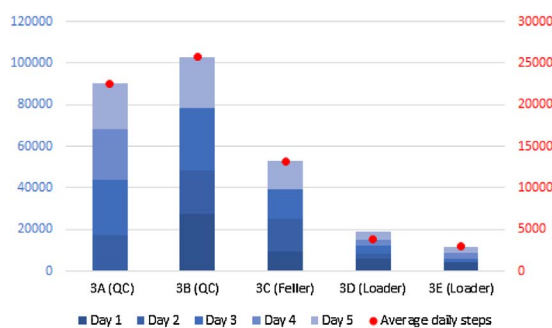
Fig. 22. Step count by participant (summer).

Fig. 23. Crew 3 summer study.

these workers enhance the step count information by indicating that fellers (who are performing high levels of manual labour as well as walking) typically work at a higher rate of physical exertion than quality control roles with similar step counts. However, it is not the case that the most physically active workers have significantly slower re-action times at the middle or end of day so we cannot infer that they are more fatigued or impaired than workers in other, less manual, roles.

Bates and Schneider (2008) used the observed differences in heart rates to identify break periods throughout the course of a workday. Used in this way it may help compliance with legislation by ensuring workers take adequate breaks. Although some similar observations of lower exertion can be found in our results for the lunch break period the data is not sufficiently contextualised or significant to warrant its po-tential use in questions of compliance. Also, heart rate in particular and physically induced fatigue in general are highly personalised to an in-dividual. Age, gender, fitness-level and overall health are all con-tributors to these, and as such the data must be considered within this context. Commercial solutions, such as the Readiband sleep tracker from Fatigue Science[5] typically develop an initial personalised baseline for each individual and consider subsequent data in relation to this.

This, however, requires additional study time (to set the baseline), re-quiring more time commitments from participants which is often pro-blematic. We discuss this further shortly.

Reaction times. When we examine the available workplace accident data (see Fig. 1 in Section 1) by time of day, we see two definite spikes in time periods of incident occurrence: the first occurred between 10 am and 11 am, the second between 2 pm and 3 pm.

Brisswalter et al. (1997) found that the effects of physical exercise led to a decrease in cognitive performance (i.e., choice reaction time), however simple reaction time alone showed no significant difference. Our results show differences between SRT and CRT, with greater var-iation in SRT results. However, neither SRT nor CRT shows a clear correlation to physical activity. As such, while workplace fatigue, manifested as slower reaction times, may be a contributor to the periods
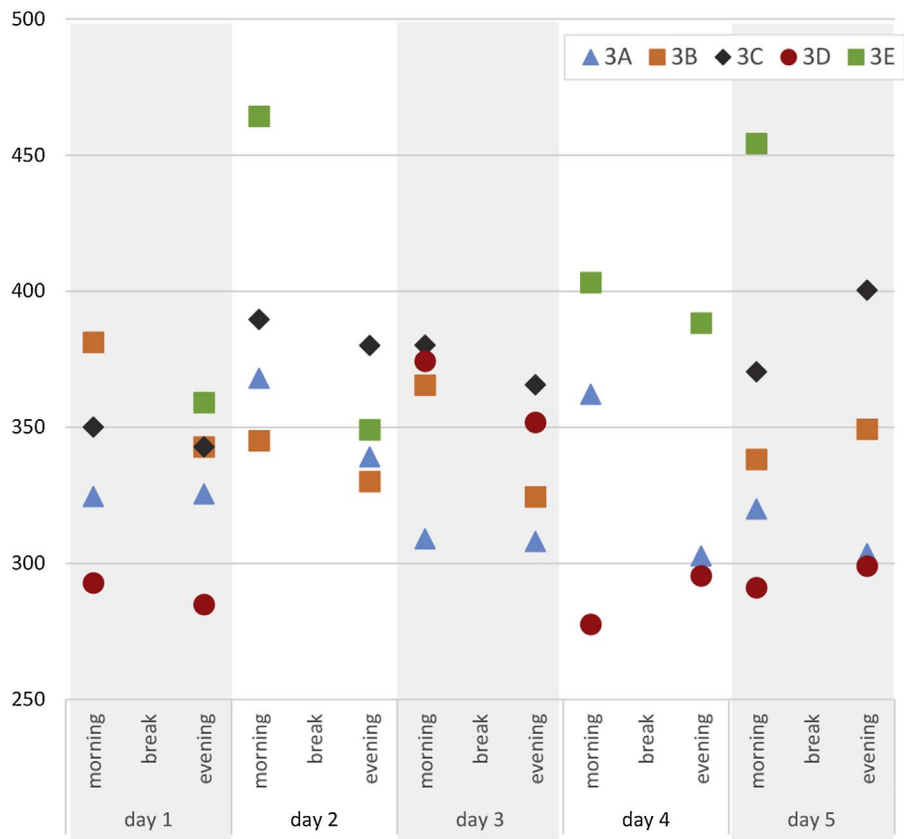
---

[5] https://www.fatiguescience.com.
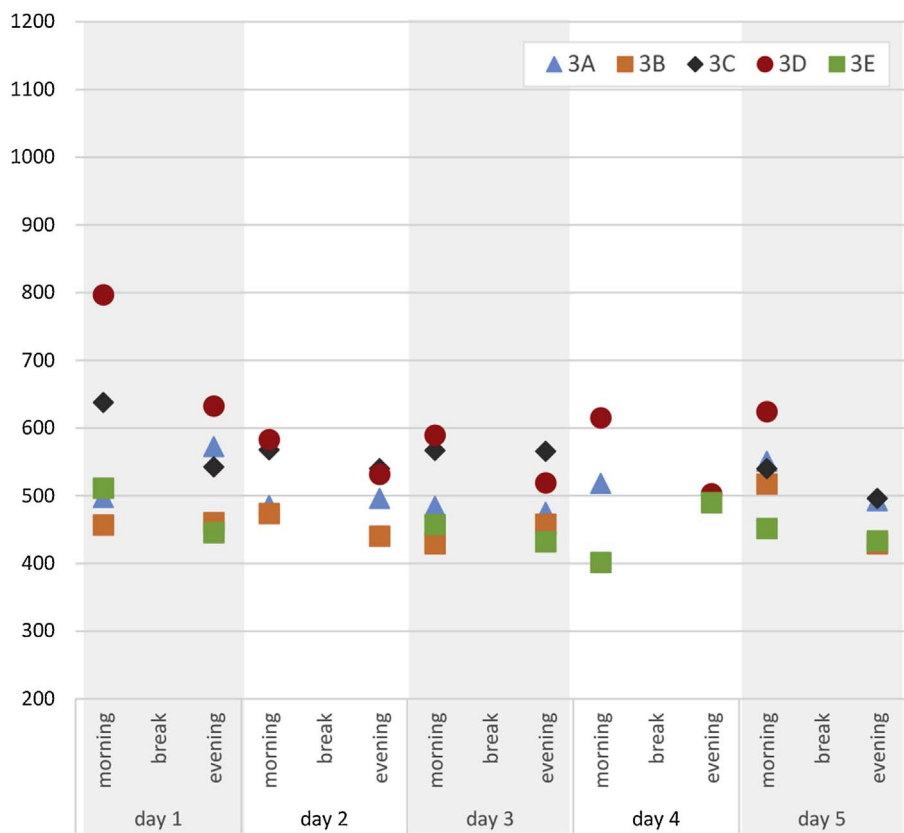
Fig. 24. Simple reaction times (summer).

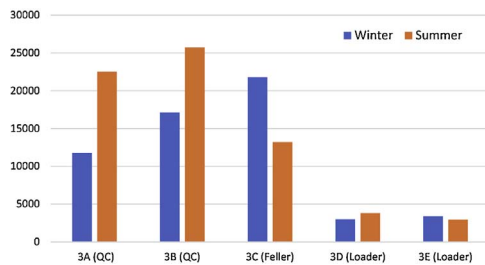**Fig. 25.** Choice reaction times (summer).

Fig. 26. Average steps winter (blue) vs. summer (orange). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
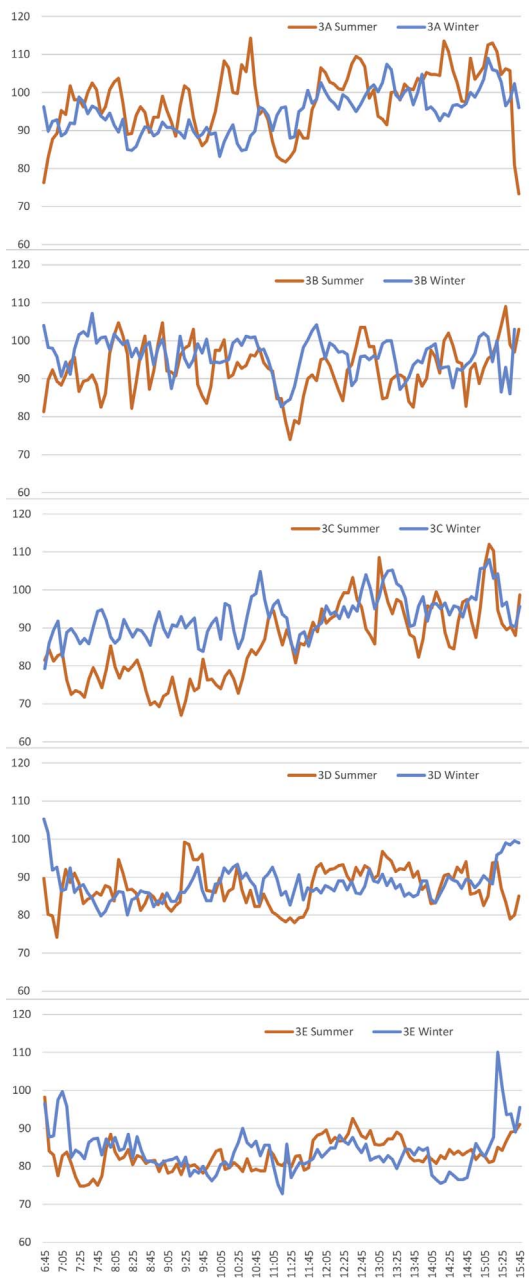

Fig. 27. Heart rate winter (blue) vs. summer (orange), from top: 3A, 3B, 3C, 3D, 3E. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of increased incident rates, we cannot currently identify a direct correlation between reaction time trends and workplace incidences. Measures for CRT and SRT are known to be highly variable across different timescales for individuals. Even with outliers removed (as discussed in Section 5) the variability of the data means that it cannot be used individually as an indicator of either fatigue or work-induced stress.

Influence of temperature. Our investigations of the effect of temperature on performance were driven by comments from work crews during our winter studies. All of the crews indicated that they felt that summer work had the greatest impact on their fatigue levels due to higher temperatures. (Pilcher et al., 2002) concluded from a meta-study that both hot ($> 32°C$) and cold temperatures ($< 10°C$) have negative effects on performance. Similar temperature extremes are reached in our studies during several mornings (e.g., in day 1,2,3, 8 and 9 for Crew 1, day 1 for Crew 2, and all five days for Crew 3). However, we found that although they feel the work is harder in the summer months, their performance is not significantly impacted detrimentally by higher temperatures. If anything, our comparisons in Section 6.2 rather shows a small decline at the lower temperatures during winter months but this is not significant.

Sleep and fatigue. As we discussed in the introduction to this work, we are aware of the importance of sleep and the role it can play in workplace fatigue. The importance of sleep in relation to waking performance has been investigated as a contributor to fatigue many times (Williamson and Feyer, 2000; Åkerstedt et al., 2002; Belenky et al., 2003) with conclusions that low duration or poor sleep quality adversely affecting ones mood, physical performance and cognitive processing abilities. Several studies have identified that increases in reaction time occur with sleep deprivation (Van Dongen et al., 2003; Lim and Dinges, 2008; Kim et al., 2011) suggesting that levels of sleep achieved by an individual can influence performance. More importantly it is further suggested that moderate sleep deprivation can impair performance similar to those levels found in alcohol intoxication (Williamson and Feyer, 2000).

Anecdotally, forestry management have expressed concerns about sleep quantity and quality of their workers. At one of our initial meetings, one individual stated that younger workers might "party all weekend" and "turn up for work on Monday morning exhausted". However, there is no evidence for this. Our early attempts to include sleep tracking as part of our studies led us to understand that there is a high level of resistance for this from workers who feel it is an invasion of their privacy. This is further supported by the forestry industry's own experience of trying to conduct a study using the Fatigue Science Readiband[6] sleep tracking and fatigue monitoring solution, which is seen as the 'gold-standard' method for collecting sleep data outside of a dedicated sleep laboratory. At the time of writing, the health and safety organisation trying to run the sleep study had been unable to recruit enough volunteers among the forestry workers to conduct the study, despite delaying it several times. This reluctance from the workers to be monitored in this way is reflected in our own ethical concerns about the collection and use of such data, leading to questions such as how will a management team deal with a worker who is identified as being fatigued due to a period of poor sleep? Further discussions of data privacy and ethical concerns in the context of monitoring forestry workers are presented in Bowen et al. (2017).

Contextualising monitoring data. There are still many 'myths' in the consideration of fatigue and accident rates and how these can be reduced by worker monitoring. We have found an increase in the desire from management teams to find ways of monitoring workers (both from those involved in our studies, and those seeking to utilise other options such as the Readiband or other proprietary solutions) but this is not necessarily coupled with evidence to support such monitoring. Our studies so far suggest that data must be both contextualised and

---
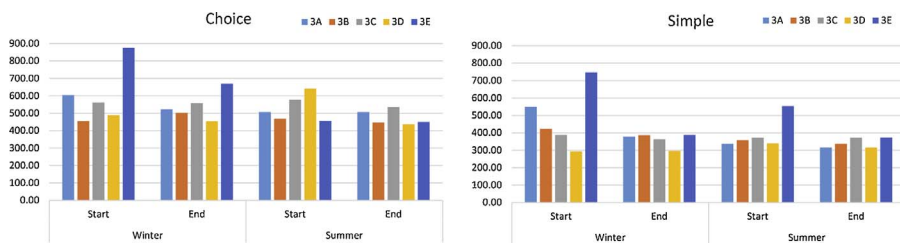
[6] https://www.fatiguescience.com/.

Fig. 28. Mean difference in reaction time winter vs. summer.

Table 5
Decision Making Time (DMT): winter vs. summer.

Participant	Role	Morning (Winter)	Evening (Winter)	Morning (Summer)	Evening (Summer)
3A	quality	55.55	145.30	170.44	191.01
3B	quality	32.18	115.89	89.41	88.39
3C	feller	174.57	195.16	164.39	131.09
3D	loader	195.44	157.55	301.61	121.25
3E	loader	129.75	281.48	39.95	263.45

individualised before it can be meaningfully used to indicate fatigue, and even then it should be understood that indicators of fatigue do not directly correlate to accidents or near misses.

When analysing the type of data we have presented here we must take into account the lack of context. For example although we can see the details of the physiological data throughout the day, if a particular data point is directly affected by an event in the field (e.g., a worker's heart rate spikes because of a near miss accident) we cannot determine this. Even the SRT/CRT testing is prone to the effect of distraction, on at least one occasion we were aware of a worker being distracted by someone else entering the break room as the were doing the response-time testing, and it is likely that more minor distractions were also present that we were unaware of. In addition although we may be more successful at identifying relationships across aggregated data, this is not useful if we wish to capture real-time data and analyse it in the field to provide instant feedback to (or about) individual workers.

Study limitations. The final consideration is the lack of control we have over when and how to perform monitoring, and the limited access we had to participants. Both workers and management are mindful of work performance targets, so we cannot interrupt worker time to any great extent. Conversely, as workers get limited break time they are resistant to undergoing any testing that eats into that time. We would have liked to have performed reaction time testing prior to their lunch break as well as at the end, but this was not possible due to the workers' need to take the break, have lunch, etc. before they would engage in our activities. The accident data presented in Fig. 1 shows two peaks, one immediately prior to the lunch break and the other mid-afternoon. Ideally we would like to have gathered reaction time data at exactly these points, but it proved impossible for us to collect more detailed data at these times other than the automatically collected physiological data.

Similarly, it would have been useful if we could have collected additional information about participants prior to the study. However, this was not possible as we only knew who the participants were when we arrived for the first day of the study. There are a number of factors that directly affect how much physical and mental exertion lead to fatigue. These include things such as age (which we were able to determine), fitness levels, general health and lifestyle factors such as good nutrition and whether or not participants are smokers. Our experience with the sleep questionnaire, which was abandoned due to lack of responses from the participants, meant that we did not pursue this as an additional way to collect data. However, for future studies it would be useful to gather as much of this data as possible from participants during the initial on-site set up phase by way of the researcher directly asking participants. We are also governed by our University Ethics Committee, who grant permission for these studies. They have rules which prevent us from paying, or otherwise inducing, participants, so we cannot use this as a way of encouraging participants to take part in out-of-work study activities.

8. Summary and conclusions

Our studies aimed to investigate the use of in-situ data collection in the New Zealand forestry industry. Specifically, we wanted to identify suitable measurements and measuring techniques from two perspectives: (1) could we reliably capture data from forestry workers over the course of their working day; and (2) could we find meaningful correlations in the data to suggest that we could identify fatigue in workers based on their reaction times and perceived workloads.

Reliable data capture. Results from our two studies show that we can collect data using lightweight off-the-shelf equipment, although there are some restrictions to this. For example, our studies have used wrist-worn commercial activity trackers to collect heart rate data, whereas heart rate variability is likely to provide more reliable data for considering fatigue. However, measurement of heart rate variability in an accurate manner requires the use of chest-straps (wrist-worn light-based heart monitors are not accurate enough in this domain) and these can be uncomfortable to wear for long periods of time by workers engaged in more physical roles. For our next series of studies (future work) we are looking at incorporating such chest-strap monitors into compression shirts so that they are more comfortable to wear.

Meaningful correlations. The data that we collect automatically (including step counts and heart rate) can be compared to reaction time tests which use simple and choice reaction time as an indicator of impairments. However, we did not find significant correlations in our data to show that we can determine fatigue-based impairments from our measurements. Not only do personal factors have a large influence on the physiological data, but there are contextual elements for both types of data (e.g., distractions when a worker undertakes the reaction-time tests, or the desire to finish quickly at the end of the day affecting mindset) which also have an effect. While this may seem to be a disappointing outcome it does provide valuable information for forestry health and safety bodies who are keen to adopt such monitoring approaches. Although some commercial solutions do promise to be able to accurately identify fatigue based on similar measures to our own, our results suggest that they should be cautious in adopting them without investing significant time to study their use in the forestry domain.

It is also important to note that our monitoring periods were short and our participant numbers small. As such our studies can only provide a snap shot of physiological and reaction time data across a limited time period. Extended data collection over longer time frames with larger groups of workers may enable us to better identify any trends that may be present as well as evaluate the reliability of the data collection by way of repeated results. Our future work includes conducting longer studies with larger groups of workers and extending the measurements taken. In addition they will include comparisons with some of the commercial solutions being considered by the industry to see if these produce results that are different or can somehow be validated as more accurate.

In the longer term we are also investigating how the data may be used as part of a larger solution based around an Internet-of-Things

solution which captures a wider variety of data (mixing automatic measures with self-reporting and ambient sensors).

Acknowledgements

References

Adams, G., Armstrong, H., Cosman, M., 2014, October. Independent Forestry Safety Review – An Agenda for Change in the Forestry Industry.

Åkerstedt, T., Fredlund, P., Gillberg, M., Jansson, B., 2002. Work load and work hours in relation to disturbed sleep and fatigue in a large representative sample. J. Psychosom. Res. 53 (1), 585–588.

Alvarez, J.A., Emory, E., 2006. Executive function and the frontal lobes: a meta-analytic review. Neuropsychol. Rev. 16 (1), 17–42.

Bates, G.P., Schneider, J., 2008. Hydration status and physiological workload of UAE construction workers: a prospective longitudinal observational study. J. Occup. Med. Toxicol. 3 (1), 1.

Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., Balkin, T.J., 2003. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. J. Sleep Res. 12 (1), 1–12.

Bowen, J., Hinze, A., Cunningham, S.J., Parker, R., 2015. Evaluating low-cost activity trackers for use in large-scale data gathering of forestry workers. Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction, OzCHI'15. ACM, New York, NY, USA, pp. 474–482 ISBN 978-1-4503-3673-4.

Bowen, J., Hinze, A., Griffiths, C., Kumar, V., Bainbridge, D., 2017. Personal data collection in the workplace: ethical and technical challenges. British Human Computer Interaction Conference, (BHCI) 10 pages.

Brisswalter, J., Arcelin, R., Audiffren, M., Delignières, D., 1997. Influence of physical exercise on simple reaction time: effect of physical fitness. Percept. Motor Skills J. 85, 1019–1027.

Competenz, 2012. Forestry Cluster Needs Analysis. . Available from: http://troq.competenz.org.nz/assets/TROQ/Documents/Forestry/Needs-Analysis-Report-Forestry-22112012-final.pdf (Online).

Deary, I.J., Liewald, D., Nissan, J., 2011. A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. Behav. Res. Methods 43 (1), 258–268.

Donders, F.C., 1969. On the speed of mental processes. Acta Psychol. 30, 412–431.

Galton, F., 1889. An instrument for measuring reaction time. Report of the British Association for the Advancement of Science, vol. 59. pp. 784–785.

Hockey, R., 2013. The Psychology of Fatigue, Work, Effort and Control. University of Sheffield ISBN 9780521762656.

HumanWork Interaction Workshop, Position Papers. Available from: https://projects.hci.sbg.ac.at/hwid/2015/position-papers (Online).

Kim, H.J., Kim, J.H., Park, K.-D., Choi, K.-G., Lee, H.W., 2011. A survey of sleep deprivation patterns and their effects on cognitive functions of residents and interns in Korea. Sleep Med. 12 (4), 390–396.

Kumar, S., Amell, Y., Narayan, T., Prasad, N., 2002. Indices of muscle fatigue. Proceeding of Human Factors and Ergonomics Society Annual Meeting, vol. 13.

Legislation, New Zealand. Acts, 2008. Minimum Wage Act, 1983. New Zealand.

Lilley, R., Feyer, A., Kirk, P., Gander, P., 2002. A survey of forest workers in New Zealand.

Do hours of work, rest, and recovery play a role in accidents and injury? J. Saf. Res. 33 (1), 53–71.

Lim, J., Dinges, D.F., 2008. Sleep deprivation and vigilant attention. Ann. N. Y. Acad. Sci. 1129 (1), 305–322.

Lin, Y.-H., Chen, C.-Y., Lu, S.-Y., Lin, Y.-C., 2008. Visual fatigue during VDT work: effects of time-based and environment-based conditions. Displays 29 (5), 487–492.

Ministry for Business and Innovation, 2014. Independent Forestry Safety Review an Agenda for Change in the Forestry Sector.

Morrel-Samuels, P., 2002. Getting the truth into workplace surveys. Harvard Business Review 80 (2), 111–118.

Murata, A., Uetake, A., Takasawa, Y., 2005a. Evaluation of mental fatigue using feature parameter extracted from event-related potential. Int. J. Ind. Ergonom. 35 (8), 761–770.

Murata, A., Uetake, A., Takasawa, Y., 2005b. Evaluation of mental fatigue using feature parameter extracted from event-related potential. Int. J. Ind. Ergonom. 35 (8), 761–770.

New Zealand Treasury, 2016. New Zealand Economic and Financial Overview 2016.

Norman, D.A., Shallice, T., 1986. Attention to Action. Springer US, Boston, MA, pp. 1–18. ISBN 978-1-4757-0629-1.

Parker, R., 2010. Technological Advances in the Analysis of Work in Dangerous Environments: Tree Felling and Rural Fire Fighting. PhD Thesis.

Payscale, 2014. Hourly Rate for Industry: Logging. Available from: http://www.payscale.com/research/NZ/Industry=Logging/Hourly_Rate (Online).

Pichot, V., Bourin, E., Roche, F., Garet, M., Gaspoz, J.-M., Duverney, D., Antoniadis, A., Lacour, J.-R., Barthélémy, J.-C., 2002. Quantification of cumulated physical fatigue at the workplace. Pflügers Archiv. 445 (2), 267–272.

Pilcher, J.J., Nadler, E., Busch, C., 2002. Effects of hot and cold temperature exposure on performance: a meta-analytic review. Ergonomics 45 (10), 682–698.

Porcari, J., Ekhwan, R., 2007. Do you do 10k a day? Am. Council Exerc. Fit. Matters 12 (4), 9–12.

Robergs, R.A., Landwehr, R., 2002. The surprising history of the "hrmax= 220-age" equation. J. Exerc. Physiol. Online 5 (2), 1–10.

Sabzi, A.H., 2012. The effect of different fatigue protocols on choice reaction time. Middle-East J. Sci. Res. 12.

Saito, K., 1999. Measurement of fatigue in industries. Ind. Health 37 (2), 134–142.

Spurgeon, A., Harrington, J.M., Cooper, C.L., 1997. Health and safety problems associated with long working hours: a review of the current position. Occup. Environ. Med. 54 (6), 367–375.

The Parliament of New Zealand, 2015. Health and Safety at Work Act 2015. Available from: http://www.legislation.govt.nz/act/public/2015/0070/latest/DLM5976660.html (Online).

van der Linden, D., Frese, M., Meijman, T.F., 2003. Mental fatigue and the control of cognitive processes: effects on perseveration and planning. Acta Psychol. 113 (1), 45–65.

Van Dongen, H.P., Maislin, G., Mullington, J.M., Dinges, D.F., 2003. The cumulative cost of additional wakefulness: dose–response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep-New York Then Westchester- 26 (2), 117–129.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accid. Anal. Prev. 43 (2), 498–515. http://dx.doi.org/10.1016/j.aap.2009.11.011. ISSN 0001-4575. Advancing Fatigue and Safety Research.

Williamson, A.M., Feyer, A.-M., 2000. Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. Occup. Environ. Med. 57 (10), 649–655.

Williamson, A.M., Feyer, A.-M., Mattick, R.P., Friswell, R., Finlay-Brown, S., 2001. Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. Accid. Anal. Prev. 33 (3), 313–326.

Personal Data Collection in the Workplace: Ethical and Technical Challenges

Judy Bowen, Annika Hinze, Christopher Griffiths, Vimal Kumar, David Bainbridge
Computer Science Department
University of Waikato, New Zealand
{bowen,hinze,kumar,bainbridge}@waikato.ac.nz,
cjgg1@students.waikato.ac.nz

Forestry is a dangerous work environment and collecting data on site to identify and warn about hazardous situations is challenging. In this paper, we discuss our attempts at creating continuous data-collection methods that are ethical, sustainable and effective. We explore the difficulties in collecting personal and environmental data from workers and their work domain. We also draw attention to the specific challenges in designing for sensor-based, wearable rugged IoT solutions. We present a case-study, comprising of a number of experiments, which exemplifies the work we have been undertaking in this domain. The case study is based on our approach to developing a robust, trusted Internet of Things (IoT) solution for dangerous work environments (specifically the forestry environment). We focus the results of this case-study on both the technical successes and challenges as well as the personal and ethical challenges that have been elicited.

1. INTRODUCTION

New Zealand has around 1.8 million hectares of plantation forests and the industry contributes roughly 4% of national GDP to the economy. Forestry also has the highest fatality and injury rate of any industrial sector in NZ (since 2008 there have been 32 fatalities) and has New Zealand's highest rate of workplace injuries with claims to the NZ accident compensation scheme (ACC) in excess of two million NZ dollars each year. An independent review of all involved in the sector (using interviews and self-reporting) identified potential contributors to the poor safety record Adams et al. (2014). These included a lack of training; worker fatigue; poor health and safety processes. As a result a number of recommendations were made based around initiatives such as increased codes of practice, wider participation in training and certification for workers, the creation of new safety action groups etc. However, there was no deeper consideration of the wider underlying causes nor practical proposals for how to identify and prevent unsafe work practices.

While the specific NZ forestry setting is unique, other outdoor-based and labour-intensive industries such as mining, haulage, all-terrain farming and fishing encounter similarly hazardous situations. Known pressure points are again fatigue, de-hydration, distraction, isolated work, remote locations, inexperienced and poorly paid staff, and time pressures. Our initial interest in this domain was motivated by finding ways to unobtrusively gather large amounts of data from forestry workers in order to generate an actual data set of work and environmental factors (rather then self-reported data) from which to understand the working environment and identify worker fatigue (a known cause of accidents and contributor to risk). There are well-known, and well-studied techniques for using biometric data to indicate and measure fatigue (we discuss these in more detail in section 2) but these are typically laboratory-based and invasive and therefore not suitable for in-situ workplace monitoring. We are interested in finding suitable technological solutions to replicate such measures in an unobtrusive fashion using technology that can be easily deployed in an outdoor setting. Ultimately we would like to use real-time data capture to monitor and understand worker metrics with a view to being able to identify hazardous situations as they arise, and intervene as appropriate. The over-arching aim, therefore, is that of reducing the high accident rate in NZ forestry.

There are many challenges inherent in collecting observational data in workplace environments. Human Work Interaction Workshop (HWID) (2015) specifically focused on design for challenging work environments and how to collect relevant data to inform such design. Common themes emerged from a variety of different work domains studied, such as safe access to industrial sites, ethical considerations of monitoring employees (including use of, and access to, data) and finding unobtrusive study methods. Our own initial studies, which looked at the use of lightweight and cheap data gathering tools (such as activity trackers) encountered similar problems. Early on in our work it became clear that the technical, ethical and

sociological challenges of implementing worker tracking (even on a small scale) required us to find alternative approaches which better addressed these.

Over the course of subsequent experiments it became clear that we needed to measure a variety of different metrics (beyond those offered by basic activity trackers) and develop a better way of understanding the personalised implications of fatigue for workers. This has led to our current approach, which we present in our case study. This seeks to adopt a new style of IoT (the *Rugged Internet of Things* or RIoT) as a possible solution. This in turn creates additional challenges, which we also discuss.

This paper therefore addresses two problems. The first is the difficulty in collecting personal and environmental data from workers and their work domain. We address this by proposing a specialised sensor-based IoT solution for outdoor environments. The second is the technical and ethical issues that arise from this proposed solution. Our contributions are the insights we provide into the problems associated with data gathering in hazardous work environments and our proposal of how to move forward using a new version of IoT (RIoT), which is not only suitable for use in rugged and unconnected environments but also considers new mechanisms for data security and privacy. The IoT solution we are developing is not only truly body/person-centric, but is also designed to support people's safety on a day-to-day basis.

2. BACKGROUND TO NZ FORESTRY

An investigation into the role that rest and recovery play in accidents and injury of workers was undertaken by Lilley et al. (2002). This relied on self-reporting and involved 367 workers responding to a self-administered questionnaire. The results showed that 78% of workers reported experiencing fatigue at work at least some of the time and the study concluded that the combination of slim margin for error and impairment due to fatigue constituted a significant risk factor within the industry. In an attempt to gain more detailed data, Parker (2010) conducted a study using wearable video cameras to capture forestry worker behaviours. This work was limited by the small number of participants (due to equipment costs) and the time and expertise required to analyse the footage to understand what was being observed. Adams et al. (2014) conducted another study using self-reporting specifically focusing on forestry, with the results outlined in our introduction.

In general, robotic solutions, which are applied elsewhere in forestry, do not work well in the extreme New Zealand terrain, although there is ongoing research in this area to try to adapt equipment or develop new machinery to remove humans from the work environment. Live observation and in-situ monitoring are not suitable in many work conditions and particularly do not work in hazardous work environments. For example, Parker's initial observational

studies elicited more about the practice of keeping on-site visitors safe than it did about typical worker behaviours. Some of the specific challenges of data gathering in NZ forestry have been reported in (Bowen et al. 2015b), but here we describe a case study we have carried out in order to identify both the practical requirements as well as the philosophical, ethical and social implications of such work when we endeavour to introduce novel technological solutions into industrial environments.

3. RELATED WORK

We focus on related work in three key areas: uses, effects and ethics of monitoring workers; measuring fatigue, activity, recovery and response times; using sensors and IoT solutions in work domains.

3.1. Tracking of Workers

Employee monitoring and tracking is not a new idea. Different approaches have been used to consider issues such as productivity, health and safety and security since the early days of the factory floor-walker (human observation of worker productivity) and the punch-in time clock used to ensure workers arrived on time and did not leave early. As technology has advanced, so too have the methods used for monitoring and tracking workers. Any form of monitoring of employees can create tension between employers and employees.

Botan (1996) reports on a survey of 465 employees on their attitudes of workplace surveillance. He found that irrespective of the motivations for the surveillance, most workers felt untrusted by their employers and that this was likely to be the first step in other management interventions that would not be in the employee's best interests. Kortuem et al. (2007) discussed the use of worker tracking specifically for Health and Safety purposes. This considered the use of a vibration monitoring technique for industrial workers aimed at reducing a condition called "Vibration White Finger". They considered whether the use of (ubiquitous computing in this instance) could play a role in making industrial workplaces safer. They found that even in an example where the monitoring was intended to keep workers safer there was still a perception that such a system could be used to exert control over employees, for example by creating accurate logs of worker activities. Such perceptions existed even when the reality of the monitoring did not include such aims.

The proliferation of personal activity trackers in recent years has given rise to a new type of worker tracking. Firstly there are companies who seek to promote the health of their workers by encouraging them to be active and provide trackers for personal employee use to support this. For example Target in the U.S have offered to give FitBit trackers to all of their workers to increase awareness

about healthier lifestyles[1] In a similar manner, although with more focus on rewarding adherence, oil company BP track step counts of workers and offer lower health care premiums to those who meet certain criteria[2]. According to technology research company Gartner, in 2013 about 2,000 companies offered their employees fitness trackers. In 2014 this rose to around 10,000, and companies such as FitBit now have dedicated partnerships with organisations to provide large numbers of trackers ad personalised data provision. The collection of such personal data and its use raises many ethical questions about how such data is used and who has access to the information - for example what happens to the employee who does not meet the fitness criteria defined by their employer? We discuss this later in the paper as it pertains to our case study and our opinions as researchers collecting such data, as well as employee reactions and 'buy in' to such initiatives.

3.2. Studies into Causes and Effects of Fatigue

The biometric measures we are proposing to capture, along with their meanings and effects, have been well studied in the field of psychology. Here we primarily focus on the following topics: definitions and effects of fatigue; activity, fatigue and recovery; fatigue and response times; heart-rate variability as an indicator of stress and fatigue.

Fatigue is typically classified into two general types, mental fatigue that affects an individual's cognitive processes and physical fatigue that affects an individual's ability to maintain physical actions. There is some contention over this division though with some researchers believing that fatigue is a single general state that is driven by physiological responses to energy expenditure of whichever category (Hockey and Ebrary 2013). Studies of physical fatigue typically require participants to undertake physically demanding tasks either for a pre-determined period of time or until they are unable to continue. A variety of measurements are compared pre- and post-task to evaluate the effect of the activity and extent of the fatigue. For example Kumar et al. (2004) measured oxygen uptake, ventilation, heart rate, blood oxygenation, blood volume and took electromyographic readings while subjects performed a physically demanding exercise, and reported a steady reduction in force exerted over the duration of the task.

Mental fatigue has been shown to affect task motivation (v. d. Linden et al. 2003) and high levels of mental fatigue have been show to result in a loss of efficiency and lower productivity of workers (Murata et al. 2005). Like physical fatigue, mental fatigue may be the result of fatiguing activities (cognitive processes) but it is also linked to disturbed, or lack of, sleep (Äkerstedt et al. 2002). While

---

[1] http://www.cnbc.com/2015/09/16/targets-fitbit-offer-to-workers-may-miss-its-mark.html
[2] http://www.forbes.com/sites/parmyolson/2015/10/20/fitbit-employers-barclays-godaddy-wellness/#6170061b3baa

physical fatigue can be measured by way of ability to exert force or perform activity (as above), measuring the effects of mental fatigue is less straightforward. One important (for our work) correlation that has been demonstrated is the effect on reaction times of individuals who are fatigued. Galton (1889) developed a simple reaction time (SRT) test which recorded a participant's response to a simple stimulus. This early test still forms the basis for several variations that have been developed to measure SRT and it is also used as the basis for the choice reaction time test (CRT) which records the time it takes a participant to choose a correct response from a number of alternatives. Of particular interest is the evidence showing that reaction time is adversely affected by both physical and mental fatigue (Brisswalter et al. 1997) suggesting we may see slower reaction times in physically demanding jobs, such as those found in forestry.

In addition to fatigue indicators such as SRT and CRT, there are changes in the autonomic nervous system when an individual is under stress (again both physical and mental). One key indicator that can identify this is heart-rate variability (HRV) which is the change in the inter-beat interval of the heart. A higher variability indicates higher levels of stress and corresponding fatigue and has been shown to be caused by work-induced cognitive stress (Chandola et al. 2008) as well as physical activity (Kaur et al. 2014). The increase in wearable technology capable of recording HRV has led to an increase in its use as a stress and fatigue measurement tool for athletes as well as ordinary individuals.

The majority of the studies described above, and many similar or complementary studies, are conducted in controlled environments (typically a laboratory setting) with specialised equipment and involve large-numbers of participants. This enables specific variables to be measured and controlled for required circumstances. For example, simulated driving laboratories can be used to investigate not just the fatiguing effects of driving in general, but rather the effects of particular driving conditions over specified periods for large numbers of test subjects (see for example Charlton and Baas (2006); Charlton and Starkey (2013)). Our intention is not try to replicate such studies or re-investigate known results from literature. Rather we want to find out if we can replicate the *results* of such studies using low-cost and light-weight measurement techniques in real-world settings. If we can do so, then we can rely on such technologies for in-situ monitoring of forestry workers with the confidence that the implications we draw from the measurements are based upon empirical studies conducted in controlled environments.

3.3. Using Sensors and IOT for Personal Monitoring

The Internet of Things is predominantly discussed in terms of a self-configuring network connecting objects in 'smart' homes and businesses, with a strong focus on the objects and environments. Typical applications are smart buildings,

smart homes, ambient intelligence, and mobile health-care. These applications assume large-scale and reasonably-stable computation and sensor constellations.

Even IoT applications that use rather fluid sensor constellations typically make these assumptions. For example, participatory sensing or crowd-sensing (Jaimes et al. 2015) are activities that engage the public to place sensors in regions of interest to gain large sample sizes. However, while the sensors, e.g. in the urban surface project (Kuznetsov and Paulos 2010), may be dynamically placed by the public, the urban computing environment itself is well-established and stable.

In contrast, our domain has a strong concern with the human body. Communication may be established via a Body Area Network (BAN) in collaboration with a Personal Area Network (PAN). Some smart home applications are treated as an extension of a body area network, e.g., for health-care applications (Gubbi et al. 2013). Again, most of these can rely on a stable network environment with which sensors can securely communicate. Body-centric systems use environmental, wearable, and implanted sensors. There are already a number of simple wearable technology applications, such as a T-shirt that visualises air quality monitoring results (Kim et al. 2010) or cycling helmets displaying heart-rate data (Walmink et al. 2014). In these cases, the sensing is instantly translated into the visualisation, without recording capability nor any links to the wearer's personal activity context. Some projects use RFID technology and IoT communication for personal health-care applications (Amendola et al. 2014) and for gathering information (temperature, humidity, and other gases) about the user's living environment. For example, Negi et al. (2011) and Adams et al. (2009) combine sensors with GPS to create a wearable personal air monitors. Wearable systems designed for outdoor use often rely on Bluetooth communication between GPS, sensors and smartphones (Honicky et al. 2008), possibly transferring data to central collection points via GPRS (Dutta et al. 2009). Inside buildings the use of GPS is limited and other indoor positioning systems are employed for location-based monitoring (Brown et al. 2016). Some of the applications focus on real-time monitoring of workers to protect them from environmental hazards, such as overexposure to air pollution (Fathallah et al. 2016).

Many of these health-related applications use IoT architectures that are akin to smart-city proposals, which are used to support people with disabilities (e.g., Domingo (2012)). Others use stand-alone body-focused systems, such as the Xbox Kinect. For example, González-Ortega et al. (2014) use 3D computer vision system for cognitive assessment and rehabilitation. These systems assume the support of powerful computing networks, often in a localised setting. Rohokale et al. (2011) proposed using a *cooperative IoT network* for rural health-care, which is

akin to ad-hoc wireless sensor networks in which each node acts as both sensor and relay. This work predominantly focuses on establishing communication with no concern for security or wearability of the equipment.

Scant attention has been paid to the communication and security of data when the IoT devices interact autonomously Roman et al. (2013). Secure and trustworthy computing typically focuses on resource-rich environments or hard security measures (e.g., encryption, signatures and certificates) that are typically energy hungry (e.g., Kumar and Madria (2015); Kothmayr et al. (2013)). Only a few studies address the specific challenges of trust in IoT (e.g., Bao and Chen (2012); Lacuesta et al. (2012)), and none take into consideration the characteristics of rugged environments, making their findings not applicable in such environments. Trust in IoT data collection at user level has recently received more public attention, with users becoming aware of the potential for mis-use of information collected (Brennan 2015). Limiting collection of data is difficult when safety requires data collection and devices record very personal data via sensors embedded in clothing or even implanted.

Interaction design in the IoT space makes it tempting to merely or overly focus on the objects – the 'things' in the Internet of Things (Jenkins 2015). Our problem domain has two interaction aspects: sensing of data and feedback to workers. In this paper, we consider the challenges of sensing, with a strong focus on the interplay between the objects and the humans involved.

4. THE RUGGED INTERNET OF THINGS (RIOT)

We believe that the new generation of lightweight, wearable technology and sensors of the Internet of Things (IoT) can help in identifying hazardous situations in work environments such as forestry, ultimately preventing fatalities. There are, however, many challenges in doing so. The use of IoT has already been embraced in some hazardous work environments, such as mining (Pye 2015). However most of these projects focus on specific environments where infrastructure is not an issue. Such ideas are built on the assumption of the continuous availability of computational power (in the form of cloud computing), high bandwidth (in the form of WiFi and cellular networks) and energy, since devices can be plugged in.

Many of those assumptions do not hold in rural, agricultural and forestry settings. Resources in these places may only be available intermittently. Such environments are characterised by lack of available bandwidth, computation constraints, energy constraints, and very importantly limited interaction between devices. Existing IoT technologies, which rely on the aforementioned assumptions, cannot cope in such rugged environments. For IoT to work successfully and safely in rugged environments

we must recognise that the standard assumptions do not work and provide alternatives. Our current experimental setup mixes fixed access points with temporary storage solutions (mobile phones) to ensure sensor data is not lost as workers move in and out of connectivity.

We therefore set out to explore the following three aspects:

Relevant data: What sort of data might be relevant in order to determine worker fatigue and unsafe situations?

Suitable Collection: Considering the data we might wish to collect, what are appropriate and effective ways of collecting such required data from forestry workers?

Analysis & use How can we analyse and use the obtained data (online or off-line) such that it would be of use in ensuring safety in dangerous work environments?

We present a summary of our analysis of relevant data to be collected in the next section. We then present in detail a case study that looks into suitable methods for collecting data in the forestry environment. Explorations of online use of the data have already begun and are also part of our future work.

5. RELEVANT DATA: ANALYSIS

As discussed above, our RIoT targets the problem of a 'smart landscape' in which disconnectedness and harsh operating conditions are the norm. Business and personal data are highly sensitive and possible interference through attempted data access or malicious data inserts have to be prevented. In addition, as it is known that workers may be suspicious about how any of the collected data is used, even if they agree to participate they may not comply. Therefore, spurious or suspicious data may be the result of either malicious interference from external entities or due to worker disruption.

Data we wish to collect can be categorised across several different dimensions. First, we consider the privacy considerations for the data and whether it should be considered as:

- *personal [P]* – only the owner should have access to the unconsolidated and unanonymised data, identification should only be possible in specified emergency scenarios
- *business-sensitive [B]* – needs to be concealed from external entities as it may reveal properties of the work environment that can be considered commercially sensitive

Secondly we consider the requirement for when data should be available for collection (frequency and availability) which can also be divided into two groups:

- *continuous [C]* – where it is essential that data is collected in an uninterrupted manner

- *infrequent [I]* – data may be provided at varying intervals throughout the day

Table 1 gives an overview of the proposed data to be collected and its categorisations. These data categories were developed based on prior work with high-performance athletes Tavares et al. (2016) and industry engagement with forestry workers Bowen et al. (2015a,b); Griffiths (2016), as well as analysis of relevant literature.

Data	P	B	C	I
Activity	X			X
Ambient temperature				X
Breathing	X	X	X	
Calorific burn	X			X
Core body temperature	X	X		X
Heart-rate variability	X	X	X	
Hill climbing	X			X
Location			X	
Reaction times	X	X		X
Sleep data	X	X		X
Vibration	X	X	X	

Table 1: *Data Categories*

6. SUITABLE COLLECTION: CASE STUDY

In order to explore suitable data collection methods for worker-related fatigue data in the outdoor forestry environment, we designed a case study that comprises three phases (see Table 2). The table provides an overview of collected data for each phase, participants and length of the study. So far, we have conducted Phases I and II of our case study, and are currently undertaking Phase III. We here first describe the case study setup and report on results of Phases I and II; it is too early to report results on Phase III – it is shown here only for completeness. We then discuss our observations, the challenges encountered and lessons learnt from the case study.

We begin by summarising each of the phases wrt. their goals, aims and participants. We then discuss the challenges and problems that were identified by the studies and how they relate to our initial problem statement. We conclude this section with a discussion of the insights obtained from our results and how they contribute to our proposed solution.

6.1. Case Study Methodology

We started by looking at how we might predict hazards by harnessing the power of a new generation of lightweight, wearable technology (such as activity trackers). We subsequently investigated different types of sensors (and wearable sensors such as those found within the LifeBEAM Smart Hat[3] for example) within an IoT.

Phase 1. This phase started with a number of experiments performed by the research team over varying periods of time to investigate the properties of activity tracker usage.

---

[3] http://life-beam.com/shop/smart-hat/

Phase	Description	Length	#Participants
I	Activity, steps and sleep patterns	16 + 5 weeks	4 + 2
II	Activity, steps and response times	3 + 3.5 weeks	1 + 15
III	Heart-rate variation, physical and mental fatigue activities	ongoing	

***Table 2:** Case Study Phases (data on external participants participants + forestry workers)*

The aim of this initial research was to discover any usability issues that might occur when using activity trackers for long-term studies (so requiring them to be worn 24/7) as well as investigating the effects of domain-specific activities (such as operating a chain-saw, driving long distances on bumpy roads, walking in forestry environment etc.) Three researchers wore devices (sometimes more than one at a time for comparison purposes) to track both daily activity and sleep for periods of 6–14 weeks. In addition one researcher kept a diary throughout the same period to enable consideration of particular data points. We then replicated this data collection with two forestry workers over a five week period. The participants were asked to wear an activity tracker (each had a different brand) 24 hours a day, initially for a period of two weeks, and then subsequently for another two weeks, then for one final week. The extensions were due to ongoing problems with the participants use of the trackers (discussed shortly).

Phase 2. The second phase began by one of the researchers undergoing a period of self-monitoring on activity and fatigue levels. Activity was measured based on step-counting, heart-rate and calorific burn (using a Fitbit HR), sleep was measured in terms of quantity and quality (again using the Fitbit HR) and the effects of fatigue were measured based on reaction times. Reaction time testing used two methods, *simple-reaction time* testing (SRT) and *choice-reaction time* testing (CRT). For this study, reaction time was measured using the 'Reaction Time' application[4] designed to measure the time taken to respond to visual stimulus (colour change) and screen touch (response). These experiments ran for a three week period and encompassed activities undertaken by a researcher, covering a mixture of workplace and study activity.

Again we then moved our data collection to in-situ forestry workers. This involved the collection of physiological data by means of a Fitbit Charge HR wrist worn monitoring device and testing of SRT and CRT. Simple and choice reaction time measurements were undertaken at commencement of the participants' work period, during their break time and on completion of the participants' work day. The Deary-Liewald Reaction Time Task application developed by the Centre for Cognitive Ageing and Cognitive Epidemology at the University of Edinburgh Deary et al. (2011) was used for this purpose. SRT testing was completed first with each participant undertaking 15 individual tests. CRT testing was performed

secondly with participants undertaking 20 individual tests. Participants were selected from three work crews based at three separate locations (all members of a crew were included where possible) who were each monitored for two periods of 3–5 days. In total there were fifteen participants who were all male, with ages between 17 and 62. Five participants were loader operators, three worked in quality control, three were manual tree fellers, two were process operators and two were log makers.

Crew 1 was a fully mechanised crew with most operations being carried out using plant and machinery. Crew 2 was a primarily manual crew with most operations being performed by workers on the ground (using chainsaws etc.) Crew 3 was a hauler crew who operate in steep terrain using cabling techniques to drag felled trees to the skid site for processing. The working day for all three crews (apart from loader operators who typically started the day 2 hours earlier and then broke for the morning meeting) starts with a meeting where tasks are assigned. Directly after this meeting we issued each participant with a Fitbit HR to wear for the day and performed the first of the reaction time tests. Work then commences for around four hours, at which point a 45 minute break occurs when we performed the second reaction time test. Work then re-commences until the end of the day when we performed the final reaction time test and collected the activity trackers for data synchronising and charging.

6.2. Summary of Results

We here summarise the case study results wrt. the data quality and suitability of the data collection method.

Phase 1. While the full results from Phase I were reported in Bowen et al. (2015a), here we discuss the key findings from both phases and show how they relate to the aims of the work described. Our experiences of Phase I showed that context-free data can be misleading (high activity levels may not be due to steps but other actions such as driving or even drinking). There were large variances in sleep-tracking accuracy (when compared with diary reports as well as measured against state of the art devices like the Readiband[5]) and devices can get in the way of some activities (uncomfortable when typing on a keyboard for long periods or irritating to the skin overnight). This suggests that choice of technology needs to be based on a number of factors and that data needs to be correlated with other variables in order to obtain a clearer picture of its meaning. While different types of tracker reported different

---

[4]available from Google Play `https://play.google.com/store/apps/details?id=com.chingy1788.reactiontime`

[5]`www.fatiguescience.com/Readiband/sleep_tracker`

values for activity tracking the differences remained consistent (so trends seemed accurate) which means that factors such as comfort and utility in the environment can be used as the dominant choice factor.

The most significant finding from the forestry worker engagement in Phase I is the level of technical difficulties that occurred—participants could not change the mode of the devices to track sleep; devices were lost; mode changes occurred frequently so that data collection was compromised; participants never charged the devices; connectors for uploading data were lost. Of course, it is possible (and indeed likely) that not all of these problems were actually technical but that there were also elements of resistance from the participants to being monitored in this fashion. As we have discussed earlier, in some sense this is not surprising, the monitoring of workers during their private time is potentially controversial. Even though our participants were volunteers and keen to take part, during the initial meeting to set up the study it was clear that there were reservations about some aspects of the monitoring (particularly the sleep monitoring) and what could be identified from the data).

Workers were assured that only anonymous and aggregated data would be available to their boss; however, the fact that their boss might be able to see any data caused considerable concern and might have been the reason for the subtle disruptions and signs of non-compliance we observed. A similar observation was made by Kortuem et al. Kortuem et al. (2007), as they explored organisational issues of industrial health and safety monitoring system. They had also observed "both a perceived lack of trust and a lack of effective two-way communication between management and operatives".

Phase 2. As the first part of Phase II involved self-monitoring, and because we had already learnt some of the lessons relating to equipment choice from Phase I, there were overall fewer issues. Our focus for Phase 2 was on identifying correlations between activity and reaction times. Specifically we wanted to see if the data we could collect would correlate with known properties of activity and fatigue (as discussed earlier). Speed of mental processing (SMP) is a means to aggregate data from simple and choice reaction time ($SMP = CRT - SRT$). A summary of the collected data for Crew 3 is shown in Figure 3. We found that the effect of activity on reaction time varied between participants. SRT showed no common pattern other than a tendency towards being slightly slower at the end of the day than at the start of the day but between those points there was no consistent effect. CRT similarly had no common pattern. The data is personalised but may also depend on role types. For example, the loader drivers 3E and 3D both show improved mental processing times as well as CRT and SRT throughout the day. People were typically found to be consistent within each day but very different from each

Table 3: Crew 3 Mean reaction times in msec

Period	Participant	Role	CRT	SRT	SMP
Start	3A	Quality Control	662	701	-39
Break			649	382	267
End			525	390	135
Start	3B	Quality Control	624	445	179
Break			544	363	181
End			534	396	148
Start	3C	Manual Feller	566	392	174
Break			566	386	180
End			569	357	212
Start	3D	Loader Operator	520	311	209
Break			456	321	135
End			458	277	181
Start	3E	Loader Operator	1067	1016	51
Break			820	569	251
End			742	424	318

other. We also observed that not only physical exhaustion may contribute to fatigue but also mental activity.

Implications for data collection As a result of Phase I we made a decision that we would not continue with sleep-tracking of the workers; instead it was decided to focus on measuring the effects that fatigue (based on activity) might have. It may be that including some type of self-reporting question regarding sleep quality at the start of the electronic reaction time test may be useful, although we should be mindful that the same ethical issues that could lead to non-compliance with the sleep monitoring may similarly affect the answers given. If workers believe they may be penalised (e.g. sent home as unfit to work) if they select an answer indicating they have slept poorly several nights in a row then they may be reluctant to provide such an answer.

The need for the researcher to be on site several times a day to facilitate the testing also meant that they were able to observe some aspects of the environment which proved to be informative, despite not being part of the overall study plan. The effect of temperature appears to have a bigger impact on reaction time (both SRT and CRT) than activity alone. The biggest effect was seen at temperatures less than 4°C with the mean differential across all participants between 2°C and 4°C on CRT being over 100 milliseconds and on SRT being around 50 milliseconds. However, this needs to be considered in conjunction with the activity data as typically the coldest part of the day is when work commences and as the temperature rises the amount of activity having been performed also increases.

The remote working conditions for each of the forestry crews and lack of facilities contributed to the effect of ambient temperature as an important variable. Work sites are at remote locations with the only welfare facilities available to the crew being the vehicles they travel to work in. There are no fresh water or toilet facilities at any of the locations. Crew 1 had access to a metal shipping container that is used as both the site office and lunch room. This

container travels with the crew from production site to production site as the crew moves around. This has an effect on workers' choice of hydration and food throughout the day. In addition the lack of a comfortable, warm, area to take breaks in, and no running water or power meant that breaks were taken as required in a perfunctory manner, rather than being used as an opportunity for workers to relax, make tea/coffee, heat up food, socialise etc. as might be seen in indoor working environments. Power is not supplied to the container and as such no heating is available and crews protect themselves from adverse temperatures by use of clothing layers. Machine operators who spend large parts of the day sitting in unheated machine cabs are particularly affected by cold temperatures, whereas the workers on the ground report finding the summer months where they are unprotected from the heat more physically challenging. Our results also indicated that mental fatigue (again seen by machine operators) appears to have a stronger effect than that of physical activity. It is clear, therefore, that all of these need to be carefully measured and considered in larger scale data-gathering activities.

Finally the individual nature of the results seen (particularly the differences in reaction time across workers) indicates that there can be no overall general benchmark applied to determine whether or not a worker is fatigued or has reduced reaction time. Rather we need individual data collected over time to act as a personal benchmark, so deviation from an individual's normal pattern of data is what is important. Again this emphasises the need to build personal data histories as a mechanism for predicting future behaviours. The requirements for collecting, storing and analysing this type of personal data, as well as incorporating environmental and business data leads to a number of further considerations that we discuss next.

7. DISCUSSION

The personal stories we uncovered and the insights gathered while doing these studies suggest that the design considerations are not those we traditionally prioritise. Any system used for monitoring workers, even when their safety is our primary concern, must be focussed on their privacy as much as anything else. It may be that we will not create the *most effective* IoT solution or will not include the *most optimal* data inputs but rather we aim to find the most *ethical, robust and secure* solution that can do the required job. There are many philosophical, ethical & social implications of collecting and using this data. Workers already do hard jobs for minimum wage, if we collect data that deems them not fit to work then what happens? There are two distinct parts to our focus, the first is that any data collected is shared (or hidden) appropriately. The second is that workers are kept safe and well.

In Phase I, one of the workers was off sick for a day and at the same time problems were encountered with his monitoring device. This again suggests that there is a fine line between what is acceptable and what is not when it comes to such personal monitoring. It is also clear that it may not always be obvious where a problem lies and if a technology error really is that or if something else is at play. As computer scientists it is tempting to focus on the things we *can* do rather than the things we *should* do. If one could market such monitoring solutions as being beneficial to workers, or the elderly or the disabled (as we see with many similar monitoring approach rationales) then we may stray into dangerous areas. Risk assessment for poorly paid workers in dangerous environments is clearly a good thing, but if the data is used to send home, or lay off, workers who do not meet the new risk criteria then we must consider the responsibilities we have in this. This is especially true when it is not as simple as saying "worker X is fatigued and will cause a serious accident if he stays at work"; the interplay of various factors is much more subtle than that. While high-performance athletes seem to accept the "lab-rat" lifestyle where all aspects of their performance may be monitored both in and out of work this comes with benefits for them which allow them to improve and attain higher standards. However most research in this area focusses on the different metrics or studies that can be used to fine-tune athletic performance rather than consider the effect this has on those being studied—particularly when under-performance or lack of adherence to training and nutrition schedules is suggested.

Similarly the choice of the components we include in our RIoT solution must be carefully considered. Off-the-shelf sensors and tracking equipment is appealing because of its availability and low cost (easy to deploy quickly to large numbers of people), but much of this is not designed to be secure or private. We have been experimenting with a hat that includes a built-in heart rate monitor which just broadcasts its data via Bluetooth continually and which can potentially be captured by anyone or anything in close enough proximity. Developers of such artefacts for personal use are not typically concerned with such data leakage which may not seem concerning when geared towards lifestyle and fitness. Even presuming we do collect data in a secure and private manner (see Table 1), we still need to ask the questions about who then has access to this data (the workers themselves, their bosses, health and safety bodies etc.) and how it is presented and used. We must be clear about our proposed use of this data and ensure that it cannot be accessed and used for other purposes (performance management of employees for example). This again requires us to treat the data in ways that may not necessarily be the most optimal in terms of the technological solution but which ensures the ethical dimension is acknowledged.

8. SUMMARY AND FUTURE WORK

In this paper we have discussed our attempts at creating continuous data-collection methods that are ethical,

sustainable and effective. We explore the difficulties in collecting personal and environmental data from workers and their work domain and discuss both the technical and ethical issues that arise. We have also presented a case-study that explores data collection methods as part of a robust, trusted Internet of Things (IoT) solution for dangerous work environments. We described considerations for relevant data and suitable collection methods, while exploring the use of lightweight sensors to monitor worker activity levels and response times as fatigue indicators.

As we discussed earlier, the nature of the data we wish to observe and gather includes highly sensitive personal data as well as business-sensitive information. This naturally requires safe methods for storage and communication among trusted partners. However, the very nature of the rugged environments we are working in means that connected devices may be transient due to power limitations, movement in and out of connectivity etc. This means that security and trust must be dealt with dynamically, and the IoT includes small lightweight sensors that do not have the capacity for on-board security. Not only does this add a layer of overhead that does not exist in typical IoT in terms of management of connectivity, it also means that trust cannot develop over time as in established IoT networks. Another element of our research, therefore, is in developing a trust model that can support this type of dynamic connectivity and can react accordingly if devices disappear or newly enter. The model should also be able to differentiate between anomalous and maliciously inserted false data. Reliance on redundancy to partially solve this problem is not necessarily suitable in a rugged environment where the infrastructure to support just the minimum required connectivity is already challenging.

We propose to protect the data through the use of a trust management system. Typically trust amongst IoT components is either confirmed by third parties (which are not available in our setting) or is developed over time (which is not suitable in our dynamic outdoor environment). We use data aggregation and composition to derive valuable safety-related information from the collected sensor data. Our trust model will analyse the data from neighbouring nodes in the IoT and classify the data as acceptable or malicious. We started with wearable technologies as a proof-of-concept for our data collection but want to go beyond just collecting data so propose to develop a sensor-based IoT which also provides feedback to workers via wearable technology such as a *smart vest*.

The next stage for our work is completing Phase III of our case study in which we experiment with monitoring heart-rate variability data and recognition of mental and physical fatigue activities. This will be incorporated into sensors that will be part of our RIoT network setup. This will then be employed in field tests to see how well the RIoT network performs and to identify potential areas of

weakness in the security aspects. We can then move on to analysing the data itself and consider how we will use them in-situ in useful ways to help reduce risk.

REFERENCES

Adams, C., Riggs, P., and Volckens, J. (2009). Development of a method for personal, spatiotemporal exposure assessment. *Journal of Environmental Monitoring*, 11(7):1331–1339.

Adams, G., Armstrong, H., and Cosman, M. (2014). Independent forestry safety review – an agenda for change in the forestry industry.

Äkerstedt, T., Fredlund, P., Gillberg, M., and Jansson, B. (2002). Work load and work hours in relation to disturbed sleep and fatigue in a large representative sample. *Journal of Psychosomatic Research*, 53(1):585 – 588.

Amendola, S., Lodato, R., Manzari, S., Occhiuzzi, C., and Marrocco, G. (2014). RFID technology for iot-based personal healthcare in smart spaces. *IEEE Internet of things journal*, 1(2):144–152.

Bao, F. and Chen, I. (2012). Dynamic trust management for internet of things applications. In *Workshop on Self-aware Internet of things*, pages 1–6. ACM.

Botan, C. (1996). Communication work and electronic surveillance: A model for predicting panoptic effects. *Communication Monographs*, 63:294–313.

Bowen, J., Hinze, A., Cunningham, S., and Parker, R. (2015a). Evaluating low-cost activity trackers for use in large-scale data gathering of forestry workers. In *Australian SIGCHI Conference (OZCHI)*, pages 474–482. ACM.

Bowen, J., Hinze, A., and Cunningham, S. J. (2015b). Into the woods. In *INTERACT*, pages 137–143. Springer.

Brennan, J. (2015). 5Ps of the Internet of Things: A user-centric approach. Keynote presentation at the European Workshop for Trust & Identity.

Brisswalter, J., Arcelin, R., Audiffren, M., and Delignieres, D. (1997). Influence of physical exercise on simple reaction time: effect of physical fitness. *Perceptual and Motor Skills*, 85:1019–1027.

Brown, K., Shaw, P., Mead, K., Kovein, R., Voorhees, R., and Brandes, A. (2016). Development of the chemical exposure monitor with indoor positioning for workplace VOC surveys. *J. of occupational and environmental hygiene*, 13(6):401–412.

Chandola, T., Britton, A., Brunner, E., Hemingway, H., Malik, M., Kumari, M., Badrick, E., Kivimaki, M., and Marmot, M. (2008). Work stress and coronary heart disease: what are the mechanisms? *European Heart Journal*, 29(5):640.

Charlton, S. and Baas, P. H. (2006). Fatigue, work-rest cycles, and psychomotor performance of New Zealand truck drivers. *New Zealand Journal of Psychology*, 30(1):32 – 39.

Charlton, S. G. and Starkey, N. J. (2013). Driving on familiar roads: Automaticity and inattention blindness. *Traffic Psychology and Behaviour*, 19:121 – 133.

Deary, I. J., Liewald, D., and Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods*, 43(1).

Domingo, M. (2012). An overview of the Internet of Things for people with disabilities. *Journal of Network and Computer Applications*, 35(2):584–596.

Dutta, P., Aoki, P., Kumar, N., Mainwaring, A., Myers, C., Willett, W., and Woodruff, A. (2009). Common sense: participatory urban sensing using a network of handheld air quality monitors. In *Embedded networked sensor systems*, pages 349–350.

Fathallah, H., Lecuire, V., Rondeau, E., et al. (2016). An iot-based scheme for real time indoor personal exposure assessment. In *Consumer Communications & Networking Conference (CCNC)*, pages 323–324. IEEE.

Galton, F. (1889). An instrument for measuring reaction time. *Report of the British Association for the Advancement of Science*, 59:784–785.

González-Ortega, D., Díaz-Pernas, F., Martínez-Zarzuela, M., and Antón-Rodríguez, M. (2014). A kinect-based system for cognitive rehabilitation exercises monitoring. *Computer methods and programs in biomedicine*, 113(2):620–631.

Griffiths, C. (2016). Investigating tools and methods for data capture of forestry workers. Master's thesis, University of Waikato.

Gubbi, J., Buyya, R., Marusic, S., and Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7):1645–1660.

Hockey, R. and Ebrary, I. (2013). *The psychology of fatigue: Work, effort and control*. Cambridge University Press.

Honicky, R., Brewer, E., Paulos, E., and White, R. (2008). N-smarts: networked suite of mobile atmospheric real-time sensors. In *Networked systems for developing regions*, pages 25–30.

Human Work Interaction Workshop (HWID) (2015). position papers. available online at `https://projects.hci.sbg.ac.at/hwid2015/position-papers`.

Jaimes, L., Vergara-Laurens, I., and Raij, A. (2015). A survey of incentive techniques for mobile crowd sensing. *IEEE Internet of Things Journal*, 2(5):370–380.

Jenkins, T. (2015). Designing the things of the iot. In *ACM Tangible, Embedded, and Embodied Interaction*, pages 449–452. ACM.

Kaur, S., Bhalla, B., Bajaj, S. K., Sanyal, S., and Babbar, R. (2014). Effect of physical and mental stress on heart rate variability in type-A and type-B personalities. *Indian Journal of Applied Basic Medical Sciences*, 15(20):59–70.

Kim, S., Paulos, E., and Gross, M. (2010). WearAir: expressive t-shirts for air quality sensing. In *ACM Tangible, embedded, and embodied interaction*, pages 295–296. ACM.

Kortuem, G., Alford, D., Ball, L., Busby, J., Davies, N., Efstratiou, C., Finney, J., White, M. I., and Kinder, K. (2007). *Sensor Networks or Smart Artifacts?*, pages 465–482.

Kothmayr, T., Schmitt, C., Hu, W., Brünig, M., and Carle, G. (2013). DTLS based security and two-way authentication for the Internet of Things. *Ad Hoc Networks*, 11(8):2710–2723.

Kumar, S., Amell, T., Narayan, Y., and Prasad, N. (2004). Measurement of localized muscle fatigue in biceps brachii using objective and subjective measures. *Muscle Strength*, pages 105–121.

Kumar, V. and Madria, S. (2015). Distributed attribute based access control of aggregated data in sensor clouds. In *Reliable Distributed Systems (SRDS), 2015 IEEE 34th Symposium on*, pages 218–227. IEEE.

Kuznetsov, S. and Paulos, E. (2010). Participatory sensing in public spaces: activating urban surfaces with sensor probes. In *ACM Conference on Designing Interactive Systems*, pages 21–30. ACM.

Lacuesta, R., Palacios-Navarro, G., Cetina, C., Peñalver, L., and Lloret, J. (2012). Internet of things: where to be is to trust. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):1–16.

Lilley, R., Feyer, A., Kirk, P., and Gander, P. (2002). A survey of forest workers in New Zealand. Do hours of work, rest, and recovery play a role in accidents and injury? *Journal of Safety Research*, 33(1):53–71.

Murata, A., Uetake, A., and Takasawa, Y. (2005). Evaluation of mental fatigue using feature parameter extracted from event-related potential. *International journal of industrial ergonomics*, 35(8):761–770.

Negi, I., Tsow, F., and Tanwar, K. (2011). Novel monitor paradigm for real-time exposure assessment. *Journal of Exposure Science and Environmental Epidemiology*, 21(4):419–426.

Parker, R. (2010). *Technological advances in the analysis of work in dangerous environments: Tree felling and rural fire fighting*. PhD thesis, Massey University, Palmerston North, New Zealand.

Pye, A. (2015). Mining companies embrace the IoT. *Engineering and Technology Magazine*, 10(5).

Rohokale, V., Prasad, N., and Prasad, R. (2011). A cooperative IoT for rural healthcare monitoring and control. In *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, pages 1–6.

Roman, R., Zhou, J., and Lopez, J. (2013). On the features and challenges of security and privacy in distributed Internet of Things. *Computer Networks*, 57(10):2266–2279.

Tavares, F., Healey, P., Smith, T. B., and Driller, M. W. (2016). Effect of training load on acute fatigue and wellness during an in-season non-competitive week in elite rugby athletes. In *Sport & Exercise Science New Zealand Annual Conference*.

v. d. Linden, D., Frese, M., and Meijman, T. (2003). Mental fatigue the control of cognitive processes. *Acta Psychologica*, 113(1):45–65.

Walmink, W., Wilde, D., and Mueller, F. (2014). Displaying heart rate data on a bicycle helmet to support social exertion experiences. In *ACM Tangible, Embedded and Embodied Interaction*, pages 97–104.