

DEFINITION OF ALARM FATIGUE AND ITS INFLUENCE ON STAFF
PERFORMANCE

by

Shuchisnigdha Deb

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Industrial and Management Engineering

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2014

©COPYRIGHT

by

Shuchisnigdha Deb

2014

All Rights Reserved

ACKNOWLEDGEMENTS

I would love to thank Dr. David Claudio for his constant support and inspiration throughout this project. I am also grateful to Dr. Nicholas Ward, Dr. Maria Velazquez and Dr. Frank Marchak for agreeing to be part of my committee, and for their support. Support from my two co-researchers, Rafael Dias and FNU Pawan Kumar were also significant for this study. I would like to thank all the ICU staff at BDH for their patience and support with my constant questions and observations. My gratitude also goes to Andrea Roberts, from quality department at BDH for making their staff and workplace so available to accomplish this project successfully. I would like to thank my editor Ms. Iris Driscoll for all her hard work and time. My parents, sisters, husband and brother-in-laws have been a constant source of motivation, and made it entirely possible for me to be here. I am thankful to my two daughters who have sacrificed a lot. I would specifically like to thank my friends Mahmudur Rahman and Tanjila afrin; without these people, this was not possible for me.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
Research Questions	6
2. LITERATURE REVIEW.....	7
Alarm Fatigue.....	8
Personality Type.....	17
Mental Workload.....	21
Affect.....	24
Boredom.....	24
Apathy.....	25
Distrust.....	26
Performance Measures	27
Literature Review Summary	27
3. METHODOLOGY	30
Approval for the Study	31
Participants	31
Location.....	32
Procedure.....	33
Health Technology Foundation Survey	34
Hierarchical Task Analysis	35
Observational Data Collection.....	37
Mental Workload Measurement	38
NASA TLX Measures:	39
SWAT Measures	40
Affect Measurement.....	42
Boredom.....	42
Apathy	43
Distrust.....	43
Performance Measures.....	44
Statistical Analyses	44
4. RESULTS.....	47
Clinical Alarm Survey.....	47
Hierarchical Task Analyses.....	49
Descriptive Statistics	52

TABLE OF CONTENTS (CONTINUED)

Research Question 2- Definition of Alarm Fatigue as Mental Workload.....	54
Regression Analyses to Define Alarm Fatigue as Mental Workload	54
Research Question 3- Definition of Alarm Fatigue as Affect	57
Regression Analyses to Define Alarm Fatigue as Affect	57
Principal Component Analyses for Alarm Fatigue Causing Factors	60
Research Question 4- Comparison between Two Definitions	62
Research Question 5- Significance of Alarm Fatigue Measures on Staff Performance	64
Relationship between the Alarm Fatigue Measures and the Performance Measures.....	65
Regression Analyses for Staff Performances Using Alarm Fatigue Measures:	65
Principal Component Analyses for Alarm Fatigue Measures.....	67
Regression Analyses for Staff Performances Using Alarm Fatigue Causing Factors.....	68
Regression Using Combined Factors.....	71
Principal Component Analyses for Combined Factors.....	74
5. DISCUSSION	76
Assessment of Existing Alarm Monitoring System	77
Objective 1- Identification of all the Possible Influential Factors for Defining Alarm Fatigue.....	79
Hierarchical Task Analysis for Unit Clerks.....	80
Hierarchical Task Analysis for Nurses	81
Descriptive Statistics for Alarm Fatigue Measures and Performance Measures	83
Objective 2- Definition of Alarm Fatigue in terms of Mental Workload	86
Objective 3- Definition of Alarm Fatigue in terms of Affect.....	89
Objective 4- Comparison between Two Definitions.....	90
Significance of Shift	92
Significance of Staff ID	93
Significance of Workload	94
Significance of Number of Alarms	95
Significance of Alarm Type.....	96
Significance of Alarm Criticality.....	97
Significance of Noise Level.....	97
Significance of Time Elapsed	98
Significance of Task Priority	99
Significance of Extraversion.....	100
Significance of Openness.....	101

TABLE OF CONTENTS (CONTINUED)

Significance of Agreeableness	101
Significance of Conscientiousness.....	102
Significance of Neuroticism	102
Objective 5- Significance of Alarm Fatigue on Staff Performance	103
Unsuccessful Attempts to Relate Alarm Fatigue to Staff Performance.....	103
Successful Models which Define Factors Related to Staff Performance	105
Significance of Shift	109
Significance of Staff ID	110
Significance of Workload	110
Significance of Number of Alarms	110
Significance of Alarm Type.....	111
Significance of Alarm Criticality.....	111
Significance of Time Elapsed	111
Significance of Task Priority	112
Significance of Noise Level.....	112
Significance of Extraversion.....	112
Significance of Agreeableness.....	112
Significance of Neuroticism	113
Significance of Openness.....	113
Significance of SWAT Measures.....	113
Significance of NASA TLX Measures	114
Significance of Boredom	114
Significance of Apathy	114
Significance of Distrust.....	115
6. CONCLUSIONS	116
REFERENCES CITED.....	123
APPENDICES	131
APPENDIX A: Approvals for the Stusy	132
APPENDIX B: Clinical Alarm Survey	136
APPENDIX C: Observational Data Collection Table.....	139
APPENDIX D: Alarm Fatigue Measuring Scales.....	142
APPENDIX E: Descriptive Statistics for Nurses	148
APPENDIX F: Correlational Analyses	150
APPENDIX G: Principal Component Analyses (PCA).....	155
APPENDIX H: Significance of Individual Factors.....	162

LIST OF TABLES

Table	Page
1. Definition of technical terms related to clinical alarms	9
2. Summary of literature review	15
3. Dimensions for MBTI.....	20
4. Dimensions for Big Five Factors	20
5. Demographics summary of participants	32
6. The summary of the steps that were conducted to perform HTA in the ICU	36
7. Rating scale definitions for the six dimensions of NASA TLX	40
8. Rating scale definitions for the six dimensions of SWAT.....	41
9. Factors and measures considered in this study	44
10. Responses by percentages to the alarm related information	47
11. Descriptive statistics for alarm fatigue measures and performance measures	52
12. Summary of regression analyses to define alarm fatigue as SWAT measures	55
13. Summary of regression to define alarm fatigue as NASA TLX measures.....	56
14. Summary of regression analyses to define alarm fatigue as Boredom	58
15. Summary of regression analyses to define alarm fatigue as Apathy	59
16. Summary of regression analyses to define alarm fatigue as Distrust	60
17. Summary of regression with components extracted from PCA.....	61

LIST OF TABLES (CONTINUED)

Table	Page
18. Summarization of influencing variables from two definitions pertaining to unit clerks	63
19. Summarization of influencing variables from two definitions pertaining to nurses	64
20. Summary of regression for Response (yes/no) using alarm fatigue measures.....	65
21. Summary of regression for Response Time using alarm fatigue measures	65
22. Summary of regression for Number of Ignored Alarms using alarm fatigue measures.....	66
23. Summary of regression with components extracted from PCA.....	67
24. Summary of regression for Response using alarm fatigue causing factors	68
25. Summary of regression for Response Time using alarm fatigue causing factors	69
26. Summary of regression for Number of Ignored Alarms using alarm fatigue causing factors	69
27. Summarization of all the influencing variables for performance measures.....	70
28. Summary of regression for Response (yes/no) using combined factors.....	72
29. Summary of regression for Response Time using combined factors	72
30. Summary of regression for Number of Ignored Alarms using combined factors.....	73
31. Summary of regression with components extracted from PCA combined factors.....	74

LIST OF TABLES (CONTINUED)

Table	Page
32. Summarization of all the influencing variables from regression with combined factors to relate alarm fatigue to staff performances.....	75
33. summary of the significances of influencing factors on the alarm fatigue measures	92
34. summary of the significances of influencing factors on staff performance	109

LIST OF FIGURES

Figure	Page
1. Conceptual model for alarm fatigue and its influence on staff performance	30
2. Representation of the ICU at BDH	33
3. Responses to the issues concerning alarms.....	49
4. Hierarchical Task Analysis for unit clerks in ICU.....	50
5. Hierarchical Task Analysis for nurses in ICU	51
6. Effect of alarm fatigue on alarm response for unit clerks.....	54
7. Diagnostics plots for SWAT measures with nurses.....	56
8. Diagnostics plots for NASA TLX measures with nurses	57
9. Diagnostics plots for Boredom in nurses	58
10. Diagnostics plots for Apathy in nurses	59
11. Diagnostics plots for Distrust in nurses	60
12. Approaches to understanding the role of alarm fatigue	77
13. Primary tasks distribution for unit clerks in ICU during observed shifts	80
14. Primary tasks distribution for nurses at ICU during observed shifts	82
15. Association of fatigue causing factors with mental workload measures showing R^2 values inside arrows for both groups of participants.....	87
16. Association of alarm causing factors with affect showing R^2 values inside arrows for both groups of participants	90
17. Unsuccessful models to associate alarm fatigue with staff performance for unit clerks	104
18. Unsuccessful models to associate alarm fatigue with staff performance for nurses.....	105

LIST OF FIGURES (CONTINUED)

Figure	Page
19. Successful models to associate alarm fatigue with staff performance for unit clerks	106
21. Proposed models on the association of alarm fatigue to staff performance for unit clerks	108
22. Proposed models on the association of alarm fatigue to staff performance for nurses	108
23: Conceptual model for unit clerks.....	120
24: Conceptual model for nurses	121

NOMENCLATURE

ECRI	Emergency Care Research Institute
JCAHO	Joint Commission on the Accreditation of Healthcare Organizations
TJC	The Joint Commission
ACCE	American College of Clinical Engineering
FDA	Food and Drug Administration
MAUDE	Manufacturer and User Facility Device Experience
AAMI	Association for the Advancement of Medical Instrumentation
HTF	Healthcare Technology Foundation
NPSG	National Patient Safety Goals
ICU	Intensive Care Unit
BDH	Bozeman Deaconess Hospital
CRCU	Coronary Respiratory Care Unit
HTA	Hierarchical Task Analysis
NASA TLX	National Aeronautics and Space Administration-Task Load Index
SWAT	Subjective Workload Assessment Technique
SOFI	Swedish Occupational Fatigue Inventory
MBTI	Myers Briggs Type Indicator
PCA	Principal Component Analysis
RMSE	Root Mean Square Error

ABSTRACT

An alarm is a warning of an approaching situation which requires a response. This study considered the influences of alarms in the clinical health environment. The Emergency Care Research Institute considered alarm hazard as the number one health technology hazard for the years 2012 through 2014. The Joint Commission set a standard for all hospitals in the US to assess alarm fatigue in their monitoring process and to develop a systematic, coordinated approach to clinical alarm system management. In order to comply with this requirement, a working definition of alarm fatigue is necessary. To the very best of our knowledge, there have been no studies proposing a quantitative definition of alarm fatigue, a way to measure it, or exploring the influence of alarm fatigue on performance deterioration. This observational study undertook the objective of defining alarm fatigue and its role on performance deterioration in a quantitative way. A survey using a questionnaire proposed by the American College of Clinical Engineering Healthcare Technology Foundation was conducted before the observations to assess the attitude of nurses toward the existing alarm monitoring system at the hospital where the research took place. An extensive literature review and Hierarchical Task Analyses were conducted in order to reveal all the possible influencing factors behind alarm fatigue. From these, alarm fatigue was defined and measured in terms of mental workload and three types of affect: boredom, apathy, and distrust. A conceptual model was developed considering the significance of working conditions and staff individuality on alarm fatigue and, consequently, alarm fatigue on staff performance. Staff performance was measured in terms of response (yes/no), response time and number of ignored alarms. Several analytical approaches were performed to find association between alarm fatigue and staff performance. The results show that, in general, performance deterioration is actually influenced by a combination of alarm fatigue with working conditions and staff individuality. In the case of nurses and response time, alarm fatigue plays no role, only working conditions and staff individuality. These findings suggest that the role of alarm fatigue as a health hazard in the clinical environment should be reevaluated.

INTRODUCTION

An alarm is a warning of an approaching situation which requires a response.

Alarm systems give audible, visual or other forms of signal to indicate a potential need or hazardous condition. Alarm sounds are common components in technological environments, such as in industry, power plants, transportation and in many other settings. This study considers the influences of alarms in a clinical health environment.

The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO), a federally approved nonprofit organization whose mission it is to continuously improve the safety and quality of healthcare delivered to the public, has provided an estimate of the number of alarms per day in US hospitals. According to JCAHO, each day hundreds of auditory alarm signals sound for every patient, thousands of alarm signals ring in every unit and tens of thousands of alarm signals blare throughout every hospital (as cited by Mitka, 2013). In addition, studies have indicated that the number of false alarms on any medical floor varies from 80-99% (Lawless, 1994; Tsien & Fackler, 1997; Clinical Alarms Task Force, 2007; Cvach, 2012). This excessive number of false positive alarms results in caregiver sensory overload and desensitization that make them disinclined to respond to real threats due to what has been called alarm fatigue. Therefore, the Emergency Care Research Institute (ECRI), a nonprofit organization for applied scientific research in healthcare improvement, considered alarm hazard as the number one health technology hazard for the years 2012 through 2014 (ECRI, 2011; 2012; 2013). Alarm hazard is primarily a result of alarm fatigue. Repeated

false alarms and constantly noisy monitors lead to nurses silencing or turning off the alarms due to alarm fatigue, which has been shown to be the main contributor in hundreds of patient deaths (Kowalczyk, 2011).

An unanticipated event in a healthcare setting which results in death or injury, but is not related to the natural course of the illness, is called a *Sentinel Event* by JCAHO. JCAHO, which was renamed The Joint Commission (TJC) in 2007, has a database of Sentinel Events that includes reports of 98 alarm-related events between January 2009 and June 2012. Of the 98 reported events, 80 resulted in death, thirteen in permanent loss of function, and five in unexpected additional care or extended stay (The Joint Commission, 2013). Ninety-four of the reported events occurred in hospitals, with the majority of those events occurring in telemetry, intensive care, general medicine, and emergency department areas. In an interview, the commission stated that those voluntary reports were a gross undercount. In reality, the estimates were close to 1,000 alarm incidents in which patients died or were injured, or faced risk of extended stays. In more than 60 percent of the cases, alarms were either inappropriately turned off or were not audible in all areas (Crites, 2013). Interestingly, the Food and Drug Administration's (FDA) Manufacturer and User Facility Device Experience (MAUDE) database reveals 566 death reports associated with alarms from January 2005 to June 2010 (The Joint Commission, 2013). Those reports do not provide complete information about the scenario and don't confirm that alarm was the main contributor to those deaths. The research of the ECRI found more than 2,200 incidents related to medical devices between

2000 and 2006, 12% of which were alarm-related problems (Clinical Alarms Task Force, 2007). Alarm-related problems may include technical failures as well as alarm fatigue. Nevertheless, JCAHO has found these numbers significant enough to address.

In April 2002, JCAHO appointed a panel of physicians, nurses, pharmacists, and other patient safety experts to advise them on developing the first set of national patient safety goals (NPSG). Improving the effectiveness of clinical alarm systems was one of six goals (The Joint Commission, 2014). In 2004, this goal was no longer an optional goal for hospital organizations and was incorporated into JCAHO standards. Since 2004, it has been mandatory by JCAHO standard for all hospitals in the US to assess alarm fatigue in their monitoring process and to develop a systematic, coordinated approach to clinical alarm system management (Clinical Alarms Task Force, 2007). In addition, JCAHO requires root cause analyses for all sentinel events (Williams, 2001). This would now include those that result from alarm fatigue. In October 2011, a two-day summit was convoked by the ECRI, the Association for the Advancement of Medical Instrumentation (AAMI), and the American College of Clinical Engineering (ACCE), along with the FDA and TJC. At that summit, multidisciplinary specialists gathered in a collaborative effort with a mission of eliminating the issue of patient safety risk related to the alarms from healthcare by 2017 (Association for the Advancement of Medical Instrumentation, 2011). The summit discussed alarm-related topics and agreed to a set of short-term and long-term objectives to reach this goal.

Nevertheless, alarm fatigue remains a high concern in healthcare. However, in order to comply with this mission and requirement, a working definition of alarm fatigue is necessary. Even though researchers and professionals have used the term alarm fatigue, this phrase has only been defined in qualitative ways without any quantitative representations that can be useful in identifying factors that cause such fatigue.

This research has the objective of creating a working definition of alarm fatigue which identifies all the influencing factors and, consequently, the influence of alarm fatigue on performance deterioration. Alarm fatigue has been defined in qualitative ways by various studies using several influencing factors that cause alarm fatigue, such as the number of false alarms (Baillargeon, 2013), sensory overload and desensitization (Pishori, 2012), work-load due to long shift and repetitive task (Weinger & Smith, 1993), noise level (Wallis, 2010), difficulty with alarm-creating equipment and their settings (Welch, 2011), emotional, psychological and personality factors (Weinger & Smith, 1993).

The most common and reasonable factor which can be used to measure the influence of alarm fatigue is the response to the alarm: whether it is ignored, delayed, silenced or turned off (Graham & Cvach, 2010; Ashton, 2011). Some studies have created flow charts, cause-and-effect diagrams to support the root-cause analysis for alarm fatigue (Clinical Alarms Task Force, 2007; Baillargeon, 2013). To the very best of our knowledge, however, there has not been a study done which identifies the significant factors behind alarm fatigue and the influence of alarm fatigue on staff performance

using statistical analysis. Furthermore, according to most of the qualitative definitions of alarm fatigue, it has been found that staff get overwhelmed and desensitized due to a high number of false alarms and continuous beeping monitors, resulting in alarm fatigue; however, no study measured this overwhelming condition or desensitization. This study undertook the research project for defining alarm fatigue and identifying the influence of alarm fatigue on performance deterioration. In addition, the use of personality factors as an influencing variable and mental workload and affect as measures of alarm fatigue were unique considerations for this kind of observational study.

In this research, statistical analyses were conducted considering mental workload and affect as dependent variables and total number of alarms, staff to patient ratio (workload), time elapsed since start of the shift, alarm type, alarm criticality, noise level, task priority and staff personality as independent variables. A survey, using a questionnaire proposed by the American College of Clinical Engineering Healthcare Technology Foundation (ACCE HTF), was administered before the observations to assess the attitude of nurses toward the current alarm system at the hospital where the observations took place. The mental workload was measured using two occupational mental workload measuring tools: Subjective Workload Assessment Technique (SWAT) (Reid, Potter, & Bressler, 1989) and NASA Task Load Index (TLX) (Barker Steege & Nussbaum, 2012). Affect considered for this study were distrust, apathy, and boredom which were measured using validated survey tools. For apathy and boredom, the short versions of the Apathy Evaluation Scale (APS) (Marin R. S., 1991) and the Boredom

Proneness Scale (BPS) (Farmer & Sundberg, 1986) were used respectively. For distrust, some questions were developed on the basis of an inter-cultural scale to measure the trust in automation (Chien, Semnani-Azad, Lewis, & Sycara, 2014). Those questions were validated by the staffs in the facility where the research took place. The purpose of the research is to define alarm fatigue in terms of mental workload and affect considering all the influencing factors that cause fatigue. This definition will help hospitals determine the risk of their staff experiencing alarm fatigue by nurses in any acute-care area. At the same time, this study intends to verify whether this alarm fatigue is the cause that deteriorates staff performance and eventually causes adverse clinical incidents in terms of patient deaths, damage to patient conditions or extended stays among others.

Research Questions

In this study, five questions were examined: 1) What are all the possible factors that influence alarm fatigue in the staff?; 2) Is it possible to define alarm fatigue in terms of the mental workload of the staff based on those factors?; 3) Similar to the previous question, is it possible to define alarm fatigue in terms of affect on the staff based on those factors?; 4) If both definitions are possible, are they analogous?; 5) How does alarm fatigue influence staff performance?

LITERATURE REVIEW

Over the last 30 years, alarm systems in clinical environment have become more widespread and have been improved through new technologies, bringing the number of different alarms from 6 in 1983 to more than 40 in 2011 (Kerr & Hayes, 1983; Borowski, et al., 2011). In spite of this progress, however, these alarms create several problems in the healthcare process and sometimes even result in patient deaths. In 1974, the Emergency Care Research Institute (ECRI) published its first sentinel event report on an ignored alarm signal of a hypothermia machine that resulted in serious patient burns. The report identified a minor drawback in the hypothermia machine in that the alarm light flashed continuously until the patient's temperature reached the desired value. Because of the constant flashing of the alarm light during a normal phase of the warming process, the alarm was ignored during a critical event (ECRI, 1974; Clinical Alarms Task Force, 2007). By 1982, researchers recognized the increasing number of monitoring signals, with “no end in sight” (Stafford, 1982). A year later, Kerr & Hayes (1983) identified that each patient could have six or more alarms, a situation that could increase the risk of caregiver confusion as to which alarm is going off and would actually be contrary to the patient's best interest.

An extensive review of literature was conducted utilizing Inspec, PubMed, Pubget, CINAHL, Science Direct and Google Scholar Search with the purpose of searching for nursing and healthcare articles, as well as human-factor articles for personality measures, mental workload measuring tools and affect scales such as

boredom, apathy, and trust/distrust measuring scales. Keywords used for the search included: alarm fatigue, clinical alarms, cardiac monitors, telemetry, mental workload, personality tests, boredom evaluation, apathy evaluation, and trust and distrust evaluation. The search was limited to English and there was no date limit set. The literature reviewed include White Papers published by ACCE, articles from the Sentinel Event Alert by TJC, a summit report written by the AAMI, FDA, TJC, ACCE and ECRI, various surveys, and narrative reviews of current literature with suggestions for alarm evaluation and management. In addition, several studies using personality tests in healthcare and research that included measurement of the mental workload of nurses were reviewed. Five observational studies on clinical alarms are included. A thorough search revealed no meta-analyses, randomized control trials, or experimental design studies related to alarm fatigue.

Alarm Fatigue

Alarm Fatigue is the desensitization of staff in a clinical environment with a large number of alarms, high noise level, long shift hours, and high workload, among other factors (Weinger & Smith, 1993; Ashton, 2011). There are several types of alarms, depending on their validity and the staff's response toward them. Definitions relevant to this proposed study are given on Table 1. Response time, considered by the study, is also defined here.

Table 1: Definition of technical terms related to clinical alarms

	Alarm	Definition	References
Non-actionable Alarms	False	A false alarm is an alarm which occurs in the absence of an intended, valid patient or alarm system trigger. Examples include motion artifact, improper detection of a P or T wave, and low perfusion during pulse oximetry.	(Clark & David, 2006; Welch, 2011)
	Nuisance	Nuisance alarms are annoying alarms that may interfere with patient care, and typically do not result from adverse or potentially adverse patient conditions and which is repeated five or more times during a one-hour observation period.	(Clark & David, 2006; Welch, 2011)
	Technical	A technical alarm is an alarm to notify clinicians that an operational aspect of the system requires attention. An example is a leads-off alarm which indicates leads/sensors have come off the patient's skin.	(ECRI, 2007; Ashton, 2011; Baillargeon, 2013)
Actionable Alarms	Valid	A valid alarm is an alarm that represents true and accurate physiological data and requires an instant response from a caregiver. An example is a true critical V-tachycardia alarm which indicates a rapid heartbeat starting in the bottom chamber of the heart that can lead to sudden death.	(Welch, 2011)
	Response Time	The time between the moments when an alarm goes off and when the nurses acknowledge that they have seen the alarm and take required action.	

At the turn of the century, before The Joint Commission recognized alarm systems as a risk to patient safety and identified the term alarm fatigue, research focused on assessing the effectiveness of alarm monitoring systems, defining the problems related to these systems and identifying the factors that influenced those problems. In this regard, Weinger & Smith (1993) published an article showing the relationships among vigilance, alarms, and integrated monitoring systems. The authors stated that caregivers are prone to error in vigilance and monitoring. They revealed many instigating factors behind those errors, such as number of false alarms, high noise level, equipment and systems, workload and task characteristics, and caregivers' emotional factors, psychological and

personality factors, stress levels and training and experience. They also believed that alarms would be effective only if properly designed and implemented.

The influences of increasing alarm reliability on alarm response frequencies were assessed in terms of speed and accuracy. The results indicated that most participants (about 90%) matched their alarm response to the expected probability of true alarms. A reduction in alarm reliability resulted in a reduced response. Two years later, in 1995, James P. Bliss along with his colleagues, conducted an observational study on alarm responses to assess the monitoring system. Alarm response frequencies were observed and analyzed using t-tests and repeated-measures MANOVA. The result indicated that about 90% of the staff do not respond to the alarms regardless of their reliability and urgency.

In 1999, Chambrin et al. conducted an innovative observational study to evaluate the existing alarm monitoring system in the adult ICU at the University Hospital of Lille, France. The researchers tracked 1,971 hours in a critical care unit and found 3188 alarms, of which only 5.7% required response. The authors did not mention the term alarm fatigue, but considered false alarms as a significant issue. They confirmed several factors which should be considered to improve the effectiveness of alarms, such as alarm type, technical factors, and caregivers' training, education and shift periods (day/night).

Since TJC (JCAHO at that time) coined the term alarm fatigue and declared improving the effectiveness of clinical alarms as a standard for hospitals in 2004, much more research has been conducted to assess existing alarm monitoring systems, define

alarm fatigue, and identify the factors which cause that fatigue. Evidence of this is presented on Table 2, page 8. In 2004, a Healthcare Technology Foundation (HTF) survey, proposed by ACCE, was conducted with 1372 healthcare professionals to assess the attitude of healthcare professionals toward the existing alarm system in their work area and to help management understand their problems in order to take corrective actions. In 2006, Clark & David published a seminal White Paper on this survey which reviewed the effective use of clinical alarms and analyzed adverse-event databases. The analysis showed that, according to the majority of the respondents, alarms occur too frequently, disrupt patient care, reduce trust on alarms, and lead the caregivers to inappropriately disable the alarms. In 2011, a follow-up HTF survey voiced the perceptions of 4,278 clinical staff. The responses of that survey were similar to the previous one conducted in 2004 (Kaplan & Saccuzzo, 2013). Later in 2011, the Clinical Alarm Summit suggested improvements and priorities for this highly evident patient-safety concern based on the result of this second survey. In addition, they emphasized that nearly 20% of the responses reported adverse events occurring in their hospitals between 2009 and 2010. Nuisance and false alarms remained the most significant problem, with little progress reported in alarm hazard reduction since the first HTF survey in 2004 (Association for the Advancement of Medical Instrumentation, 2011; Clark & David, 2011).

Until this point, very few observational studies, as opposed to surveys and reviews, have been conducted on the hazards of clinical alarms and how they can be

managed. In 2003, a nine-bed Coronary Respiratory Care Unit (CRCU) in a tertiary care Canadian university hospital took a new approach by doing an observational study to examine the influences of remote telemetry on nurses and patients (Billinghurst, Morgan, & Arthur, 2003). Critical care nurses in the unit were responsible for monitoring the telemetry of patients in other units in the hospital. A total of 420 hours of data were collected for the CRCU and 52 hours of data for the remote telemetry. The researchers (2003) found that 80.2% of the total alarms were non-actionable and there was a wide range of 60% to 100% in nurse detection and response to valid alarms. The authors concluded that remote cardiac telemetry placed unnecessary demand on CRCU nurses' workload and negatively influenced the care provided to CRCU patients. These conditions are commonly seen in American hospitals where acute care nurses are responsible for the monitoring of telemetry on other medical floors.

In 2012, Tasneem Pishori at the University of Connecticut wrote a Master's thesis investigating the routine processes of alarm monitoring in a telemetry unit. The researcher simulated the existing system in order to quantify the sources of noise contributing to nursing fatigue. Real-time data were gathered through an observational study to understand the state of existing response strategies through an observational study. Data were also obtained on the number of patients, the number of staff, numbers and types of alarm and noise sources, staff responsible for alarm response and their back-ups, and additional responsibilities of staff which may prevent them from responding to alarms quickly. By altering several factors and parameters, the simulation model showed

the positive influence of reducing certain activities from the workflow on increasing the idle time of the nurses at the central monitoring area. The purpose was to make the nurses free enough to concentrate on their primary task of reviewing alarms, documenting them and communicating any issues to the appropriate caregiver.

According to a non-observational evocative study, done by Sendelbach & Funk (2013), in order to reduce the alarm-related incidents due to alarm fatigue, it is necessary to address the common causes behind the fatigue. Their study emphasized education, customization and /or minimization of alarms, adequate staffing, and preparation of skin and proper placement of electrodes to reduce the number of false and non-actionable alarms. The authors also recommended designing an experiment using a randomized control trial and statistical analysis focusing on potential outcomes, in order to find the influence of different factors. However, they did not conduct any observational study nor do any experimental design.

Also in 2013, a risk assessment study was performed by Baillargeon at an acute-care teaching hospital in Providence, RI. The researcher quantified all alarms from two medical surgical units over six-week periods for one-hour blocks during day and night shifts in order to determine the percentage of false, nuisance, technical and valid alarms. The test evaluated alarm response and included various questions related to risk assessment, such as the proximity of alarms, frequency, and staff member preparedness. Alarm frequency was calculated and average response time to critical and leads-off alarms were determined. (Leads-off alarms are alarms which indicate that leads/sensors

have come off the patient's skin, cutting off the readings.) The results showed that the percentages of false, nuisance and technical alarms added up to 52%, which the researcher considered as a contributor to alarm fatigue. In fact, response times were found at times to be over 10 minutes. Baillargeon's study concluded that nurses were at risk for experiencing alarm fatigue based on high alarm frequency, increased number of false and nuisance alarms, and a delayed response to leads-off conditions.

Despite all the research being performed on this topic, none of the studies discussed above attempted to find a quantitative or statistical definition of alarm fatigue nor measure alarm fatigue using nurses' mental workload and/or affect even though alarm fatigue might be highly related with mental and emotional processes. In the light of this gap in the literature, this current observational study was designed to define alarm fatigue identifying all possible influencing factors, which can be specified as working conditions and staff individuality. The factors that were considered influential to alarm fatigue in previous studies are summarized in Table 2. The summary demonstrates the reason for collecting each factor and the procedure to collect each of them for this alarm definition study, and references to the articles in which that factor has previously been collected as a part of an observational study or considered as an influencing factor in the research. In addition, Table 2 shows personality type, mental workload and affect that were measured using standardized survey tools during this research. These three variables (personality type, mental workload, and affect) will be further discussed in the following sections.

Table 2: Summary of literature review

Name of the variables to be collected	Why to collect	How to collect	Who has considered but not collected this variable as an influencing factor	Who has collected this variable for their own research
Total Number of Alarms	To determine the number of actionable vs. non-actionable alarms	Direct observation of all alarms from all shifts (day and night)	(Weinger & Smith, 1997; Graham & Cvach, 2010; Sendelbach, Funk & Marjorie, 2013)	(Chambrin, et al., 1999; Billinghamurst, Morgan, & Arthur, 2003; Pishori, 2012; Baillargeon, 2013)
Staff to Patient Ratio	To determine nurse workload	Direct observation	(Sendelbach, Funk & Marjorie, 2013)	(Billinghamurst, Morgan, & Arthur, 2003; Pishori, 2012)
Alarm Type	To determine whether any particular type of alarm is the main contributor to most of the non-actionable alarms	Direct observation	—	(Chambrin, et al., 1999; Pishori, 2012; Baillargeon, 2013)
Time Elapsed	To check whether the time elapsed influences alarm response performance	Direct observation using stop watch	—	(Pishori, 2012)
Noise Level	To check whether the noise level is higher than the recommended decibel level	Noise meter	(Weinger & Smith, , 1993)	(Pishori, 2012)
Nurse ID	To determine the between-subject variability	Daily staff schedule at ICU		(Pishori, 2012; Baillargeon, 2013)
Shift (Day/Night)	To determine whether shift time influences alarm response performance	Direct observation	(Sendelbach, Funk & Marjorie, 2013)	(Chambrin, et al., 1999)
Personality Type	To determine whether personality traits influence the rate at which staff becomes overwhelmed	Using Big Five Personality test	(Weinger & Smith, 1993)	—

Table 2: Summary of literature review (continued)

Name of the variables to be collected	Why to collect	How to collect	Who has considered but not collected this variable as an influencing factor	Who has collected this variable for their own research
Affect	To measure staff apathy and boredom occurring from a long shift of alarm monitoring	Surveys using Boredom Proneness Scale, and Apathy Evaluation Scale	(Weinger & Smith, 1993; Cvach, 2012)	—
Mental Workload	To measure staff perception of alarm fatigue during their shift	SWAT used during work and NASA TLX used after shifts	(Weinger & Smith, 1993)	—
Response (Yes/No)	To determine the impact of alarm fatigue on staff in order to investigate the association between staff performance and alarm fatigue	Direct observation	—	(Baillargeon, 2013)
Response time		Direct observation using stop watch	—	(Chambrin, et al., 1999; Billingham, Morgan, & Arthur, 2003; Pishori, 2012; Baillargeon, 2013)
Number of Ignored Alarms		Direct Observation	(Weigner & Smith, 1997; Ashton, 2011)	(Baillargeon, 2013; Pishori, 2012)

This summary of literature review considered preliminary studies and the studies which have been conducted to investigate the factors that can cause alarm fatigue. In addition, some of these studies collected data on staff performances and blamed alarm fatigue for any performance deterioration. Based on the reviews of these studies, this research was undertaken to create a statistical definition for alarm fatigue, identify the methods to measure it and determine its influence on performance deterioration.

Personality Type

Review of the current literature indicates that while personality traits may play a significant role in alarm fatigue, very little research has been done to explore this relationship. Nevertheless, tools do exist for determining personality traits which could be used for such research. A personality test is a questionnaire or other standardized instrument designed to reveal aspects of an individual's character or psychological makeup. Depending on their personality traits, people create their own work ethic and environment, and influence their job satisfaction through mental, emotive and behavioral processes. Individuals mentally interpret characteristics of their jobs differently according to their personality (Judge, Heller, & Mount, 2002).

Personality has been shown to affect various areas of life, such as coping with stress, dealing with crises, and job performance (Barry & McCarthy, 2001; Meeusen, Brown-Mahoney, van Dam, van Zundert, & Knape, 2010; Franklin, Johnson, White, Franklin, & Smith-Olinde, 2013). It is reasonable, therefore, to consider the influences of personality as they apply to concerns with alarm fatigue, in areas such as coping with a constantly noisy environment, devoting long-hours to alarm monitoring, making careful decisions under stress, and adapting to new and changing technologies. Important research has been done showing the relationship of personality traits to information processing and decision-making. Results have shown that test performances are both positively and negatively correlated with extraversion (Ackerman & Heggestad, Weigner & Smith, 1997), negatively correlated with neuroticism (Zeidner & Matthews, 2000),

positively correlated with conscientiousness (Moutafi, Furnham, & Crump, 2003), and positively correlated with openness to experience (Chamorro-Premuzic, Moutafi, & Furnham, 2005). In addition, many studies have reported that personality type can play a role in preferred auditory levels and have found that the Extraversion and neuroticism traits have influencing influences on what is acceptable in noise level, and even on the detection of sound (Geen, 1984; Franklin, Johnson, White, Franklin, & Smith-Olinde, 2013; Hutchinson, Duffy, & Kelly, 2005). Therefore, in defining alarm fatigue, personality may be a large contributing factor in an environment that requires alarm monitoring and decision making while simultaneously sifting out a large number of non-actionable alarms in a constantly beeping noisy background.

Nurses are responsible for the daily monitoring and management of the quality of care delivered to patients. The role of critical care nurses includes immediate detection and intervention when patients' clinical conditions change, with the help of direct observations as well as alarms. Expert nursing practice requires not only psychomotor and affective skills, but also complex thinking and judgmental processes, such as making inferences and synthesizing information to choose a course of action. Nurses who work in the intensive care unit (ICU) can be exposed to considerable work-related stress (Coomber, et al., 2002; Levy, 2004), are frequently distressed (Mealer, Shelton, Rothbaum, & Moss, 2007), as well as fatigued (Surani, Hesselbacher, Guntupalli, Surani, & Subramanian, 2014). In addition, a high degree of emotional exhaustion among nurses has been shown to predict lower self-rated performances and higher intention to quit

work (Freyaldenhoven, Smiley, Muenchen, & Konrad, 2006). Therefore, for the proposed research study on defining alarm fatigue, personality traits will be an important factor to consider and measure with the proper personality test.

The first personality tests were developed in the 1920s and were intended to ease the process of personnel selection, particularly in the armed forces. Since these early efforts, a wide variety of personality tests have been developed, notably the Myers Briggs Type Indicator (MBTI) and the Big Five Personality Test. Recently studies have applied these personality tests in the healthcare arena to evaluate the relationship of nurses' personalities to their job and work environment (Meeusen, Brown-Mahoney, van Dam, van Zundert, & Knape, 2010; Franklin, Johnson, White, Franklin, & Smith-Olinde, 2013).

The MBTI was originally developed in 1958 based on Jung's theory of psychological types to describe people's perceptions and judgments in different ways. The test separates personalities into four dichotomies: Extraversion (E) vs. Introversion (I), Sensing (S) vs. Intuition (I), Thinking (T) vs. Feeling (F), and Judging (J) vs. Perception (P). In 1986, Thompson and Borrello conducted a statistical analysis for MBTI, using a limited population of 359 participants and 95 items, and strongly supported the construct validity of the MBTI test. Three years later, McCrae & Costa argued that the MBTI questionnaire is not a valid instrument for measuring these four personality traits. In 1996, Jackson et al. added that the number of participants should be

five to ten times the number of items used in this type of questionnaire to validate it. As can be seen, there is not a strong consensus on the effectiveness of this tool.

Table 3:Dimensions for MBTI

Extraversion (E)	(I) Introversion	→
Sensing (S)	(N) Intuition	→
Thinking (T)	(F) Feeling	→
Judging (J)	(P) Perception	→

Table 4:Dimensions for Big Five Factors

Extraversion	Introversion
Openness	Closedness
Agreeableness	Hostility
Conscientiousness	Lack of Conscientiousness
Emotional Stability	Neuroticism

On the other hand, the Big Five personality model has received much more favorable reviews. It was originally derived in the 1970's by two independent research teams- Paul Costa and Robert McCrae (at the National Institutes of Health) and Warren Norman (at the University of Michigan) with Lewis Goldberg (at the University of Oregon). This test divides personalities along five dichotomies: Extraversion vs. Introversion, Openness vs. Closedness to Experience, Agreeableness vs. Hostility, Conscientiousness vs. Lack of Conscientiousness, and Emotional Stability vs. Neuroticism. This model has been widely accepted as a valid and reasonably generalizable taxonomy for personality structure (Digman, 1990), and has been used by numerous researchers as a framework to explore the criterion-related validity of personality in relation to job performance (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Tett & Murphy, 2002; Rothmann & Coetzer , 2003). An experiment compared both the tools to test personality and found similarities between the four dimensions of the MBTI and four of the five dimensions of the Big Five Factors

(Tobacyk, Livingstone, & Robbins, 2008). Table 3 and Table 4 show the dimensions for the two personality tests and the arrows show the similarity between the dimensions.

Elements of the Big Five can be seen in nearly all personality measurement systems. Because of the consistency he had seen in the results using the Big Five Factors, Goldberg suggested that any model for structuring individual differences would have to encompass similar dimensions (1981). Pertinent to this study, Lakin et al. found in 2007 that the Big Five Factor personality traits of Neuroticism and Extraversion were the only personality variables associated with fatigue. In the light of the overwhelmingly positive support for the Big Five model, this study used the Big Five Factors to test the personality of the nurses in order to determine whether personality affects how quickly nurses get overwhelmed.

Mental Workload

As with personality, although tools exist which can be used to study the role of mental workload on clinical alarm fatigue, very little research has been done in this arena to date. There are several different methods currently in use to measure the mental workload of caregivers in healthcare. One of them is the National Aeronautics and Space Administration-Task Load Index (NASA-TLX). This is a multidimensional scale for overall mental workload as a function of mental demand, physical demand, temporal demand, performance, effort and frustration dimensions, with each of these dimensions on a continuum (Hart & Staveland, 1988). These subscales can be measured on their own

(“Raw TLX”) or combined in a pairwise comparison that measures the relative importance of each subscale to the participants (Hart & Staveland, 1988). The evaluation consists of scales divided into twenty equal intervals, manually marked between labels at each end from Low to High or Good to Poor. This tool has been found to be reliable (Battiste & Bortolussi, 1988), sensitive (Hill, et al., 1992), and valid (Hart & Staveland, 1988; Rubio, Díaz, Martín, & Puente, 2004) and has been used successfully in several research arenas, including in healthcare with nurses, as a retrospective mental workload measuring tool (Young, Zavelina, & Hooper, 2008; Hoonakker, et al., 2011).

Another method that has been developed for mental workload measurement is the Subjective Workload Assessment Technique (SWAT). Initially this tool is described by Nygren (1982). Afterwards, in 1988, Reid and Nygren described it more completely. SWAT was then found as a measure of mental workload that is valid (Haworth, Bivens, & Shively, 1987), sensitive (Reid, Eggemeier, & Shingledecker, 1982), reliable (Griscomb, 1985) and relatively unobtrusive (Eggemeier, 1988). This is also a multidimensional scale, but its dimensions of time load (T), mental effort load (E), and psychological stress load (S) are at three discrete levels: low, medium and high. This tool requires scaling the subjective ratings provided by the participants through either a comparison of the relative importance of all levels over all three dimensions (27 distinct comparisons of each level and dimension with every other combination) or a pair-wise comparison (Luximon & Goonetilleke, 1998). The later kind of SWAT using pair-wise comparison is called continuous SWAT or C-SWAT and is time saving and gives similar

results to the SWAT with card sorting. In addition, SWAT can be used in two ways: during the tasks and/or after the completion of the task (Reid, 1989).

A third mental workload measurement tool is the Swedish Occupational Fatigue Inventory (SOFI). This has been recently developed and validated to contribute to the understanding of fatigue by investigating the qualities of fatigue in people from different occupations (Ahsberg, Gamberale, & Kjellberg, Weigner & Smith, 1997; Ahsberg & Gamberale, 1998; Ahsberg, 2000; Leunga, Chana, & Heb, 2004). The SOFI consists of 20 items belonging to the following dimensions: Lack of Energy, Physical Exertion, Physical Discomfort, Lack of Motivation, and Sleepiness (Ahsberg, 2000). There is, however, very little research supporting or contradicting the effectiveness of SOFI in this kind of research.

In 2004, a laboratory study conducted by Rubio et al. compared NASA TLX with SWAT and another mental workload measuring tool. The result showed that NASA TLX was highly correlated to SWAT (with a correlation of 0.97). However, even though SWAT has been tested with a psychological model of human-perceived information processing demand, studies such as Hart and Staveland (1988) and Hill et al. (1992) have shown that NASA-TLX is superior to SWAT in terms of sensitivity, especially for low mental workloads. Since NASA TLX is not designed to be given during a task, a combination which uses SWAT during the task and NASA TLX upon completion, seemed ideal for this study.

Affect

At all levels of ability, apathy, distrust, and boredom are part of a human's life. Sometimes these affect can bring a little inconvenience and sometimes they can cause serious injury, damage, and even loss of life. In the case of clinical alarm monitoring, every single alarm needs attention. However, a high number of false alarms and a constantly beeping environment reduce trust in alarms and cause apathy and boredom in staff. These three types of affect were considered the measures of alarm fatigue in this study along with the mental workload.

Boredom

Boredom is an insufficient workload, unpleasant, transient affective state caused by insufficient workload in which the individual feels a lack of interest and difficulty concentrating on their activity (Fisher, 1993). This is an aversive subjective state of low arousal (Mikulas & Vodanovich, 1993). During the alarm monitoring process, due to large numbers of non-actionable alarms, staff become indifferent and display a lack of interest in responding to alarms. This state is the result of the monotonous occurrence of non-actionable alarms which causes desensitization. This can be defined as alarm fatigue in terms of boredom. The Boredom Proneness Scale (BPS) was developed to measure boredom in a way that emphasizes individuals' connectedness with the environment and also the willingness to access adaptive resources and recognize urgencies (Farmer & Sundberg, 1986). This study intends to evaluate the boredom of caregivers in a noisy

environment with a high number of false alarms. Therefore, a short version of the Boredom Proneness Scale (BPS) was used in this study to measure the boredom of ICU staff.

The original BPS was developed by Farmer & Sundberg in 1986 with 28 items. Several studies provided evidence of the validity of BPS through comparisons with other measures of boredom and similar constructs (Ahmed, 1990; Farmer and Sunberg, 1986; Rupp & Vodanovich, Weigner & Smith, 1997; Watt & Blanchard, 1994). In 2010, a short version of this tool, having 12 items, was proposed and validated by Vodanovich, Wallace, & Kass. Ten items were taken directly from that short version of the original BPS for this study; two questions were rejected for their irrelevance to this study.

Apathy

A means about and *pathos* means emotions. So apathy means complete carelessness. This is defined as a lack of feeling, interest, or concern. It is a state of indifference, or the suppression of emotions such as concern, excitement, motivation, and/or passion. In an alarm monitoring system, caregivers get overwhelmed by a high number of repetitive non-actionable alarms and desensitized by high noise level. As a consequence, they become demotivated and immune to any alarm. This is the state of apathy which can be a measure of alarm fatigue.

This study decided to apply the clinician version of the Apathy Evaluation Scale (AES) to measure apathy in the ICU staff. The original Apathy Evaluation Scale was developed and validated by Marin in 1991. It has 18 items requiring subjective responses

which are to measure apathy in an individual. Much research has been conducted in healthcare in order to provide evidence for the reliability and validity of AES (Marin, Biedrzycki, & Firinciogullari, 1991; Hsieh, Chu, Cheng, Shen, & Lin, 2012). In 2007, Lueken with his colleagues proposed and validated a short version of this the Apathy Evaluation Scale. The short version has ten items which were used for this research to measure staff apathy by means of a retrospective survey after each observational shift.

Distrust

Distrust is the feeling of not being able to rely on something and/or someone. In the case of alarm monitoring, a large number of false or nuisance alarms leads the staff to stop relying on all alarms. This situation results in an undesirable response, such as a delayed or no response, silencing, or turning off alarms, changing threshold parameters outside the safe limits, with the possibility of very dangerous consequences. Therefore, distrust can be a very important variable to define and measure alarm fatigue.

Literature review did not give any established and valid tool to measure distrust for the hospital staff during their alarm monitoring task. This research created a survey tool in order to measure staff distrust toward alarms. This survey tool had three items. Two of the three questions were created on the basis of an Inter-Cultural Scale to measure trust in automation (Chien, Semnani-Azad, Lewis, & Sycara, 2014). A third question was added on expertise opinion. The questions were validated by two registered nurses and two Unit Clerks along with the manager in the ICU at BDH.

Performance Measures

Several studies tried to assess alarm fatigue in terms of staffs' response (yes/no) (Graham & Cvach, 2010; Pishori, 2012; Baillargeon, 2013), response time to alarms (Chambrin, et al., 1999; Billinghamurst, Morgan, & Arthur, 2003; Pishori, 2012; Baillargeon, 2013), and number of ignored or missed alarms (Graham & Cvach, 2010; Baillargeon, 2013; Pishori, 2012) in order to determine the deterioration in staff performance for an adverse working conditions. This performance deterioration was measured as alarm fatigue by those research works.

In clinical environments, it is mandatory to respond to each alarm according to their criticality and necessity. The staff always has other important duties to carry out along with alarm monitoring, such as documentation, attending phone calls, and medication. Therefore, response time or response (yes/no) varies extremely over the shifts. Furthermore, according to the findings of this study, these performances cannot be proper measures for alarm fatigue. This study considers these factors (response, response time and number of ignored alarms) as performance measures and intends to identify the correlation between alarm fatigue and these staff performances. These measures were collected during direct observation, each time an alarm went off.

Literature Review Summary

This literature review contributes to the present study in several steps. It focused, first, on the assessment of the effectiveness of clinical alarms; secondly, on the search for

all possible factors that lead to unit clerks and nurses being fatigued; thirdly, on direct observation methods for collecting data on those factors; fourthly, on the personality tests that have been used to investigate the personality types of people in stressful decision-making occupations; fifthly, on mental workload measuring methods; and then finally, on the surveys available to measure the pertinent affect: boredom, apathy, and distrust.

The research established without a doubt that alarm fatigue due to the current clinical alarm systems poses severe risk to patient safety. As far as we know, this was the first study defining alarm fatigue by investigating all the possible influencing factors as working conditions behind this fatigue through task analysis. The factors and responses which were found from the literature review and considered in this study are summarized in Table 2. For some of them standardized tools, such as Personality type tests, mental workload measurements, and effect scales, were used. For the Personality Type Tests, the literature review showed that the MBTI has some contradictory research conclusions concerning its validity and does not have the dimension ‘Neuroticism’, a very important trait for this kind of study. Therefore, this was not used in this study. On the other hand, the Big Five Factor is a valid and updated method, having all the necessary traits needed to investigate personality types and was therefore appropriate for this study. For the mental workload measurements, research has shown that, as compared with the NASA TLX and SWAT, SOFI does not have enough applications in the research arena to measure the mental workload of workers. Therefore, this was not considered in this study as a tool to measure mental workload. This study decided to do C-SWAT during the

working periods of the nurses to get instant responses on their feelings about their mental workload and the RAW TLX at the end of the staffs' shift to get more sensitive retrospective responses on the overall shift load. For the affect scales, boredom, apathy, and distrust were measured using relevant, reliable and valid subjective survey tools which are very popular in the healthcare research arena. The use of personality factor as an influencing variable and mental workload and affect as measures of the alarm fatigue were unique considerations for this kind of observational study.

METHODOLOGY

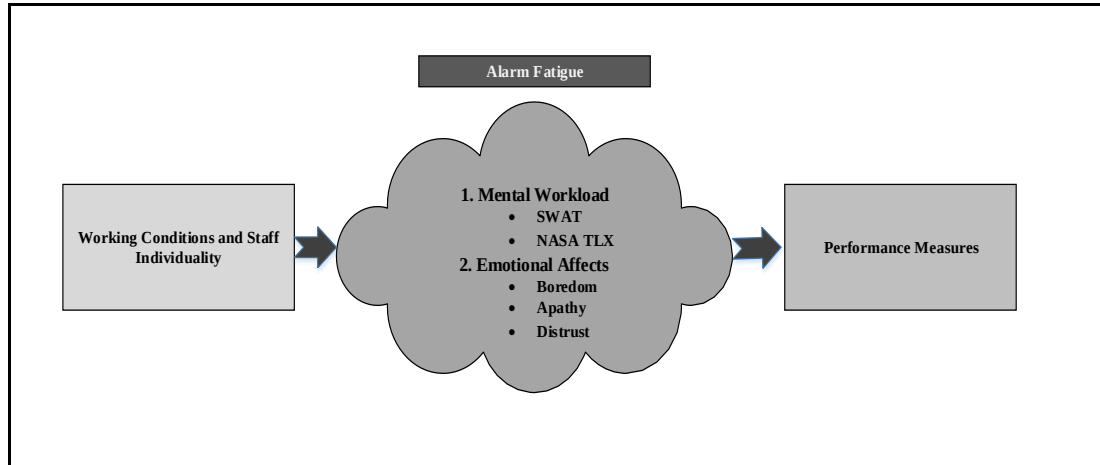


Figure 1: Conceptual model for alarm fatigue and its influence on staff performance

This chapter describes the design of the proposed study. This description begins with the literature review which revealed that alarm fatigue can be defined in terms of mental workload and affect. In order to further break down these two definitions into their influencing variables, several surveys and direct observations were conducted. Direct observations also yielded performance measures as influenced by alarm fatigue. Figure 1 shows a conceptual model for the method of study. Statistical analyses were then used to estimate the significance of the influencing factors in order to propose quantifiable definitions of alarm fatigue and the influence of alarm fatigue on staff performance. The following sections present the approvals obtained from IRB and the Bozeman Deaconess Hospital, followed by brief discussions of the clinical alarm survey and task analyses, direct observations, questionnaires and statistical analyses that were used in this research.

Approval for the Study

Approval from the Montana State University Institutional Review Board (IRB) was obtained prior to the beginning of all surveys and data collection. The Quality Department of Bozeman Deaconess Hospital (BDH) also endorsed the study. In addition, the manager of the Intensive Care Unit (ICU) and head of the nurses at BDH gave approvals for the surveys that were to be conducted. This IRB approval along with informed consent forms are attached in Appendix A.

Participants

The potential participants include all staff who are responsible for the alarm monitoring in the Intensive Care Unit (ICU) in an acute care hospital. Data for six unit clerks and 18 registered nurses were collected for this study. The inclusion criteria consisted of all unit clerks and nurses employed as permanently recruited staff in the ICU during the time of the observations. The staff scheduling policy at ICU for this particular facility included one registered nurse for every two patients (exceptions were available depending on patient criticality) and one unit clerk for one twelve-hour shift. Demographic information for both unit clerks and nurses are summarized in Table 5.

The unit clerk is responsible for monitoring the two telemetry monitors from the central monitoring area: one monitor for the ICU patients and another monitor for the other surgical and medical floor patients. They also answer phone calls. Unit clerks are responsible for responding to most of the alarms. The registered nurses take care of

patients, monitor the ventilator and IV pump alarms in each patient's room, and answer patients' calls as well. All the participants were assigned a participant identity number (ID) for the convenience of data analysis of this research. All personal information of the participants was maintained highly confidential and was not reported with the results. Patients were not observed for this study and no patient information was collected.

Table 5: Demographics summary of participants

Demographic Category		Statistics/Percentage	
		Unit Clerk	Nurse
Gender	Female	83.33%	84.20%
	Male	16.67%	15.80%
Age (years)	Average (SD)	30.33 (8.04)	44.11(16.20)
Education	High School	16.67%	0%
	Bachelor	83.33%	100%
Experience (years)	Average (SD)	7 (2.77)	10.33 (11.48)
Experience at ICU (years)	Average (SD)	5.5 (3.02)	8.08 (7.78)
Specialized training on alarm monitoring		100%	36.36%
Experience of participating in alarm related research		None	None

Location

The center monitoring station has two desks; one of which is assigned for the unit clerk. The unit clerk is responsible for monitoring two telemetry monitors in the central monitoring area. The ventilator and IV pump alarms for each patient are monitored from that patient's room. There are two other rooms adjacent to the monitoring area: staff break room having medicine storage, and the room for pharmacist and physicians. The unit has two twelve-hour shifts. For nurses, one shift starts from seven o'clock in the morning to seven o'clock at night and the second shift starts from seven o'clock at night

to seven o'clock in the morning. Whereas for unit clerks, one shift starts from six o'clock in the morning to six o'clock at night and the second shift starts from six o'clock at night to six o'clock in the morning.

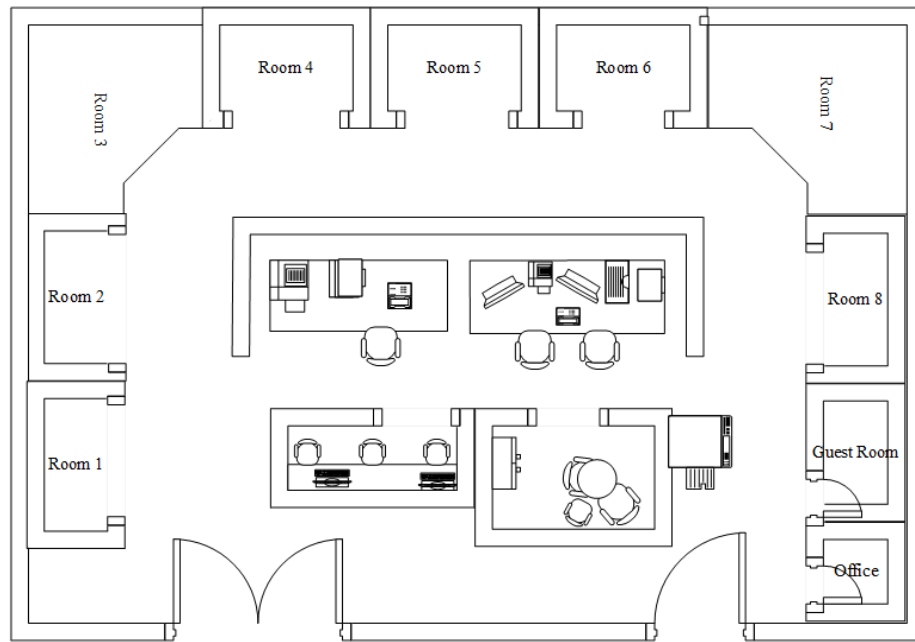


Figure 2: Representation of the ICU at BDH

Procedure

The proposed research conducted several surveys and observations. The study started with a current state assessment survey, the Healthcare Technology Foundation survey. Following this survey, Hierarchical Task Analyses (HTA) were conducted for both unit clerks and nurses for both shifts to determine whether all influencing factors were considered. In order to gather data for most of those factors, some direct observations were taken through work sampling. Two surveys were conducted with all

the participants using two mental workload measuring tools. One of these mental workload measuring surveys (SWAT) was conducted during the observation and the other (NASA TLX) was conducted after completion of the shifts. A personality type test was also conducted to find out the personality type of the staff. Affect was measured for all participants in terms of boredom, apathy and distrust using relevant valid scales. Distrust was measured for each participant using questionnaires during direct observational periods. Although observations were conducted for 1080 hours but 630 hours observations were used during analyses. The missing data for rest of the observations on personality dimensions limit the use of those observed hours in further statistical analyses. Statistical analyses were performed to determine the significance of the factors on the response variables (mental workload and affect) and response variables with the performance measures (response and response time).

Health Technology Foundation Survey

The survey that was used was created by the Healthcare Technology Foundation and conducted twice with physicians across the USA in 2004 and 2011 (Clinical Alarms Task Force, 2007; Clark & David, 2011; Association for the Advancement of Medical Instrumentation, 2011). This study started with that current state's assessment survey to assess the attitudes of staff toward the existing alarm monitoring system in the ICU at BDH. The survey consists of two parts: demographic information on job title, years of experience, working unit, and working facility; and alarm-related information, such as how often nuisance alarms occur, how noisy the environment is, and if it is possible to

differentiate the alarm sounds and identify the sources from among others. To investigate the variability between the responses, other information was asked of the ICU staff using a demographic information form with this survey. The information includes age, education, experience working as a nurse, experience working in the ICU, and previous experience participating in this kind of alarm management study. The demographic form along with the survey is attached in Appendix B. These two forms were distributed to all participants and they were requested to fill these out within four weeks of the distribution.

Hierarchical Task Analysis

Task analysis is the breaking of a task into its component parts. It is a useful way of observing how people interact with various aspects of their working environment. This is the process of learning about typical task performers by observing them in action to understand in detail how they perform their tasks and achieve their intended goals. Task analysis includes a detailed description of both manual and mental activities, task and element durations, task frequency, task allocation, task complexity, environmental conditions, necessary material and equipment, and any other unique factors involved in or required for one or more person to perform a given task (Diaper & Stanton, 2003). Since task analysis is so comprehensive, it was important to include it in this research. This task analysis provided better understanding of the chosen facility and the processes involved. This thorough research on the process and staff helped to determine the factors that cause alarm fatigue and thereby to produce a working definition.

Table 6: The summary of the steps that were conducted to perform HTA in the ICU

Step	Description
1. Decide the purpose of the analysis	1. Define the goals in the ICU 2. Identify all the tasks along with each activity to accomplish the goals 3. Determine the workload for individual staff
2. Get agreement between the stakeholders on the definition of task goals and criterion measures	1. Stakeholders include the ICU manager, staff, supervisors, experts from the Bio-Med and the Quality departments, human factors engineers 2. Concentrate on system values and output 3. Agree on performance indicators and criteria
3. Identify sources of task information and select means of data acquisition	1. Consult all the possible sources available: direct observation, walk-through, protocols and manuals, expert interviews
4. Acquire data and draft decomposition table or diagram	1. Account for each task in terms of input, action, feedback and goal attainment criteria and identify plans 2. Divide tasks into mutually exclusive and exhaustive sub-tasks. 3. Consider not only what should happen but also what might happen. Estimate the probability (p) and cost (c) of failures
5. Recheck validity of decomposition with stakeholders	1. Stakeholders will be invited to confirm the analysis of decomposition by validating the identified goals, and performance criteria

The creation of HTA is generally ascribed to Annett and Duncan (1967). This method produces a hierarchy of three levels of tasks: Goal, which is the system state that a person wishes to achieve; Task, which is a structured set of activities following some sequence which allows a person to achieve the goal; Operation or Action, which is what a person must actually do to carry out each individual activity. In this study tasks were analyzed for nurse and unit clerks separately for both shifts (day and night). The top-level goal of the system is to respond to the alarms at ICU. The task steps necessary to do this are listed on the next level of the hierarchy. Plans indicate the activities or sub-goals that should be carried out in order to achieve the goal. These activities were further broken

down into operations at the lower levels. The order in which these are carried out is determined by the plan. The principal steps that were carried out in doing HTA are given in Table 6.

Observational Data Collection

Observational data were collected using the data collection sheet created by graduate students from College of Nursing and Industrial Engineering Department at MSU (attached in Appendix C). This data collection sheet was validated by taking five pilot observations. There were twelve six-hour direct observational blocks including day and night shifts to collect data on the number of alarms, alarm types, alarm criticality, time elapsed since the start of a shift, staff to patient ratio (workload), and response time to alarms, along with response (yes/no) and task priority, the variables determined through the HTA over two months. These twelve shifts (six for days and six for nights) were randomly selected over two-months and work sampling was conducted to generate random numbers to select six observation slots for each of those shifts. Each observation slot was 15-minutes long.

Work-sampling is a valid and popular method which was used in a hospital setting by Wright in 1954. This application was one of the first examples of using the work sampling method in the case of research in healthcare. Afterwards, this method was applied to research with nursing personnel (Abdellah & Levine, 1954), health professionals (Reid R. A., 1975), emergency departments (Liptak, Super, Baker, & Roghmann, 1985), and other healthcare arenas. Dickson (1978) looked at fixed-time

intervals for work-sampling in contrast to random observations and found no significant differences.

Two researchers were collecting data from two monitors during direct observations: one observer collected data for ICU patients from the telemetry monitor, as well as from phone calls and noise level, and the other observer collected data for the patients from other floors, as well as from the alarms in the ICU rooms including the ventilators, IV pumps, call-lights along, and med-door-open alarms. The study collected twelve replications and eight observations are included in the data analysis. During the observation period, data were collected primarily on the number of audible alarms for each alarm category along with their criticality and nurses' response to those alarms. Task priority was also noted for each record. The frequency of phone calls, call lights and med-door-open alarms were also recorded with the response times. The noise level at the central monitoring station was collected using a noise meter for each alarm. The number of patients and number of nurses were recorded from the floor and daily staff schedule. The response times were collected using stop watches.

Mental Workload Measurement

This study was designed to create a working definition for alarm fatigue due to audible alarms in the ICU of an acute care hospital. This fatigue is correlated to information processing skills of humans and their tolerance of noise level. Therefore, mental workload is a very important response to collect for this research.

NASA TLX Measures: The National Aeronautics and Space Administration-Task Load Index (NASA-TLX), developed by Hart and Staveland (1988) is a trusted tool which assesses the mental workload imposed by a task. The NASA-TLX includes a two-part evaluation procedure making absolute and comparative ratings of a task performed by the participants based on six separate factors contributing to perceived workload. The six workload factors established by Hart and Staveland through factor analytic procedures are defined in Table 7.

These subscales can be measured on their own (“Raw TLX”) or combined in a pair-wise comparison that measures the relative importance of each subscale to the participants (Hart & Staveland, 1988). According to Hart (2006), this Raw TLX is simpler to use and gives similar results for the total mental workload score using weighing method. In this study the Raw TLX was used to measure mental workload. The NASA TLX forms (attached in Appendix D) were distributed at the end of the observed shifts to all the staff involved to collect their comments about the overall shift workload.

Participants were instructed to read a description of each scale before making any ratings. All the participants rated their experience with a given task on 20-step scales for each of the six dimensions. A score from 0 to 100 (rounded to the nearest 5 points) was obtained on each scale. An overall mental workload score for NASA TLX was obtained on a scale of 0 through 100 for each participant by calculating an average of those six ratings for six dimensions.

Table 7: Rating scale definitions for the six dimensions of NASA TLX

Dimension	Description
Mental Demand	How much mental and perceptual activity was required for these tasks (e.g., for monitoring, thinking, deciding, calculating, remembering, notifying, etc.)?
Physical Demand	How much physical activity was required (e.g. for controlling, activating, running from place to place, etc.)?
Temporal Demand	How much time pressure was felt due to the rate at which the task elements occurred? Was the rate slow and leisurely or rapid and frantic?
Performance	How successfully were the goals of the task accomplished the goals of the task and what was the satisfaction level with the performance in accomplishing these goals?
Effort	What was the level difficulty of the work (mentally and physically) required to accomplish the level of performance?
Frustration Level	How much insecurity, discouragement, irritation, stress, and annoyance versus security, gratification, contentment, relaxation, and complacency was felt during the task?

SWAT Measures: The Subjective Workload Assessment Technique (SWAT) is another multidimensional tool with three dimensions for time load (T), mental effort load (E), and psychological stress load (S) at three discrete levels: low, medium and high. SWAT was carried out in three steps. The first step was to scale development in order to rank the three different dimensions. During this stage, the participants were asked to do pair-wise comparison for three pairs (attached in Appendix D) of three dimensions according to his or her perceptions of increasing workload. Following scale development, in the second step, participants were asked to rate (1, 2 and 3 as Low, Medium and High respectively) the alarm monitoring task with regard to time load, mental effort load and psychological stress load dimensions. In the third step, each of the three dimension

scorings was converted into a numeric score from 0 to 100 using the scale developed in the first step.

Table 8: Rating scale definitions for the six dimensions of SWAT

Dimension	Measure	Description
Time Load	Low	Often have spare time. Interruptions or overlap among activities occur infrequently or not at all
	Medium	Occasionally have spare time. Interruptions or overlap among activities occur intermittently
	High	Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time
Mental Effort Load	Low	Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention
	Medium	Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required
	High	Extensive mental effort or concentration required. Very complex activity requiring total attention
Physiological Stress Demand	Low	Little confusion, risk, frustration, or anxiety exists and can be easily accommodated
	Medium	Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant willpower is required to maintain adequate performance
	High	High to very intense stress due to confusion, frustration, or anxiety. Extreme determination and self-control required

The SWAT rating can be used in two ways: during the tasks and/or after the completion of the task (Reid, 1989). This study used SWAT during the tasks to obtain an instant feeling of mental workload for the staff. Table 8 shows all the dimension ratings for SWAT and their definitions. The ICU staff was asked several times during the observation hours to rate three dimensions of SWAT. Averages were calculated for each of the participants for each observation slot.

Personality Type Survey

The Big Five Factor personality test is available online and very easy to take. All the participants received the test with the link (<http://www.outofservice.com/bigfive/>) via email. The test has forty-five basic questions to decide personality type based on the five personality dimensions: Extraversion, Openness, Agreeableness, Conscientiousness, and Neuroticism. There are also four demographic questions that are considered helpful in deciding the personality type properly. Each test taker sent the link of his/her result page to the principal researcher which only gives the personality type, not the test answers. The personality type dimensions for each of the staff were recorded along with participant ID for the convenience of data analysis.

Affect Measurement

In this study, three different types of affect were measured for all participants. Those three types of affect were boredom, apathy and distrust. These three measures were considered to define alarm fatigue along with the mental workload. Three questionnaires to collect data on these three measures are attached on Appendix E. The procedures to collect data on these three types of affect and the technique to calculate the measures are described in following sections.

Boredom: The shorter version of the Boredom Proneness Scale (BPS) with ten questions was selected for this study. This questionnaire was given to each participant after the completion of their shift so that they could fill it out without interrupting their

duties. The participants were requested to rate all ten questions based on that shift experience. Ratings were performed on a five-point Likert scale from "1" (strongly disagree) to "5" (strongly agree). Boredom was calculated out of 100 for each participant by summing up all the ratings from one survey and multiplying the total number by two.

Apathy: The shorter version of the Apathy Evaluation Scale (AES) with ten questions was given to each participant after the completion of their shift along with the boredom survey. The participants were requested to rate all ten questions based on that shift experience. Ratings were performed on a four-point Likert scale from "1" (strongly disagree) to "4" (strongly agree). Apathy was calculated out of 100 for each participant in order to normalize all the dependent variables. AES items are scored in such a way so that high AES scores indicate more apathy, i.e. less motivation. This requires recoding the item which is stated with positive (+) or "healthy" syntax (DCRC). Therefore, the fourth question was recoded. Recoding means changing item codes so that 1=4, 2=3, 3=2, 4=1.

Distrust: Three questions created for distrust were asked to the staff during observation without interrupting their tasks. They were requested to give answers as "Yes", "I don't know" or "No" which were rated as "1", "2" and "3" respectively. These responses were calculated in the scale of 100. Similar to the AES, for the distrust scale, high scores indicate more distrust. Item three was recoded during calculation of distrust.

Performance Measures

Three performance measures were collected during direct observations for each alarm. The purpose was to investigate the influence of alarm fatigue on these performance measures and compare the findings with previous research, and verify the conceptual model considered by this study.

Statistical Analyses

Table 9: Factors and measures considered in this study

Type of variable	Name of variables with corresponded levels
Factors	1. Alarm Type (14 Levels) 2. Alarm Criticality (Three Levels: Low, Medium, High) 3. Task Priority (Five Levels)
Continuous variables	4. Number of Alarms 5. Workload as Nurse to Patient ratio 6. Time Elapsed (since the start of the shift) 7. Noise Level 8. Extraversion 9. Conscientiousness 10. Neuroticism 11. Agreeableness 12. Openness
Blocking factors	1. Shift (Day / Night) 2. Staff ID
Alarm fatigue measures	1. Mental workload: SWAT, NASA TLX 2. affect: Boredom, Apathy, Distrust
Performance measures	1. Response (yes/no) 2. Response Time (for the responded alarms) 3. Number of Ignored Alarms

This observational study found seven independent variables from literature review: number of alarms, workload (represented by staff to patient ratio), alarm type,

alarm criticality, time elapsed since the start of the shift, noise level, and personality type; two blocking variables: shift periods and participants; one variables from Hierarchical Task Analyses: task priority; two response variables: mental workload and affect and three performance measures: response to alarms (yes/no) (found from HTA), number of ignored alarms, and response time for responded alarms. Table 9 shows all the categorical factors and continuous variables. The analyses for unit clerks and other nurses were performed separately as the ratio of the response to alarm for unit clerk and nurses varies significantly.

Several statistical analyses were performed for defining alarm fatigue in this study. The first approach was to do regression analyses in order to find the significance of the factors with response variables. As there were multiple categorical factors, blocking factors, and continuous variables in this study, correlation analyses were necessary for the data pertaining to both unit clerks and nurses. If two variables had a correlation larger than 0.9, one of them was removed from the further analyses.

Before doing the regression analyses, the normality, linearity and homogeneity assumptions were tested for all the five alarm fatigue measures and three performance measures. Box-Cox and Johnston transformations were applied if needed. In the case of failure to normality, linearity, and homogeneity assumptions, a variable reduction analyses, Principal Component Analysis (PCA) was conducted for five alarm fatigue measures with all the independent variables and for three performance measures with five alarm fatigue measures. PCA is a powerful tool for reducing a number of observed

variables into a smaller number of artificial variables combining all the similar factors together, which account for most of the variance in the data set. PCA identified the cluster of artificial variables which were used in regression analyses after the interpretation of the variance of factors on each new component. On the case of any failure of producing successful models with the components extracted from PCA, finally, non-parametric, non-linear regression analyses were designed to be performed using Multivariate Adaptive Regression Splines (MARS). This is a software that identified the important variables and their importance scoring, in order to determine the significance of each alarm fatigue measure with respect to all the independent variables (working conditions and staff individuality) and performance measures with respect to (a) alarm fatigue measures, (b) factors including working conditions and staff individuality, and (c) combined including both working conditions and alarm fatigue measures.

This study was designed to define alarm fatigue in terms of all the possible contributing factors to that fatigue. The research will be helpful to assess the alarm fatigue using that definition for all the hospitals across the USA in order to meet TJC's requirement for the improvement of clinical alarm management.

RESULTS

Clinical Alarm Survey

The clinical alarm survey created by the Healthcare Technology Foundation was distributed to all potential participants in the Intensive Care Unit (ICU) at Bozeman Deaconess Hospital (BDH). The purpose of that survey was to assess the staff attitude toward existing alarm monitoring system on that facility. Seventeen out of thirty seven responses were received. All the responses were analyzed together: 16% of the total responses were from unit clerks and the remaining 84% were from nurses. Table 10 shows important responses to the alarm related information section for the existing alarm monitoring system. Figure 3 shows the responses from the staff to the issues concerning alarms in ICU at BDH. The responses from Table 10 and Figure 3 would reveal the necessity of conducting further research in the chosen facility.

Table 10: Responses by percentages to the alarm related information

Statements	Agree and Strongly Agree	Neutral	Disagree and Strongly Disagree
1. The purpose of clinical alarms is to alert staff of hazardous patient condition	100.0%	—	—
2. Alarm sounds and/or visual displays should differentiate alarm priority	100.0%	—	—
3. Alarm sounds and/or visual displays should be distinct based on source	84.6%	15.4%	—
4. Alarms should impact multiple senses (audible, visual, proprioceptive, etc.)	92.3%	7.7%	—
5. Nuisance alarms occur frequently	84.6%	7.7%	7.7%
6. Nuisance alarms disrupt patient care	84.6%	15.4%	—

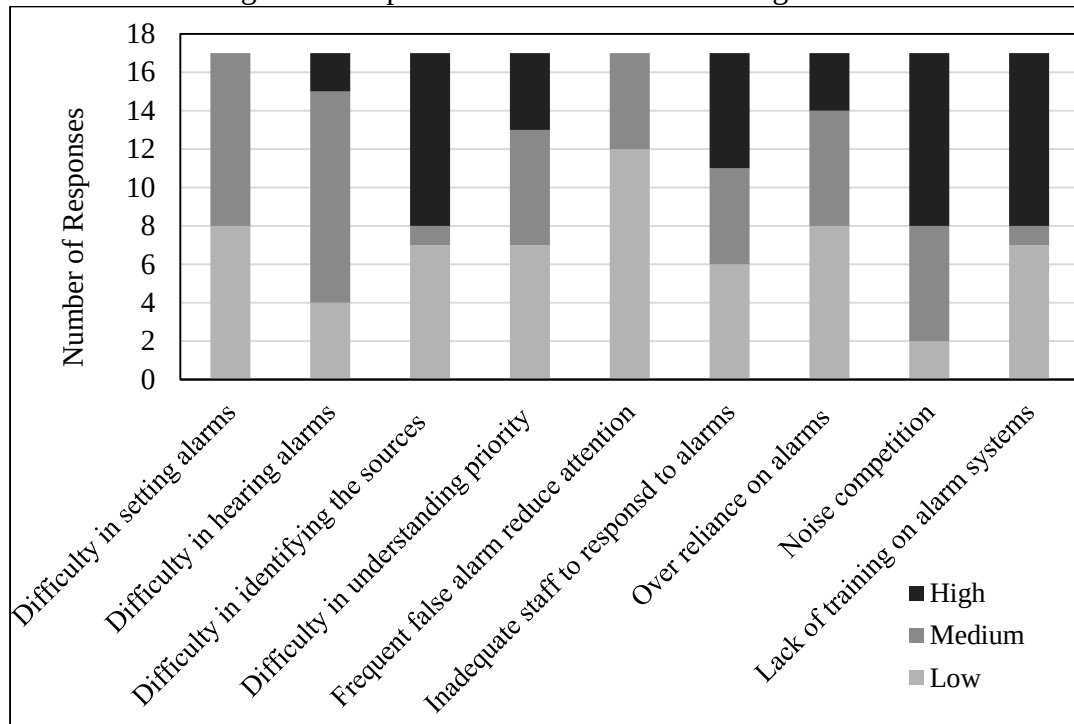
Table 10: Responses by percentages to the alarm related information
(continued)

Statements	Agree and Strongly Agree	Neutral	Disagree and Strongly Disagree
7. Nuisance alarms reduce trust in alarms and cause care givers to turn them off	53.8%	46.2%	—
8. The alarms used on my floor/area are adequate to alert staff	69.2%	15.4%	15.4%
9. There have been frequent instances where alarms could not be heard	53.8%	7.7%	38.5%
10. The staff is sensitive to alarms and responds quickly	30.8%	23%	46.2%
11. It can be confusing to determine which device is in alarm	46.2%	15.3%	38.5%
12. A central alarm management staff is helpful	69.2%	7.7%	23.1%
13. Alarm integration and communication systems via wireless device (e.g., pager, cell phone) are useful in improving alarm management and response	23.1%	53.8%	23.1%
14. Smart alarms would be effective in reducing false alarms and improving response	84.6%	15.4%	—
15. Policies and procedures exist within the facility to regulate alarms and they are followed	23.1%	30.7%	46.2%
16. There is a requirement in your institution to document that the alarms are set and are appropriate for each patient	30.8%	23%	46.2%

The second section of this survey collected data on the issues concerning alarms. This section provided insights into how the ICU staff rated the relative influences of the various challenges they faced with clinical alarm management. The respondents rated each response here from 1 to 9. For the convenience of analysis, these ratings were

categorized as ‘Low’ (for ratings 1, 2, and 3), ‘Medium’ (for ratings 4, 5, and 6), and ‘High’ (for ratings 7, 8, and 9) Figure 3 shows all the responses for this section.

Figure 3: Responses to the issues concerning alarms



Hierarchical Task Analyses

The hierarchical task analyses (HTA) was performed for ICU staff. This task analysis helped to find influencing factor/factors that can cause alarm fatigue which was not found from literature review. HTA for alarm monitoring were conducted for both nurses and unit clerks at the ICU. Two HTA diagrams, one for nurses and one for unit clerks were drawn and reviewed by two nurses, one unit clerk and the ICU manager at BDH. These hierarchies are shown in Figure 4 and 5.

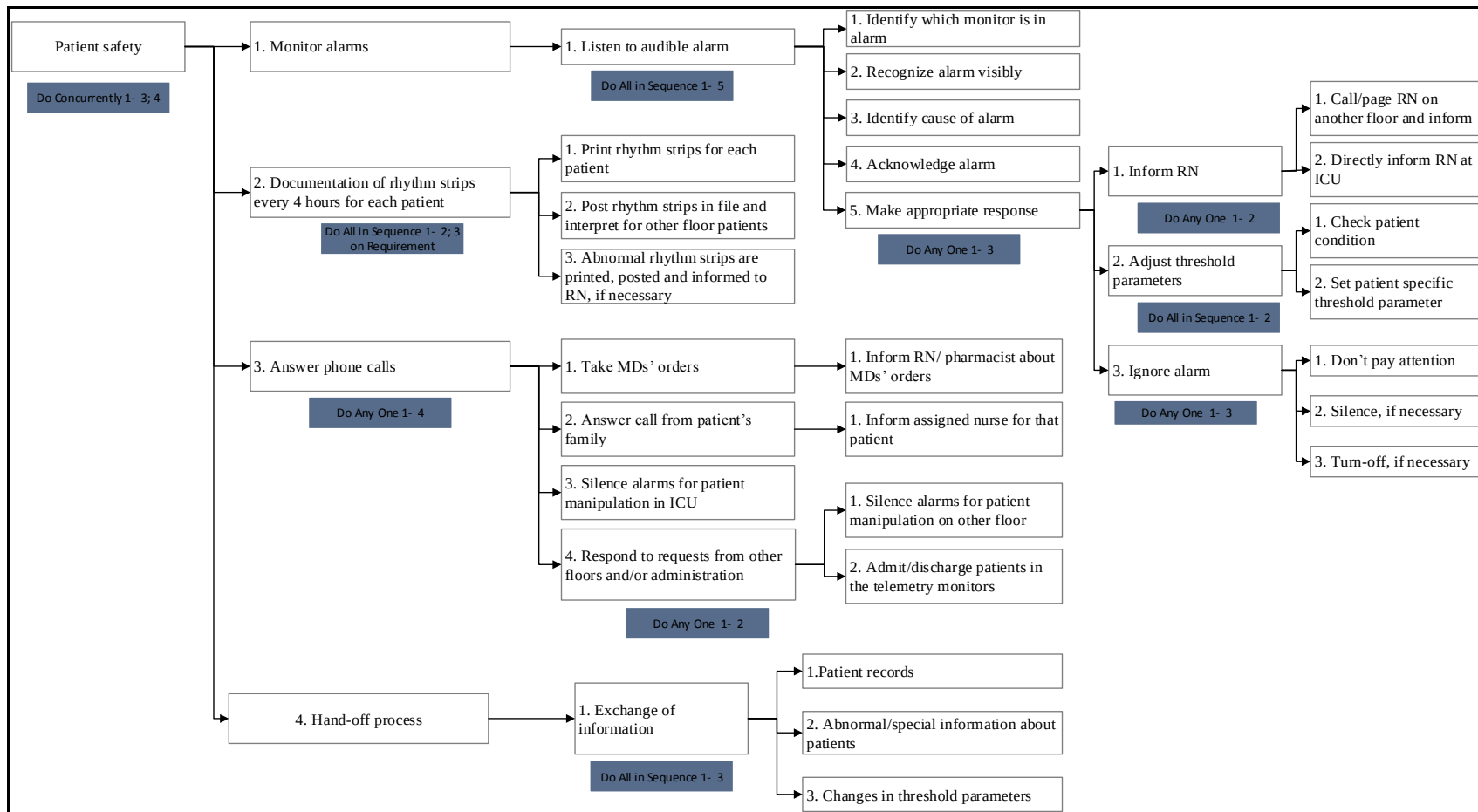


Figure 4: Hierarchical Task Analysis for unit clerks in ICU

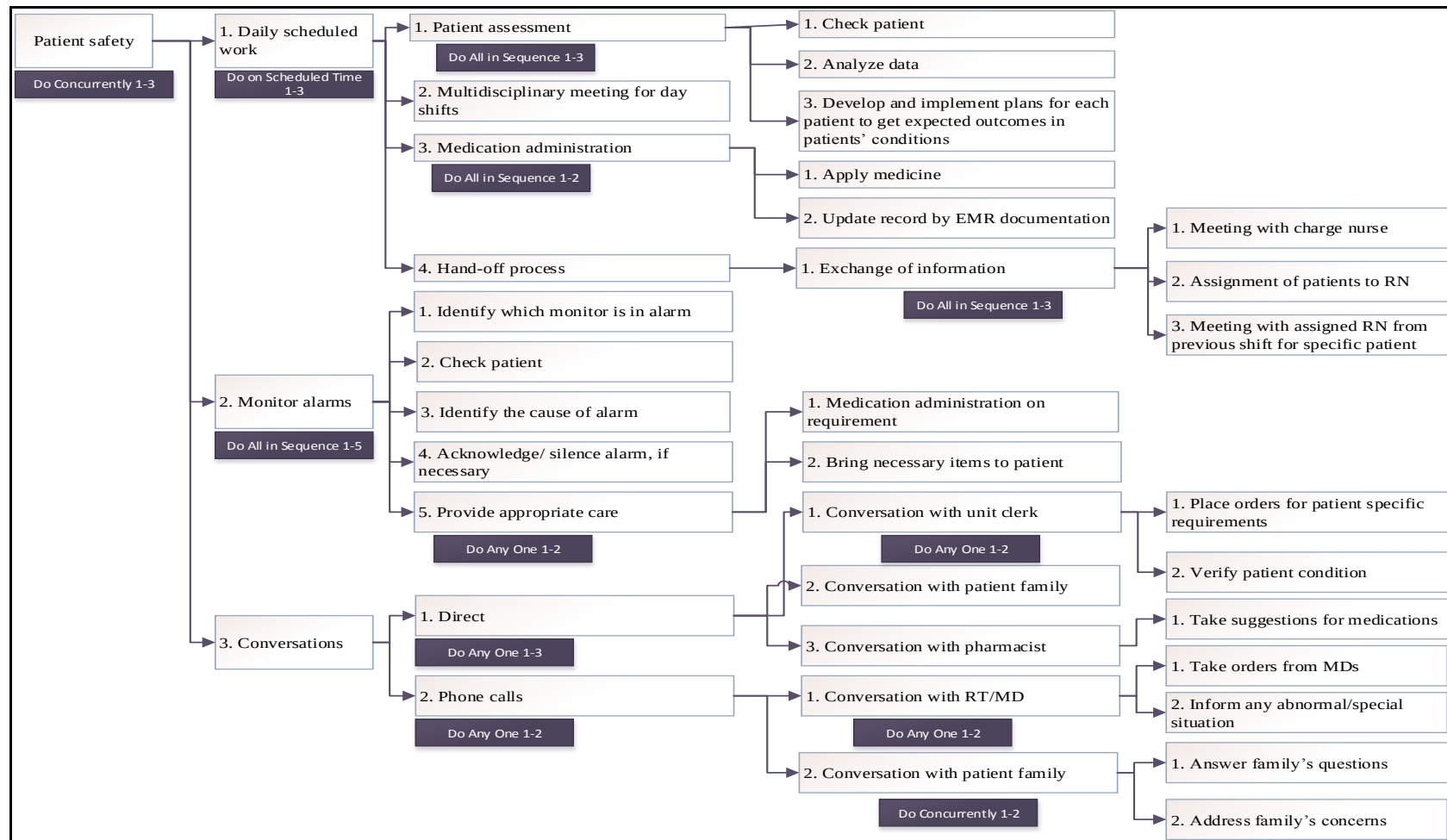


Figure 5: Hierarchical Task Analysis for nurses in ICU

Descriptive Statistics

Descriptive statistics for alarm fatigue measures (mental workload and affect measures) and performance measures (response, response time to each alarm, and number of ignored alarms) collected during the study are summarized in Table 11 for both unit clerks and nurses. These statistics show the average values for alarm fatigue measures and performance measures along with 95% confidence intervals (CI) for them. The mental workload and the affect for unit clerks are found to be higher during night shifts than those during day shifts, whereas for nurses most of these measures are higher during day shifts.

Table 11: Descriptive statistics for alarm fatigue measures and performance measures

Participant	Measures for alarm fatigue	Day Shift	
		Mean (SD)	95% Confidence Interval
Unit Clerk	SWAT	49.38 (11.72)	(48.44, 50.31)
	NASA TLX	43.77 (18.17)	(42.32, 45.22)
	Distrust	60.04 (9.04)	(59.32, 60.76)
	Boredom	62.70 (1.42)	(62.59, 62.81)
	Apathy	81.88 (7.04)	(81.31, 82.44)
	Response Time	5.04 (5.86)	(4.37, 5.72)
	Percentages of Ignored Alarms		52.48%
Nurse	SWAT	53.58 (23.51)	(45.31, 61.86)
	NASA TLX	61.27 (8.30)	(57.32, 65.21)
	Distrust	51.61 (11.93)	(47.41, 55.81)
	Boredom	59.35 (4.86)	(57.04, 61.66)
	Apathy	83.79 (9.84)	(79.12, 88.47)
	Response Time	6.5 (3.66)	(4.43, 8.57)
	Percentages of Ignored Alarms		29.41%

Table 11: Descriptive statistics for alarm fatigue measures and performance measures
(continued)

Participant	Measures for alarm fatigue	Night Shift	
		Mean (SD)	95% Confidence Interval
Unit Clerk	SWAT	50.79 (12.34)	(49.63, 51.96)
	NASA TLX	45.88 (10.69)	(4.87, 46.88)
	Distrust	72.30 (5.56)	(71.78, 72.82)
	Boredom	66.60 (3.56)	(66.26, 66.93)
	Apathy	88.82 (4.28)	(88.42, 89.22)
	Response Time	6.85 (11.33)	(5.27, 8.43)
	Percentages of Ignored Alarms	53.23%	
Nurse	SWAT	45.88 (10.44)	(42.20, 49.55)
	NASA TLX	47.86 (16.60)	(42.02, 53.71)
	Distrust	57.53 (4.66)	55.89, 59.17)
	Boredom	64.65 (6.64)	(62.31, 66.98)
	Apathy	83.03 (10.15)	(79.46, 8.60)
	Response Time	9.47 (9.04)	(5.17, 13.77)
	Percentages of Ignored Alarms	45.16%	

Figure 6 shows the change in mental workload (measured by SWAT), affect (measured as distrust level) and performance (measured as percentage of ignored alarms) for the first and the second half of the shifts for unit clerks. Similar trends in the bar charts for these three variables would reveal that mental workload and affect influence staff performance during alarm monitoring. Although no statistical significance was observed. The surveys and observations showed this similar trends and resulting influences on staff performance for unit clerks, but not for nurses. The figure for nurses with these results is attached in appendix E.

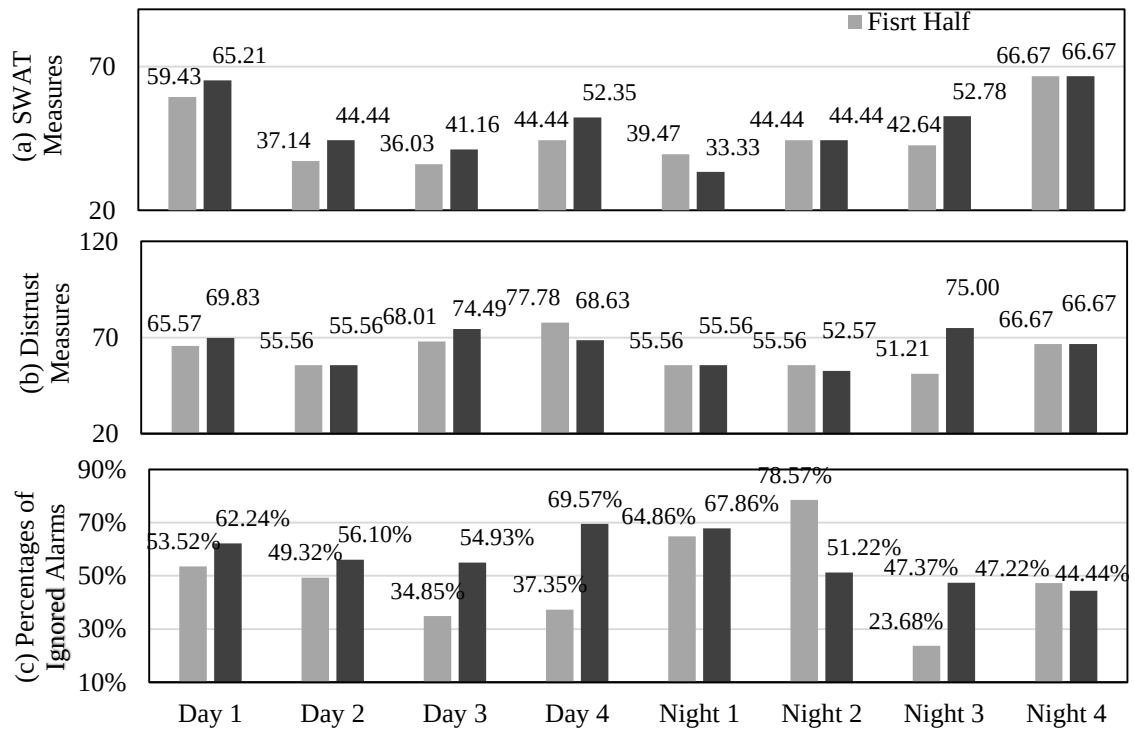


Figure 6: Effect of alarm fatigue on alarm response for unit clerks

Research Question 2- Definition of Alarm Fatigue as Mental Workload

Alarm fatigue is an important issue in healthcare for acute care medical floors.

Staff is overwhelmed by working conditions during alarm monitoring. This study measured this sense of being overwhelmed in terms of mental workload.

Regression Analyses to Define Alarm Fatigue as Mental Workload

Because of having considered a large number of factors in this study, correlation analyses were performed for both unit clerks and nurses with the factors measured as working conditions and staff individualities. These analyses are shown in Appendix F1.

The factors do not show any correlation among each other. Table 12 and 13 summarize the tests of normality and linearity assumptions for two mental workload measures and their appropriate regression models for unit clerk and nurse. Multivariate analyses of regression splines were also conducted for the models for which linear regressions were possible. The tables list the significant variables found from regression analyses, using the * symbol to mark those variables which are common in both models (linear regression and MARS). Diagnostics plots drawn for linear models exhibit the graphs for homogeneity and normality assumptions.

Table 12: Summary of regression analyses to define alarm fatigue as SWAT measures

Regression Model	Unit Clerk		Nurse	
	Real Data	After Transformation	$R^2 = 0.729$ $RMSE = 10.324$ Normality Test (Shapiro-Wilk): $p = 0.054$	
Linear Regression	$R^2 = 0.876$	$\lambda = 2.00$	Variables that have influence	P value
	$RMSE = 4.229$	$R^2 = 0.889$	Shift	0.034
	Normality Test: $p = 0.0001$	$RMSE = 412.531$	Number of Alarms*	0.009
		Normality Test: $p = 0.0003$	Alarm Criticality	0.024
			Noise Level*	0.019
			Agreeableness*	0.039
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.997$		$R^2 = 0.993$	
	$RMSE = 0.64068$		$RMSE = 1.56$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Shift	100	Openness	100
	Workload	48.72	Neuroticism	87.42
	Time Elapsed	45.36	Noise Level*	79.16
	Staff ID	16.02	Workload	69.58
	Task Priority	14.29	Time Elapsed	49.10
	Agreeableness	9.90	Number of Alarms*	27.54
	Noise Level	6.08	Extraversion	27.21
	Alarm Type	5.38	Agreeableness*	25.34

* Indicates the significant variable found in both linear and the MARS models

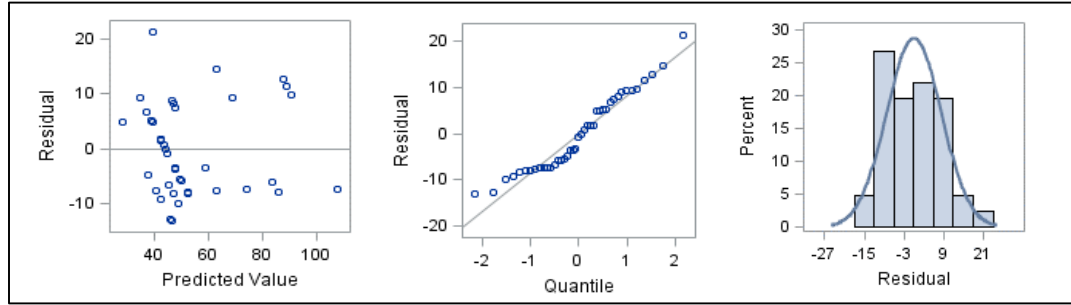


Figure 7: Diagnostics plots for SWAT measures with nurses

Table 13: Summary of regression to define alarm fatigue as NASA TLX measures

Regression Model	Unit Clerk		Nurse	
Linear Regression	Real Data	After Transformation	$R^2 = 0.757$ $RMSE = 10.039$ Normality Test (Shapiro-Wilk): $\rho=0.749$	
	$R^2 = 0.760$	$\lambda = 4.27$	Variables that have influence	P value
	$RMSE = 7.6$	$R^2 = 0.899$		
	Normality Test: $\rho= 0.0003$	$RMSE = 5.9517$	Staff ID	0.0138
		Normality Test: $\rho<0.0001$	Number of Alarms*	0.0252
			Alarm Type	0.0133
			Alarm Criticality	0.0038
			Extraversion*	0.001
		Openness	0.012	
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.916$		$R^2= 0.97$	
	$RMSE = 4.495$		$RMSE= 2.72$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Agreeableness	100	Agreeableness	100
	Shift	62.55	Conscientiousness	84.98
	Time Elapsed	56.57	Noise Level	69.01
	Workload	21.62	Task Priority	60.53
	Extraversion	20.53	Neuroticism	47.97
	Noise Level	14.63	Extraversion*	37.44
	Number of Alarms	12.96	Time Elapsed	26.99
	Alarm Type	11.32	Number of Alarms*	17.89
Task Priority	8.03	Workload	16.03	

* Indicates the significant variable found in both linear and the MARS models

As compared with the linear models, the MARS models have better statistics for R-square and Root Mean Square Error (RMSE). Workload, time elapsed, noise level, and some personality dimensions have influence on both SWAT and NASA TLX measures.

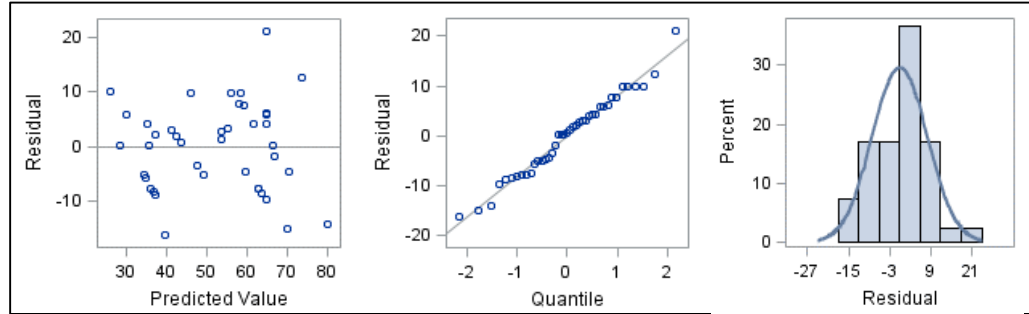


Figure 8: Diagnostics plots for NASA TLX measures with nurses

Research Question 3- Definition of Alarm Fatigue as Affect

Existing working conditions along with staff individualities can lead to desensitization and indifference. The large number of false alarms, noisy environment, and workload (Staff: Patient) along with other factors can create distrust toward alarms and can make the staff feel bored or apathetic. Therefore, one of the research objectives of this study was to define alarm fatigue in terms of affect, including distrust, boredom, and apathy.

Regression Analyses to Define Alarm Fatigue as Affect

Table 14, 15, and 16 summarize all the tests for normality and linearity assumptions for these three types of affect and their appropriate regression analysis models for both unit clerks and nurses. The significant variables from regression analyses are also listed in those tables with available statistics from each model. Diagnostics plots were drawn for linear models which exhibit the tests for homogeneity and normality assumptions.

Table 14 shows that personality type has influence on boredom for both participants. MARS shows a better model for nurses as compared with linear model which generated less RMSE value.



Table 15: Summary of regression analyses to define alarm fatigue as Apathy

Regression Model	Unit Clerk		Nurse	
	Real Data	After Transformation		
Linear Regression	$R^2 = 0.977$	$\lambda = 5.00$	$R^2 = 0.934$	
	$RMSE = 1.047$	$R^2 = 0.9764$	$RMSE = 2.622$, Normality Test: $\rho=0.723$	
	Normality Test	RMSE =	Variables that have influence	P value
	$P < 0.0001$	258692260	Shift	0.0007
		Normality Test	Staff ID	0.0066
		$\rho < 0.0001$	Workload*	0.0235
			Alarm Type*	<.0001
			Alarm Criticality	0.0120
			Noise Level*	0.0442
			Time Elapsed*	0.0004
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 1.00$		$R^2 = 0.986$	
	$RMSE = 6.92E-6$		$RMSE = 1.67$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Conscientiousness	100	Extraversion	100
	Shift	61.12	Openness*	81.70
			Workload*	57.63
			Noise Level*	52.32
			Conscientiousness*	49.22
			Alarm Type*	31.40
			Task Priority	27.01
			Time Elapsed*	20.87

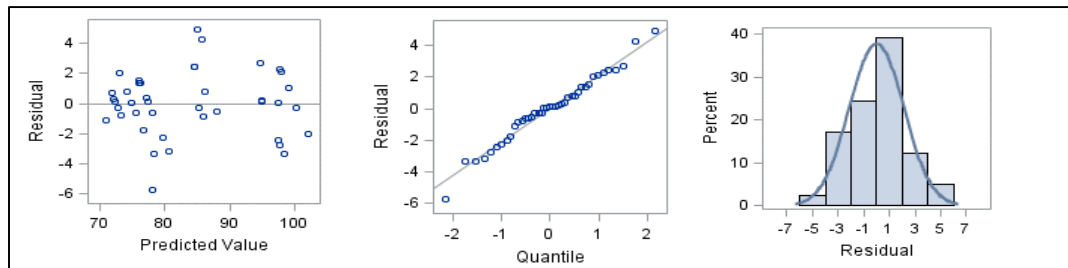


Figure 10: Diagnostics plots for Apathy in nurses

From Table 15, conscientiousness is found to be an important factor leading staff eventually to apathy during alarm monitoring. Table 16 exhibits that a large number of alarms creates distrust in staff towards alarms.

Table 16: Summary of regression analyses to define alarm fatigue as Distrust

Regression Model		Unit Clerk	Nurse	
Linear Regression	Real Data	After Transformation	$R^2 = 0.729$ $RMSE = 3.911$ Normality Test (Shapiro-Wilk): $p=0.120$	
	$R^2 = 0.896$ $RMSE = 3.177$ Normality Test (Shapiro-Wilk): $P<0.0001$	$\lambda=1$ No transformation was possible	Variables that have influence	P value
			Shift	0.0002
			Staff ID	0.0019
			Workload	0.0239
Multivariate Adaptive Regression Splines (MARS)			Noise Level*	0.0096
			Alarm Type	0.0348
	$R^2 = 0.998$ $RMSE = 0.37378$		$R^2 = 0.98$ $RMSE = 0.739$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Shift	100	Extraversion	100
	Workload	31.67	Noise Level*	97.28
	Time Elapsed	24.86	Number of Alarms	76.11
	Number of Alarms	10.04	Time Elapsed	47.50
	Agreeableness	6.50	Agreeableness	41.26
			Task Priority	31.50

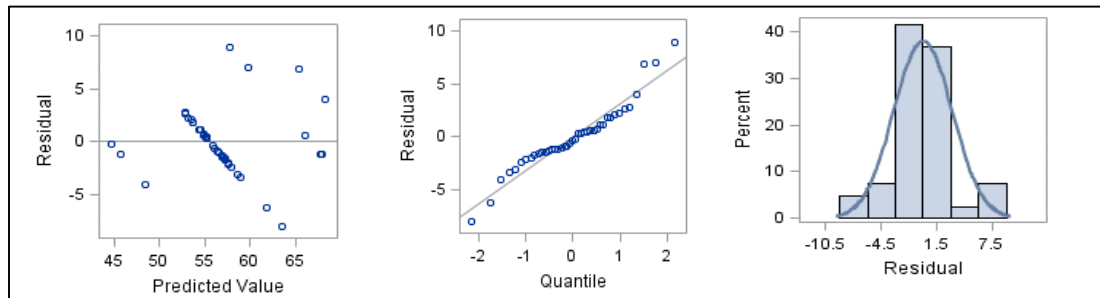


Figure 11: Diagnostics plots for Distrust in nurses

Principal Component Analyses for Alarm Fatigue Causing Factors

Principal Component Analysis (PCA) was performed with all the factors considered as working conditions to create fewer components and achieve better models for linear regression. Results are attached in Appendix G (Table G1, Table G2, Table G3,

and Figure G1). Five components having eigenvalues greater than one (Table G1 and Figure G1) were extracted for both groups. Linear regression analyses conducted with those components are summarized in Table 17 with their linearity and normality assumptions. Influencing components along with the important variables are also listed in this table. The summary exhibits that, for unit clerks, all the measures of alarm fatigue failed normality assumptions, whereas for nurses, SWAT and boredom gave linear regression with low R^2 values and high Root Mean Square Errors. Other three models give good associations for nurses.

Table 17: Summary of regression with components extracted from PCA

	Alarm fatigue Measures	Statistics	Principal Components that have influence	Variables that have influence
Unit Clerk	SWAT	$R^2 = 0.876$ $RMSE = 4.22$ Normality Test: $\rho < 0.0001$	Violation of Normality Assumption	
	NASA TLX	$R^2 = 0.760$ $RMSE = 7.59$ Normality Test: $\rho < 0.0001$	Violation of Normality Assumption	
	Boredom	$R^2 = 0.993$ $RMSE = 0.269$ Normality Test: $\rho < 0.0001$	Violation of Normality Assumption	
	Apathy	$R^2 = 0.977$ $RMSE = 1.05$ Normality Test: $\rho < 0.0001$	Violation of Normality Assumption	
	Distrust	$R^2 = 0.896$ $RMSE = 3.177$ Normality Test : $\rho < 0.0001$	Violation of Normality Assumption	

Table 17: Summary of regression with components extracted from PCA (continued)

	Alarm Fatigue Measures	Statistics	Principal Components that have influence	Variables that have influence
Nurse	SWAT	$R^2 = 0.544$ $RMSE = 13.398$ Normality Test: $\rho=0.045$	PC4, PC5	Shift Staff ID Task Priority Number of Alarms Extraversion
	NASA TLX	$R^2 = 0.371$ $RMSE = 13.014$ Normality Test: $\rho=0.0001$	Violation of Linearity and normality Assumptions	
	Boredom	$R^2 = 0.668$ $RMSE = 3.869$ Normality Test: $\rho=0.60$	PC1, PC3, PC4, PC5	Shift Staff ID Workload Alarm Type Alarm Criticality Time Elapsed Task Priority Number of Alarms Neuroticism Openness Conscientiousness Agreeableness
	Apathy	$R^2 = 0.275$ $RMSE = 8.7013$ Normality Test: $\rho=0.068$	Violation of Linearity Assumption	
	Distrust	$R^2 = 0.283$ $RMSE = 5.128$ Normality Test : $\rho=0.259$	Violation of Linearity Assumption	

Research Question 4- Comparison between Two Definitions

All the influencing variables from working conditions and staff individualities used to create alarm fatigue definitions in terms of mental workload and affect are presented in Table 18 for unit clerks and in Table 19 for nurses. These lists of variables

illustrate the comparison between the two definitions. Definitions with mental workload and affect for unit clerks, as shown in Table 18, have six common variables. Both the definitions contain some positive dimensions of personality type. Unit clerks are influenced by working conditions in defining alarm fatigue as mental workload, whereas the definition as affect, they are more influenced by staff individualities.

Table 18: Summarization of influencing variables from two definitions pertaining to unit clerks

SWAT	NASA TLX	Boredom	Apathy	Distrust
Shift	Agreeableness	Shift	Shift	Shift
Staff ID	Shift	Extraversion	Conscientiousness	Workload
Workload	Time Elapsed	Openness		Time Elapsed
Alarm Type	Workload			Number of
Noise Level	Extraversion			Alarms
Time Elapsed	Noise Level			Agreeableness
Task Priority	Number of Alarms			
Agreeableness	Alarm Type			
	Task Priority			
Influencing variables for mental workload		Influencing variables for affect		
Shift**		Shift**		
Staff ID		Number of Alarms**		
Workload**		Workload**		
Number of Alarms**		Time Elapsed**		
Alarm Type		Extraversion**		
Noise Level		Openness		
Time Elapsed**		Agreeableness**		
Task Priority		Conscientiousness		
Extraversion**				
Agreeableness**				

** Indicates the factors common to mental workload and affect

For nurses, Table 19 shows two analogous definitions containing all the possible factors considered as working conditions and staff individualities. Although for the definition as mental workload, the SWAT measures specify the influence of only few alarm-related factors along with shift and a personality type dimension, NASA TLX

includes almost all of the variables. In the case of the definitions as affect, apathy includes all of the variables considered for this study.

Table 19: Summarization of influencing variables from two definitions pertaining to nurses

SWAT	NASA TLX	Boredom	Apathy	Distrust
Shift	Staff ID	Number of	Shift	Shift
Number of	Workload	Alarms	Staff ID	Staff ID
Alarms	Number of	Workload	Workload	Number of
Alarm Criticality	Alarms	Time Elapsed	Alarm Type	Alarms
Noise Level	Alarm Type	Noise Level	Alarm Criticality	Workload
Agreeableness	Alarm Criticality	Extraversion	Noise Level	Alarm Type
	Noise Level	Openness	Time Elapsed	Noise Level
	Time Elapsed	Conscientiousness	Task Priority	Time Elapsed
	Task Priority	Neuroticism	Extraversion	Task Priority
	Extraversion	Agreeableness	Openness	Extraversion
	Openness		Conscientiousness	Agreeableness
	Agreeableness		Neuroticism	
	Conscientiousness		Agreeableness	
	Neuroticism			
Influencing variables for mental workload		Influencing variables for affect		
Shift		Shift		
Staff ID		Staff ID		
Number of Alarms		Number of Alarms		
Workload		Workload		
Alarm Type		Alarm Type		
Alarm Criticality		Alarm Criticality		
Noise Level		Noise Level		
Time Elapsed		Time Elapsed		
Task Priority		Task Priority		
Extraversion		Extraversion		
Openness		Openness		
Agreeableness		Agreeableness		
Conscientiousness		Conscientiousness		
Neuroticism		Neuroticism		

Research Question 5- Significance of Alarm Fatigue Measures on Staff Performance

Alarm fatigue is considered as an important issue causing a large number of incidents every year in the acute care arena. This study aims to define alarm fatigue and

verify these definitions by identifying their relationship to alarm monitoring performances.

Relationship between the Alarm Fatigue Measures and the Performance Measures

According to the conceptual model for this study, alarm fatigue is the result of a number of influencing factors which are considered as working conditions and staff individualities. In turn, alarm fatigue can deteriorate staff performance and impair patient care. The following section assess the validity of the conceptual model.

Regression Analyses for Staff

Performances Using Alarm Fatigue Measures: The correlation Analyses of alarm fatigue measures presented in Appendix F (see Table F2) do not give any evidence of correlation between any two of variables.

Table 20: Summary of regression for Response (yes/no) using alarm fatigue measures

Regression Model			Unit Clerk		Nurse	
Binary Logistics	Real Data	After Transformation	Real Data	After Transformation	Real Data	After Transformation
	$R^2 = 0.043$	$\lambda = 0.50$	$R^2 = 0.349$	$\lambda = -2.00$	$R^2 = 0.349$	$\lambda = -2.00$
	$RMSE = 0.032$	$R^2 = 0.043$	$RMSE = 0.309$	$R^2 = 0.214$	$RMSE = 0.309$	$R^2 = 0.214$
	Normality Test: $\rho < 0.0001$	$RMSE = 0.032$ Normality Test: $\rho = 0.0001$	Normality Test: $\rho = 0.14$	$RMSE = 7538$ Normality Test: $\rho = 0.069$	Normality Test: $\rho = 0.14$	$RMSE = 7538$ Normality Test: $\rho = 0.069$
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.014$ $RMSE = 0.4956$		$R^2 = 0.21$ $RMSE = 0.4407$		$R^2 = 0.21$ $RMSE = 0.4407$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Boredom	100	SWAT	100	SWAT	100
	Apathy	100	Apathy	71.26	Apathy	71.26
			NASA TLX	71.26	NASA TLX	71.26

Table 21: Summary of regression for Response Time using alarm fatigue measures

Regression Model	Unit Clerk		Nurse	
	Real Data	After Transformation	Real Data	After Transformation
Linear Regression	$R^2 = 0.012$	$\lambda = -0.14$	$R^2 = -0.026$	$\lambda = 0.00$
	$RMSE = 8.49$	$R^2 = 0.012$	$RMSE = 7.45$	$R^2 = 0.023$
	Normality Test: $\rho = 0.0001$	$RMSE = 8.49$	Normality Test: $\rho < 0.0001$	$RMSE = 4.03$
		Normality Test (Shapiro-Wilk): $\rho = 0.0001$		Normality Test: $\rho = 0.333$
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.08$		$R^2 = 0$	
	$RMSE = 8.1887$			
	Variables that have influence	Importance scoring		
	SWAT	100		
	NASA TLX	66.96		

Table 22: Summary of regression analyses for Number of Ignored Alarms using alarm fatigue measures

Regression Models	Unit Clerk		Nurse	
	Real Data	After Transformation	Real Data	After Transformation
Linear Regression	$R^2 = 0.231$	$\lambda = -0.50$	$R^2 = 0.084$	$\lambda = 1$
	$RMSE = 27.12$	$R^2 = 0.371$	$RMSE = 0.315$	No transformation was possible
	Normality Test: $\rho = 0.0003$	$RMSE = 85.83$	Normality Test: $\rho = 0.117$	
		Normality Test (Shapiro-Wilk): $\rho < 0.0001$		
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.70$		$R^2 = 0.98$ $RMSE = 0.456$	
	$RMSE = 16.92$			
	Variables that have influence	Importance scoring		
	NASA TLX	100		
	Distrust	23.34	Boredom	100
	SWAT	19.79	SWAT	38.11`

Table 20, 21, and 22 summarize all the tests for normality and linearity assumptions for three performance measures and show the appropriate regression analysis models for both types of participants. The significant variables from regression analyses are also listed in those tables. Tables show violations in either normality or

linearity assumptions for response and response time for both participants, but not for the number of ignored alarms.

Principal Component Analyses

for Alarm Fatigue Measures: Principal Component Analyses were also conducted

for the alarm fatigue measures. Tables G4, G5, and Figure G2 show the principal components with eigenvalues. Linear regression analyses were performed with the two extracted components (from PCA) which are summarized in Table 23. The summary exhibits that, for both unit clerks and nurses, all the performance measures failed both linearity and normality assumptions.

Table 23: Summary of regression with components extracted from PCA

	Performance Measures	Statistics	Principal Components that have influence
Unit Clerk	Response	$R^2 = 0.033$ $RMSE = 0.062$ Normality Test: $p < 0.0001$	Violation of Linearity and Normality Assumption
	Response Time	$R^2 = 0.028$ $RMSE = 27.71$ Normality Test: $p < 0.000$	Violation of Linearity and Normality Assumption
	Number of ignored alarms	$R^2 = 0.197$ $RMSE = 27.71$ Normality Test: $p < 0.0001$	Violation of Linearity and Normality Assumption
Nurse	Response	$R^2 = 0.204$ $RMSE = 0.309$ Normality Test: $p = 0.002$	Violation of Linearity and Normality Assumption
	Response Time	$R^2 = 0.028$ $RMSE = 7.63$ Normality Test: $p = 0.0003$	Violation of Linearity and Normality Assumption
	Number of ignored alarms	$R^2 = 0.014$ $RMSE = 1.24$ Normality Test: $p = 0.008$	Violation of Linearity and Normality Assumption

Regression Analyses for Staff

Performances Using Alarm Fatigue Causing Factors: As the models for staff

performance with respect to the alarm fatigue measures showed very poor linearity and normality for both response and response time, another relationship was tested for the performance measures with all the alarm fatigue causing variables of working conditions and staff individualities. These analyses were used to define alarm fatigue in terms of mental workload and affect. Table 24, Table 25, and Table 26 summarize the regression analyses of performance measures for both unit clerks and nurses using working conditions and staff individualities as independent variables.

Table 24: Summary of regression for Response using alarm fatigue causing factors

Regression Models	Unit Clerk		Nurse	
Binary Logistic Regression	Real Data	After Transformation	$R^2 = 1$ $RMSE = 0.309$ Normality Test (Shapiro-Wilk): $\rho=0.14$	
	$R^2 = 0.154$ $RMSE = 0.062$ Normality Test (Shapiro-Wilk): $\rho<0.0001$	$\lambda = 0.5$ $R^2 = 0.394$ $RMSE = 0.024$ Normality Test (Shapiro-Wilk): $\rho<0.0001$	Variables that have influence	p value
			Workload	0.041
			Alarm Type*	0.013
Multivariate Adaptive Regression Splines (MARS)			Task Priority*	0.009
			Noise Level	0.0003
	$R^2 = 0.349$ $RMSE = 0.402$		$R^2 = 0.99$ $RMSE = 0.04$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Alarm Type	100	Task Priority*	100
	Shift	61.32	Alarm Type*	25.16
	Agreeableness	58.61	Number of Alarms	25.04
	Task Priority	51.50		
	Workload	50.01		
	Alarm Criticality	40.46		
	Time Elapsed	36.08		
	Noise Level	25.55		
	Number of Alarms	23.05		
	Staff ID	9.23		

The summary in Table 24 shows no improvement of response in linearity and normality of regression models (using the linear regression and non-linear, non-parametric regression) for unit clerks, even after transformation. However, nurse data fits for linear regression for this relationship.

Table 25: Summary of regression for Response Time using alarm fatigue causing factors

Regression Models	Unit Clerk		Nurse	
	Real Data	After Transformation	Real Data	After Transformation
Linear Regression	$R^2 = 0.010$	$\lambda = 0.02$	$R^2 = -0.001$	$\lambda = -0.04$
	$RMSE = 8.50$	$R^2 = 0.394$	$RMSE = 7.36$	$R^2 = -0.026$
	Normality Test: $\rho < 0.0001$	$RMSE = 0.089$ Normality Test $\rho < 0.0001$	Normality Test: $\rho = 0.02$	$RMSE = 3.82$ Normality Test $\rho = 0.023$
	$R^2 = 0.797$		$R^2 = 0.896$	
	$RMSE = 3.84$		$RMSE = 2.52$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
Multivariate Adaptive Regression Splines (MARS)	Task Priority	100	Noise Level	100
	Neuroticism	86.07	Neuroticism	83.10
	Alarm Type	43.29	Agreeableness	49.30
	Number of Alarms	41.69	Time Elapsed	30.92
	Agreeableness	34.27	Extraversion	30.91
	Shift	34.08		
	Time Elapsed	25.72		
	Workload	25.22		
	Noise Level	23.53		
	Staff ID	12.24		
	Extraversion	10.42		
	Alarm Criticality	7.49		
	Openness	4.88		

Table 25 illustrates better relationship between response time and working conditions and staff individualities in terms of R-square, RMSE, and normality statistics for both type of participants. Along with all the fatigue causing factors, the list of significant variables exhibits the influence of personality type dimensions on response time which is not evident on response.

Table 26: Summary of regression for Number of Ignored Alarms using alarm fatigue causing factors

Regression Models		Unit Clerk	Nurse	
Linear Regression	Real Data	After Transformation	$R^2 = 0.722$ $RMSE = 0.702$ Normality Test: $\rho=0.904$	
	$R^2 = 0.785$	$\lambda = 0.00$	Variables that have influence	p value
	$RMSE = 14.33$	$R^2 = 0.859$		
	Normality Test: $\rho<0.0001$	$RMSE = 3.42$ Normality Test: $\rho<0.0001$		
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.999$ $RMSE = 0.868$		$R^2 = 0.889$ $RMSE = 0.442$	
	Variables that have influence	Importance scoring	Variables that have influence	Importance scoring
	Time Elapsed	100	Time Elapsed*	100
	Shift	33.68	Noise Level	45.98
	Workload	21.75	Openness	24.52
	Agreeableness	14.19	Task Priority	24.52
	Number of Alarms	12.81	Workload	16.94
	Openness	9.57		
	Task Priority	8.42		

Table 26 also shows a better relationship between the number of ignored alarms and working conditions and staff individualities for both unit clerks and nurses using MARS models and in terms of R-square values and root mean square errors.

Table 27 summarizes all the influencing variables for three performance measures (response [yes/no], response time, and number of ignored alarms). Unit clerks are shown to have more influencing variables for all three than nurses do for relating staff performance with alarm fatigue.

Table 27: Summarization of all the influencing variables for performance measures

Unit Clerk			Nurse		
Response (Yes/No)	Response Time	Number of Ignored Alarms	Response (Yes/No)	Response Time	Number of Ignored Alarms
Alarm Type	Task Priority	Time Elapsed	Workload	Noise Level	Time Elapsed
Shift	Neuroticism	Shift	Alarm Type	Neuroticism	Noise Level
Agreeableness	Alarm Type	Workload	Task Priority	Agreeableness	Openness
Task Priority	Number of	Agreeableness	Noise Level	Time Elapsed	Task Priority
Workload	Alarms	Number of	Number of	Extroversion	Workload
Alarm	Agreeableness	Alarms	Alarms		
Criticality	Shift	Openness			
Time Elapsed	Time Elapsed	Task Priority			
Noise Level	Workload				
Number of	Noise Level				
Alarms	Staff ID				
Staff ID	Extroversion				
	Alarm				
	Criticality				
	Openness				

Regression Using Combined Factors

Correlation Analyses for investigating the relationship between staff performances and combined factors (working conditions, staff individuality and alarm fatigue measures) were conducted and a correlation was found between apathy and conscientiousness (correlation coefficient: 0.904) for unit clerks and is attached in Appendix F as Table F3. These correlation coefficients are marked with black boxes in the correlation matrix. The nurse data does not show any correlation (Table F4 in Appendix F). In order to find a better relationship between alarm fatigue measures and performance measures, another relationship was tested. This was the relationship between the variables from alarm fatigue causing factors along with alarm fatigue measures and performance measures.

Tables 28, 29, and 30 show better models for the previous models summarized in Table 24, 25, and 26 respectively for both unit clerks and nurses.

Table 28: Summary of regression for Response (yes/no) using combined factors

Regression Models		Unit Clerk		Nurse	
Binary Logistic Regression	Real Data	After Transformation		Real Data	After Transformation
	$R^2 = 0.198$	$\lambda = 0.5$		$R^2 = 1.00$	$\lambda = -2.00$
	$RMSE = 0.062$	$R^2 = 0.394$		$RMSE = 0.309$	$R^2 = 1.00$
	Normality Test (Shapiro-Wilk): $\rho < 0.0001$	$RMSE = 0.029$ Normality Test: $\rho < 0.0001$		Normality Test: $\rho < 0.0001$	$RMSE = 0.261$ Normality Test: $\rho < 0.0001$
Multivariate Adaptive Splines (MARS)	$R^2 = 0.515$ $RMSE = 0.34925$		$R^2 = 0.993$ $RMSE = 0.0401$		
	Variables that have influence	Importance Scoring		Variables that have influence	Importance Scoring
	Alarm Type	100		Task Priority	100
	Workload	81.65		Alarm Type	24.16
	Staff ID	68.36		Number of Alarms	24.05
	Shift	68.31			
	SWAT	64.59			
	Boredom	59.61			
	Task Priority	56.90			
	Apathy	55.73			
	Time Elapsed	49.41			
	Agreeableness	46.66			
	Number of Alarms	42.09			
	Alarm Criticality	42.01			
	NASA TLX	31.33			
	Distrust	31.25			
	Extraversion	21.09			

Table 29: Summary of regression for Response Time using combined factors

Regression Models		Unit Clerk		Nurse	
Linear Regression	Real Data	After Transformation		Real Data	After Transformation
	$R^2 = 0.009$	$\lambda = 0.02$		$R^2 = -0.292$	$\lambda = -0.04$
	$RMSE = 8.51$	$R^2 = 0.014$		$RMSE = 8.37$	$R^2 = 0.012$
	Normality Test: $\rho < 0.0001$	$RMSE = 8.20$ Normality Test (Shapiro-Wilk): $\rho < 0.0001$		Normality Test: $\rho = 0.004$	$RMSE = 6.43$ Normality Test: $\rho = 0.0001$

Table 29: Summary of regression for Response Time using combined factors (continued)

Regression Models		Unit Clerk	Nurse
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.896$ $RMSE = 2.74$		$R^2 = 0.47$ $RMSE = 5.768$
	Important Variables	Importance scoring	Important Variables
	Task Priority	100	Number of Alarms
	SWAT	96.31	Neuroticism
	Alarm Type	47.33	
	Shift	38.83	
	Number of Alarms	37.69	
	Time Elapsed	32.27	
	Workload	31.75	
	Alarm Criticality	28.22	
	Neuroticism	17.48	
	Staff ID	17.17	
	Noise Level	16.63	
	Distrust	14.35	
	Boredom	12.25	
	NASA TLX	11.44	
	Openness	7.06	

Table 30: Summary of regression for Number of Ignored Alarms using combined factors

Regression Models		Unit Clerk	Nurse
Linear Regression	Real Data	After Transformation	$R^2 = 0.697$ $RMSE = 0.7331$, Normality Test: $p=0.215$
	$R^2 = 0.849$ $RMSE = 12.04$ Normality Test: $p<0.0001$	$\lambda = 2.00$ $R^2 = 0.98$ $RMSE = 2.56$ Normality Test: $p<0.0001$	Variables that have influence
			p value
			Time Elapsed* 0.0001
Multivariate Adaptive Regression Splines (MARS)	$R^2 = 0.999$ $RMSE = 0.821$		$R^2 = 0.96$ $RMSE = 0.2631$
	Important Variables	Importance Scoring	Important Variables
	Time Elapsed	100	Time Elapsed*
	Shift	31.73	Alarm Type
	SWAT	28.60	Apathy
	Number of Alarms	23.55	Noise Level
	Workload	12.12	Workload
	Distrust	10.93	NASA TLX
	NASA TLX	4.95	SWAT
			Openness

Principal Component

Analyses for Combined Factors: Principal Component Analyses with all the

factors including working conditions, staff individuality, and alarm fatigue measures, extracted six components for both type of participants. Linear regression models with those components also do not satisfy the linearity and normality assumptions except with the number of ignored alarms. The eigenvalue matrix along with the extracted components are shown in Table G6, Figure G3 attached in Appendix G.

Table 31: Summary of regression with components extracted from PCA combined factors

	Performance Measures	Statistics	Principal Components that have influence	—
Unit Clerk	Response	$R^2 = 0.077$ $RMSE = 0.062$ Normality Test: $\rho < 0.0001$	Violation of Linearity and Normality Assumption	—
	Response Time	$R^2 = 0.013$ $RMSE = 8.53$ Normality Test: $\rho < 0.000$	Violation of Linearity and Normality Assumption	—
	Number of ignored alarms	$R^2 = 0.714$ $RMSE = 16.54$ Normality Test: $\rho < 0.0001$	Violation of Normality Assumption	—
Nurse	Response	$R^2 = 0.302$ $RMSE = 0.309$ Normality Test: $\rho = 0.002$	Violation of Linearity and Normality Assumption	—
	Response Time	$R^2 = 0.0021$ $RMSE = 12.27$ Normality Test: $\rho < 0.0001$	Violation of Linearity and Normality Assumption	—
	Number of ignored alarms	$R^2 = 0.674$ $RMSE = 0.811$ Normality Test: $\rho = 0.156$	PC1, PC2	Apathy, Distrust, Workload, Time Elapsed, Boredom, Staff ID, Alarm Type, Alarm Criticality, Task Priority

Table 32 shows most of the entered variables as important variables for unit clerks, whereas there are much fewer for nurses. Moreover, the influence of personality type is also seen to a much lesser extent in nurses than in unit clerks.

Table 32: Summarization of all the influencing variables from regression with combined factors to relate alarm fatigue to staff performances

Unit Clerk			Nurse		
Response (Yes/No)	Response Time	Number of Ignored Alarms	Response (Yes/No)	Response Time	Number of Ignored Alarms
Alarm Type	Task Priority	Time Elapsed	Task Priority	Neuroticism	Time Elapsed
Number of Alarms	Number of Alarms	Number of Alarms	Number of Alarms	Number of Alarms	Apathy
Workload	SWAT	Shift	Alarm Type		Alarm Type
Staff ID	Alarm Type	SWAT			Noise Level
Shift	Shift	Workload			Workload
Boredom	Time Elapsed	Distrust			NASA TLX
Task Priority	Workload	NASA TLX			SWAT
Apathy	Alarm				Openness
Time Elapsed	Criticality				
Agreeableness	Neuroticism				
SWAT	Staff ID				
Alarm	Noise Level				
Criticality	Distrust				
NASA TLX	Boredom				
Distrust	NASA TLX				
Extraversion	Openness				
	Extraversion				

DISCUSSION

To comply with the mission to mitigate alarm-related incidents from all hospitals in the US by 2017, many hospitals are taking initiatives to assess their existing alarm monitoring systems and investigate the existence of alarm fatigue in their staff. The literature review revealed that researchers have focused on the working conditions for alarm monitoring in an attempt to tie it to staff performance. These studies then name alarm fatigue as the cause which deteriorates staff performance. In other words, researchers have been measuring staff performance and then relating any possible deterioration to alarm fatigue by determining this fatigue to be the cause. This approach can be summarized as an effect-and-cause method in which they first find the effect on performance and then blame it on alarm fatigue.

However, in order to assess alarm monitoring systems and to make to any necessary improvement in staff performance, it is essential to be able to actually measure the cause first and then see what effects it has on the performance measures. This study intends to define alarm fatigue in terms of working conditions and staff individuality and then relates alarm fatigue to staff performance to see whether the deterioration in staff performance is caused by alarm fatigue.

Figure 12 shows the approach the research considered in the literature review has taken to understand the role of alarm fatigue, as compared to the approach of this study. Straight arrows shows the relationships for this study and the curved arrows show the

relationship stated by several previous studies (Cvach, 2012; Pishori, 2012; Baillargeon, 2013). The following discussion explains the results from this study.

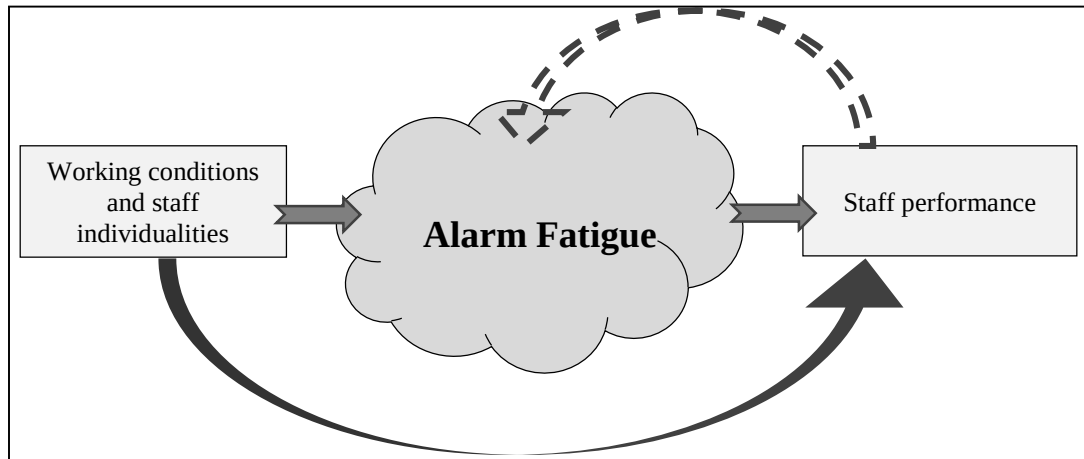


Figure 12: Approaches to understanding the role of alarm fatigue

→ Intension of current study

→ Findings of literature review research

Assessment of Existing Alarm Monitoring System

The HTF survey shows that the existing alarm monitoring system in the intensive care unit at Bozeman Deaconess Hospital produces a large number of false alarms which is one of the main causes initiating alarm fatigue in staff. The section which found some alarm-related information from ICU staff yielded some similarities and some dissimilarities between respondents. A majority of respondents (>84%) agreed or strongly agreed with the statements divulging the purpose of clinical alarms and the necessity for prioritized and easily-differentiated audible and visual alarms. Likewise, a large percentage of respondents confirmed the occurrence of frequent nuisance alarms (84.6%) which disrupt patient care (84.6%), and reduce trust in alarms causing caregivers

to disable alarms (53.8%, neutral: 46.2%). No participants disagreed to that last statement. Many of the staff agreed that smart alarms could help to minimize some of these nuisance alarms (84.6%) and that a central alarm management system could be useful (69.2%).

Responses were divergent on the statements about the complexity of hearing, recognizing alarms, and responding to alarms. Table 10 presented in the result section, shows that the percentages of those who agreed and disagreed on these questions were fairly similar. Furthermore, differing opinions were also observed on the questions concerning integrating alarm information with communications systems (e.g., pagers, cell phones), improving policies and procedures to regulate alarms properly and developing patient specific alarm settings. This divergence may indicate that the technology surrounding alarm monitoring systems requires training and experience to master.

The second section of the clinical alarm survey provided insights into how the staff rate the relative influences of various issues faced within the clinical alarm management system (see Figure 3 in result section). For most of the items, responses were well-distributed across the range of importance. Difficulty in identifying the source of an alarm, noise competition, and lack of training on the alarm system were found to be the most important of the surveyed issues from the respondents. Inadequate number of staff was rated as one of the top four issues. Difficulty in hearing alarms, over-reliance on alarms, and difficulty in prioritizing alarms were also found in the highly important

category by ICU staff. Difficulty in setting alarms and, contrary to the literature review findings, frequent false alarms were rated a medium important issues in this survey.

From the above description it is clear that nuisance alarms occur frequently and disrupt patient care. Although the participants agreed that those nuisance alarms reduce trust and cause the caregivers to turn them off, but surprisingly, this frequent occurrence of nuisance alarms was rated as least important issue by the staff. Instead, other concerns took priority in initiating disruptions in alarm management and causing alarm fatigue in the staff. Hearing, recognizing, and prioritizing alarms along with the lack of proper training in alarm management were discovered as the most important issues from this HTF survey on the existing alarm monitoring system. Establishment of a smart alarm system and better regulation of alarm monitoring policies were also seen as necessary, according to staff opinion.

Objective 1- Identification of all the Possible Influential Factors for Defining Alarm Fatigue

The literature review found several important factors which can influence alarm fatigue. These factors include shift (day/night), staff individuality along with personality type, workload (staff to patient ratio), number of alarms, time elapsed, alarm type, alarm criticality, and noise level as fatigue-causing factors, and response time as performance factor. Patient care involves several tasks other than alarm monitoring. These tasks could have influence on the monitoring task and on the alarm fatigue as well. Therefore, it was decided to perform hierarchical task analyses for the ICU staff. During task analyses, a

significant issue observed was the difference in the responsibilities of unit clerks and nurses. These two types of participants have completely different assignments toward patient safety and distributed responsibilities for responding to alarms. As a consequence, the decision was made to perform all the analyses separately, dividing the ICU staff into two types of participants: unit clerks and nurses. HTA was also conducted separately for both unit clerks and nurses in order to ensure the identification of all the possible factors that influence alarm fatigue.

Hierarchical Task Analysis for Unit Clerks

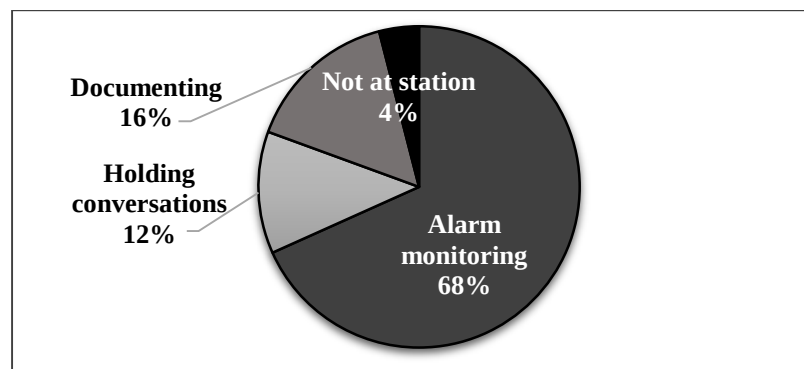


Figure 13: Primary tasks distribution for unit clerks in ICU during observed shifts

Unit clerks are responsible for responding to most of the alarms. The primary tasks of the unit clerk is to monitor two telemetry monitors from the central monitoring area: one monitor for the ICU patients and another monitor for the other surgical and medical floor patients. They also record and post rhythm strips for each patient every four hours, and answer phone calls. At the end of every shift, they hand over their responsibilities to the unit clerk for next shift. There are several important sub-tasks for

these primary tasks which are shown in the task hierarchy for unit clerks in Figure 4 (see result section).

This task analysis found that unit clerks have several tasks other than alarm monitoring. The hierarchy in Figure 4 exhibits that these tasks are equally important, although Figure 13 showing the primary tasks distribution of unit clerks for all the observed shifts reveals that alarm monitoring requires more time. When a unit clerk documents rhythm strips or answers a phone call and then has to leave the station to convey the information to a nurse, s/he cannot respond to alarms quickly. Answering phone calls is vital. Physicians' orders and concerns, patient family's queries and concerns, information about patient admission, discharge, or manipulation and other essential information come through phone calls. Similarly, documentation is necessary to have all the patient information available to physicians, nurses, and pharmacists for their best treatments. During documentation, printing strips, posting them in record book, and interpreting the strips require the clerks' complete attention. Likewise, during the hand-off process, the clerk from the previous shift explains all the essential critical information about each patient to the clerk for incoming shift. S/he has to describe every critical incident that has happened to each patient for the entire previous shift. All these situations cause the incidents of delayed responses and/or ignored alarms.

Hierarchical Task Analysis for Nurses

Nurses also need to take care of a significant amount of different tasks related to patient safety, besides alarm monitoring. Task distribution for nurses is shown in Figure

14. Usually for every two patients, one nurse is assigned. They miss alarms from one patient when they stay at the other patient's room or are engaged in other tasks, such as having conversations over the phone or directly with physicians, other nurses, pharmacists, unit clerks, or the patient's family. Sometimes they also need to attend meetings to make plans for patient care or the hand-off process. When a nurse attends one patient and gets a highly critical alarm from another patient, then s/he needs to prioritize tasks. Sometimes if there is another free nurse, s/he steps in and attends to the critical situation. In all these tasks, alarms can be easily ignored or response delayed. Moreover, difficulty in hearing alarms and recognizing the alarm (from which patient, what type and how critical) are also significant causes for ignored alarms, and/or delayed response.

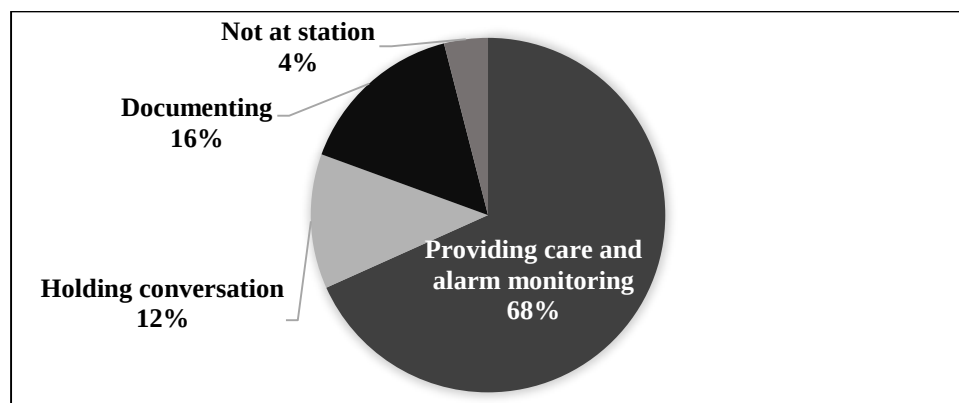


Figure 14: Primary tasks distribution for nurses at ICU during observed shifts

Therefore, task distribution is an important matter to study and task priority is an important factor to include in this study. Prioritizing tasks requires unit clerks and nurses to sometimes respond in delay or even ignore other alarms, which creates dilemma, increases mental workload and causes fatigue. Thus, along with response time, response

(Yes/No) to alarms and number of ignored alarms should be included for this observational study as performance variables. This task analysis is shown in Figure 5 in the result section

Descriptive Statistics for Alarm Fatigue Measures and Performance Measures

For the descriptive statistics the alarm fatigue measures were SWAT, NASA TLX, Boredom, Apathy, and Distrust and the performance measures were response time and number of ignored alarms. Response (yes/no) was not included in the statistics for performance measures because it was not significantly different than the number of ignored alarms. The descriptive statistics exhibit that all the alarm fatigue measures and performance measures for this research were higher during night shifts for unit clerks. Similarly, for nurses, two of the affect (boredom, and distrust) along with the performance measures were found to be higher during night shifts. The fact that most of these alarm fatigue measures have higher values and staff performance deteriorations in terms of longer response times and a higher number of ignored alarms during night shifts can be the reason for occurring higher percentage of nuisance alarms at night (night shift: 86% and day shift: 74.83%). During night shifts while the patients rest, there is less patient movements; nevertheless, there are still a large number of nuisance alarms, creating fatigue in staff in terms of mental workload and affect. The nighttime nuisance alarms are due to standard threshold parameter settings, which are not patient specific, but too risky to manipulate at night.

In contrast to unit clerks, nurses were found to have higher values on mental workload measures during day shifts. Generally, patient care is provided by physicians and nurses from respiratory therapy, physiotherapy, and other necessary laboratory testing departments during day shifts. The ICU nurses are responsible to support those physicians and nurses from other departments. During patient manipulation, a large number of nuisance alarms occur, making the environment noisy. Nurses are aware of the (lack of) criticality of these alarms as they know that these are created from patient movements. According to patient safety regulations, alarms cannot be silenced or turned off during patient manipulation; however, the HTF survey indicated that these policies are not followed. Furthermore, nurses are often called to help with the manipulation of patients other than their own, increasing their own workload. If an alarm from one of their own patients should sound during this time, they are called upon to prioritize tasks. Thus, the data showing nurses as being highly mentally loaded and indifferent towards alarms during day shifts, even more than unit clerks, may be attributed to noise level, workload (staff to patient ratio), task priority, and other possible influencing factors from patient care responsibilities.

For both types of participants, measures for apathy show very large values, indicating a high level of indifference, i.e., less motivation in responding to alarms. Since the apathy evaluation scale was used to get answers from staff at the end of the observed shifts, these high values support the fact that staff become desensitized and fatigued. Comparing the two types of participants, the percentages of ignored alarms and response

times to alarms are always (day/night) higher for unit clerks as they are responsible for monitoring and responding to most of the alarms.

Figure 6 from result section reveals the positive relationships between mental workload (in terms of SWAT), affect (in terms of distrust), and staff performance (in terms of percentage of ignored alarms) for unit clerks. The measuring of these three variables were time-dependent and occurred frequently during a shift. The resulting graphs explain that mental workload and affect have influence on unit clerk performance during alarm monitoring for most of the shift times with two exceptions, on night 1 and day 4. These exceptions are not significant enough to make any determination. During night 1, the SWAT measures decreased from 39.47 to 33.33, distrust level remained unchanged, but the percentage of ignored alarms increased from 64.86 to 67.86. This situation may be the result of the increase in number of patients over that shift (admission of two new patients in the ICU). The set-up process for a new patient causes a large number of nuisance alarms, so the unit clerks and nurses know they do not have to be seriously concerned about those alarms. As a result, however, they become desensitized to those alarms. In the same way on day 4, there was sudden increase in the number of patients (15 to 21, two were admitted into the ICU and the remaining four were admitted onto other medical floors), thus again creating a large number of nuisance alarms. In the case of the ICU patients, the unit clerks were aware of having new patients and related the nuisance alarms to them, but in the case of the other floor patients, information was not always prompt enough in coming to make the unit clerks aware of the source. As a

result, they became overwhelmed by so many alarms. The more frequent the alarms, the greater the chance of ignoring them. The clerks were not distrusting the alarms, but they became mentally overloaded in responding to each and every alarm.

Objective 2- Definition of Alarm
Fatigue in terms of Mental Workload

The definition of alarm fatigue started with a correlation analysis. The research identified a long list of possible influencing factors for alarm fatigue. Therefore, the correlation analyses were performed for both unit clerks and nurses in order to reduce the list of variables for having correlation in the pairs of factors. These correlation analyses with alarm fatigue causing factors gave evidence that none of the considered variables should be removed from further analyses (See Table F1 in Appendix F); they all played a role and needed to be considered in defining alarm fatigue for each of the participants. Linear regression analyses investigating the significant factors in defining alarm fatigue show acceptable models for nurses but not for unit clerks, even after transformations (using Box-Cox and Johnston transformations) for all five alarm fatigue measures. The graphs for homogeneity assumptions do not show any specific trend or funneling affect, which satisfies the fulfillment of this assumption for the linear regressions for nurses. Normality test results in terms of p values for those regressions, along with the histograms, satisfy normality assumptions as well.

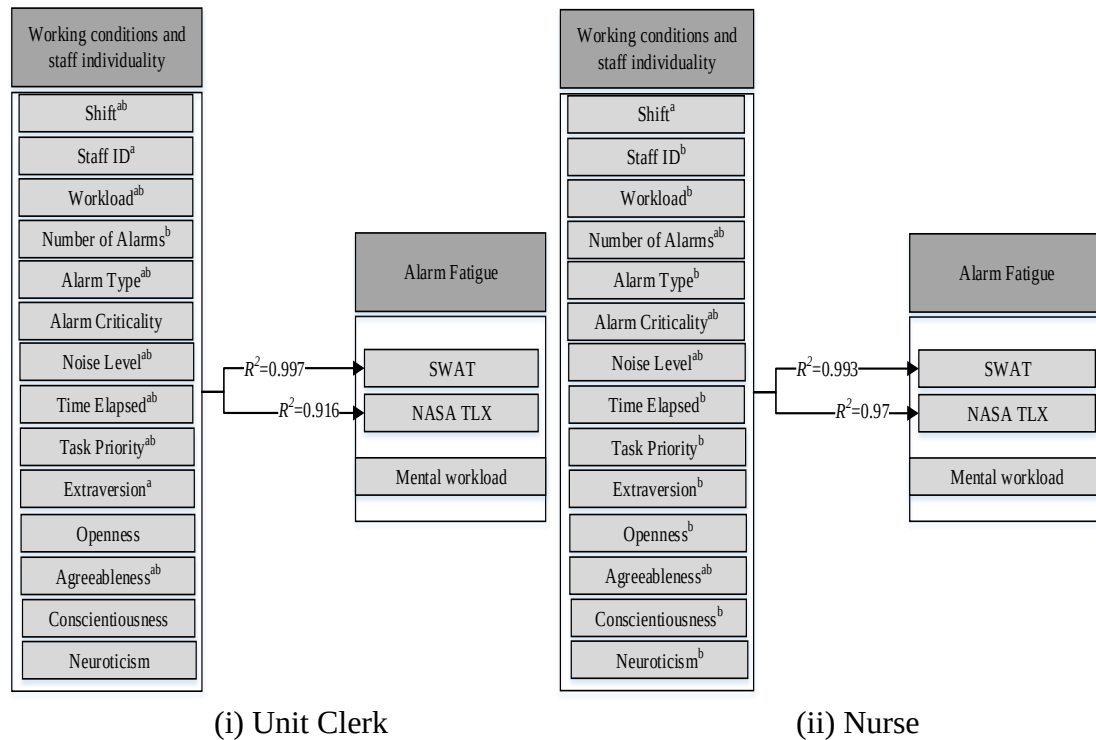


Figure 15: Association of fatigue causing factors with mental workload measures showing R^2 values inside arrows for both groups of participants
a Indicates the influential factor for SWAT
b Indicates the influential factor for NASA TLX

Non-parametric, non-linear regression analyses using the Multivariate Adaptive Regression Splines (MARS) software were conducted for both unit clerks and nurses to see if they would give improved results. MARS produced good models pertaining to unit clerks for all the alarm fatigue measures. Additionally, MARS generated better models for mental workload measures and distrust in terms of lower *RMSE* and higher R^2 values than the linear regression models, for nurses. Figure 15 shows the influencing variables from each model along with their R^2 values for both groups. The results from linear regression and MARS had some differences in outputs for the investigation of significant

factors. MARS produced better models with better *RMSE* and R^2 values, which therefore yielded more accurate results. As this study is willing to explore all the possible alarm fatigue causing factors in defining alarm fatigue, all the variables found from both types of analyses were considered in the definition.

The factors that influence mental workload measures include mostly working conditions related factors rather than staff personality dimensions in the case of definitions pertaining to unit clerks. On the other hand, for nurses, the significant variables show a balance between the working condition and staff individuality factors which influence mental workload measures. In addition, the two mental workload measures show very close similarities in terms of significant variables for defining alarm fatigue for both types of participants. This validates defining alarm fatigue in terms of mental workload through both NASA TLX and SWAT. Currently, it is the responsibility of the unit clerks to constantly monitor alarms from two telemetry monitors. As a result, they are the people who are actually exposed to the stress from continuous flashing, noise, and necessity of response. Their personality types do not cause any difference in their mental workload. They have to respond to the alarms as a part of their duty and must distinguish the alarms and protect themselves from noise and visual stress at the same time; they have no actual patient contact, however. On the other hand, nurses stay with the patients almost all of the time. They are responsible for only a few types of alarms, of which only IV pumps and ventilator alarms are frequent. These alarms are highly dependent on patient criticality and nurses may take different views in responding

to those alarms according to patient criticality, task priority and experience. Therefore, for nurses, staff individuality has an apparent influence on their mental workload.

Objective 3- Definition of Alarm Fatigue in terms of Affect

For affect, similar to mental workload measures, working condition related factors show a greater influence than personality type dimensions on unit clerks. The time elapsed and the number of patients or the number of alarms in a shift are factors which increase the distrust level toward alarms for unit clerks. Only the personality dimension agreeableness was found to be significant for distrust in unit clerks. The more suspicious (not agreeable) the person is, the more rapidly distrust level for alarms increases. On the other hand, for other two types of affect considered for this study, the unit clerks showed the influence of personality type dimensions and shift (day/night). Motivation or conscientiousness in responding to alarms is strongly dependent on personality type. The time of the shift influences the degree of indifference; during day shifts, there are several nurses and other people with whom unit clerks can talk and share their emotions, but during night shifts they need to keep the silence for patient convenience and become bored and indifferent with the monotonous activity.

For nurses, similar balanced factors between working conditions and staff individuality were found to be significant for boredom and apathy. Most of the time, nurses need to stay near the patients' rooms and perform the same kind of duties again and again to ensure patient safety. The monotony of the job along with the individual

personality of the nurse create boredom and eventually indifference toward their task of responding to alarms. For distrust, nurses have similar significant factors as unit clerks which can be explained in a similar way. Figure 16 shows the influencing variables from each model along with their R^2 values for both groups.

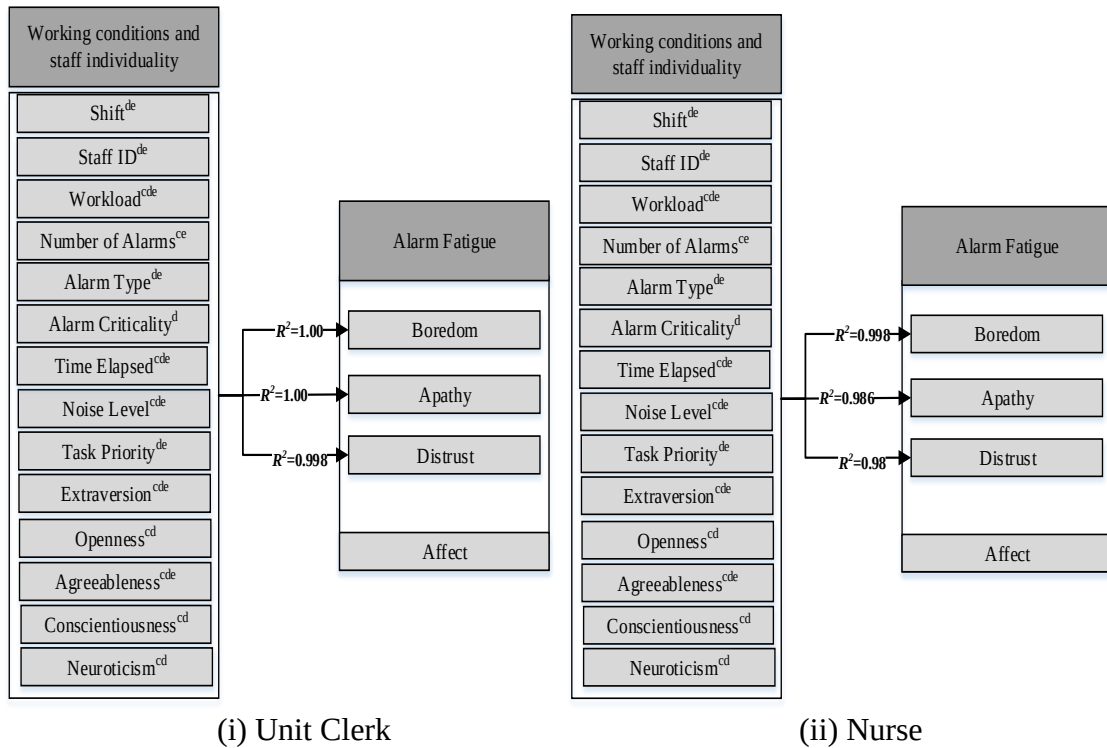


Figure 16: Association of alarm causing factors with affect showing R^2 values inside arrows for both groups of participants

c Indicates the influential factor for Boredom

d Indicates the influential factor for Apathy

e Indicates the influential factor for Distrust

Objective 4- Comparison between Two Definitions

The comparison between the two definitions (derived from mental workload and affect) for unit clerks do not show similarities. As described earlier, mental workload

measures are influenced by almost all the factors considered as working conditions except alarm criticality. Staff ID along with two personality type dimensions (extraversion and agreeableness) were found to have influence on mental workload. On the other hand, affect is influenced by most of the positive dimensions of personality type along with shift, number of alarms, workload and time elapsed. Therefore, from this description it can be concluded that the two definitions of alarm fatigue pertaining to unit clerks are not analogous.

On the other hand, for nurses, both mental workload and affect definitions include all the possible influencing factors considered for this study and yield the same definitions. Alarm fatigue in nurses is dependent not only on the alarms they are responsible for in the patient rooms, but also on the alarms coming from the two telemetry monitors in the central monitoring area. They are not responsible for the alarms coming from the central monitoring area, but these do increase the overall noise level for the nurses. Furthermore, they have so much to do throughout the whole shift, especially during day shifts, that their mental workload and affect measures are influenced by all of the possible factors. The factors that affect mental workload and three types of affect are described below. Graphs showing individual associations of these factors with alarm fatigue measures are attached in Appendix H. Table 33 shows the type of associations (positive/negative) of alarm fatigue measures with the continuous factors in working conditions and staff individualities for both types of participants.

Table 33: summary of the significances of influencing factors on the alarm fatigue measures

Working conditions and staff individualities	Unit Clerk				
	SWAT	NASA TLX	Boredom	Apathy	Distrust
Workload (Staff:Patient)	-	-			0
Number of Alarms		+			+
Time Elapsed	+	+			-
Noise Level	-	-			
Extraversion		+	+		
Openness			+		
Agreeableness	0	+			+
Conscientiousness				-	
Neuroticism					
Working conditions and staff individualities	Nurse				
	SWAT	NASA TLX	Boredom	Apathy	Distrust
Workload (Staff:Patient)		0	+	+	+
Number of Alarms	+	+	-		-
Time Elapsed		-	0	-	0
Noise Level	+	+	-	+	+
Extraversion		+	+	+	-
Openness		+	+	+	
Agreeableness	-	0	+	+	-
Conscientiousness		-	-	0	
Neuroticism		-	-	-	

+ Indicates positive association

- Indicates negative association

0 Indicates no change

Significance of Shift

Shift has influence on mental workload and affect for both types of participants (see Figure H1 in Appendix H). From task analyses, it was found that the responsibilities for nurses vary a lot between day and night shifts. During the daytime, the nurses need to attend an interdisciplinary meeting to discuss patient conditions and to make plans for patient specific care. They also need to provide support to the physicians and nurses from

respiratory therapy, physiotherapy and testing labs for their assigned patients along with their daily scheduled duties. The working environment also has different scenarios for days and nights. During day shifts, the facility remains noisy due to all the people on the floor, including nurses from other departments, physicians, pharmacists and visitors for the patients. The alarms need to be kept at a higher volume so as to make them audible above the crowd. Unit clerks are less affected by the higher noise level during the day, and additionally, they benefit from the extra support from other caregivers. On the other hand, nights are quieter, but full of risks, without the quick support from physicians and other medical caregivers. Also, if there is any breakdown in the monitoring system, technical support is not available at night. Fewer people and a monotonous job of concentrating on monitoring alarms make both nurses and unit clerks bored and tired. In addition, while unit clerks have the responsibility of responding to alarms, nurses have very little responsibility. At the end of the shift, both the participants become indifferent and mentally fatigued by this critical task of monitoring patient conditions.

Significance of Staff ID

Individuals have varying perceptions of events depending on their education, experience and personality type. Staff ID was found to have influence on mental workload for both unit clerks and nurses, but not affect for nurses only. Unit clerks have the responsibility of monitoring and responding to the alarms coming into two telemetry monitors in front of them. Even if they ignore the nuisance alarms, the environment gets noisy. Therefore, for unit clerks, although their mental workload is influenced by their

individuality, they do not feel distrust, bored or less motivated to respond to the alarms because responding is their duty. They need to acknowledge the alarms to make the facility less noisy and favorable for critical patient care, although they do not need to make an active response for most of those alarms. On the other hand, nurses need to take care of a lesser number of alarms from the patient rooms only. Their responses are highly based on their insights and experiences and their status (busy/ idle). Most of the time, nurses remain extremely busy with their scheduled tasks along with patient specific requirements. Eventually at some stage, they experience burnout and experience distrust or become highly indifferent in responding to alarms. The experience of this burnout phase varies from person to person. Therefore, for nurses, the influence of staff ID was found in the case of both definitions.

Significance of Workload

Workload had significance on both mental workload and affect for both types of participants (see Figures H3 and H11 in Appendix H). Unit clerks reported decreased mental workload with decreased workload (staff to patient ratio), which is obvious. The interesting part is the graph shows unchanged distrust level with the change in workload. Therefore, it can be said that workload individually did not have any significance on distrust. Nurses show similar results for mental workload. With the decrease in workload, a slight decrease in NASA TLX measure was found. On the other hand, for affect, all the measures are negatively influenced by workload. When workload decreases, nurses don't have enough work to do without passing idle time. This situation makes them bored and

eventually apathetic which is the state of being indifferent about responding to alarms. This state creates distrust in nurses as well which restrain them from responding to alarms.

Significance of Number of Alarms

The number of alarms is a factor that increases with time and the graphs attached in Appendix H (Figures H2 and H12) show some interesting results for the significance of this factor. Both types of participants show a positive influence of the number of alarms on their mental workload measures. With the increase in the number of alarms, the mental workload for both nurses and unit clerks increases, which is expected. However, the more alarms they respond to, the more they experience fatigue in doing this monotonous job while taking care of other scheduled tasks at the same time. The number of alarms has influence on the affect for both unit clerks and; unit clerks show its significance on distrust only and nurses show its significance on boredom and distrust. Interestingly, for nurses this factor has a negative influence on distrust and boredom, whereas for unit clerks it shows a positive influence on distrust. Generally, unit clerks are responsible to monitor most of the alarms and they often experience a lots of nuisance alarms, false alarms that go off frequently. Therefore, as the number of those alarms increases, their distrust level increases. There was no connection found for boredom or apathy regarding unit clerks as they are bound to respond to the alarms in front of them to make the environment less noisy and to make the following alarms distinguishable. On the other hand, usually nurses experience fewer number of alarms as they spend most of

the time nearer to the patient rooms and are responsible for monitoring only a few alarms, most of which are true. When they experience a larger number of alarms, they usually believe them and take the necessary steps to mitigate the alarms. Hearing alarms is not as monotonous for nurses as it is for unit clerks. Therefore, an increase in the number of alarms does not raise the distrust or boredom level for nurses.

Significance of Alarm Type

Alarm type is another factor that had influence on mental workload for both unit clerks and nurses. On the other hand, alarm type does not have influence on affect for unit clerks, and only on distrust and apathy (not boredom) for nurses. Usually, unit clerks need to take care of a large variety of alarms (up to fourteen kinds on each telemetry monitor). Some of these alarms, such as leads-off, battery low or no signal along with difficult-to-read EKG signals, generally go off so frequently that the unit clerks become fatigued and turn them off or silence them. In addition, unit clerks only convey abnormal information about any patient to nurses and document this, but they never take any active action themselves regarding patient care. In the case of nurses, they experience a very low number alarms, and most of the time, they try to respond to them with active measures to stop the alarm. The constant attention of nurses to acknowledge alarms and to mitigate their frequent occurrence causes fatigue in them mentally and emotionally. It was found during the observations that at the end of the shifts nurses no longer want to respond to those alarms and only do their scheduled tasks.

Significance of Alarm Criticality

Alarm criticality was found to have influence on both types of definitions only for nurses. For unit clerks, most of the alarms they experience are nuisance alarms. Most of the alarms they have to respond to are not highly critical and the only critical situations they experience come from phone calls. Therefore, criticality of alarms does not make them fatigued mentally or emotionally. On the other hand, nurses experience a very low number of alarms, critical or not, in their patient rooms and they try to respond to them actively to mitigate their frequent occurrence. But their response is determined by their task prioritization which is dependent on the alarm criticality. For example, for a call light or highly critical ventilator or IV pump alarm, they respond immediately, making these the priority. When the number of such alarms increases, task prioritization becomes more difficult, creating dilemmas which increase mental and emotional fatigue.

Significance of Noise Level

Both unit clerks and nurses show the significance of noise level on mental workload; however, only nurses show the influence of noise level on affect (See Figures H4 and H13 in Appendix H). As would be expected, an increase in noise level increases the mental workload for nurses, but surprisingly, an increase in noise level decreases mental fatigue for unit clerks. Nurses have the responsibility of patient care and sequentially scheduled tasks. They often have extra workload from other unscheduled tasks. Therefore, noise disrupts their thoughts and work plan and makes them mentally fatigued and emotionally affected. As a result, the increase in their distrust and apathy

levels can be attributed to a higher noise level. Nevertheless, increase in noise level decreases the boredom level for nurses. Boredom is the result of a monotonous job or having nothing to do. But a high noise level makes the nurses conscious of approaching danger in patient conditions and limits their boredom. Unit clerks, on the other hand, are habituated to the noisy environment as their main responsibility is to hear and recognize what type of alarm is on which monitor. Therefore, a lower noise level from a lower number of alarms makes them uncomfortable and feeling at risk of missing critical alarms. Furthermore, night shifts are overall less noisy and were found to have less mental workload for nurses and more for unit clerks, which also supports the significance of noise level on both types of participants.

Significance of Time Elapsed

Time elapsed was found to have its significance on two definitions pertaining to unit clerks and nurses (see Figures H5 and H14 in Appendix H). Both SWAT and NASA TLX measures for unit clerks showed the influence of time elapsed on their mental workload. Graphs (attached in Appendix H) show that as the time passes, mental workload for unit clerks increases. On the other hand, for nurses, NASA TLX measure shows that time elapsed has negative influence on mental workload. With time the mental workload of nurses decreases. This result was quite surprising but not difficult to explain. In the beginning of every shift, nurses need to attend two meetings; one with the charge nurse from previous shift to get information about patients' conditions admitted in the ICU at that time, and the other with the nurse from previous shift who was assigned for

the same patient the new shift's nurse would be assigned for. During day shifts, they also need to attend an interdisciplinary meeting to create plans for patient care at ICU. At the beginning of the shift, nurses remain careful about criticality of patients. As the time goes, they become habituated with the patient conditions and requirements. In addition, patient care from other departments are conducted during the beginning of the shifts, which adds some extra workload on nurses. Therefore, nurses' results showed decrease in the mental workload with time. For alarm fatigue definition with affect, time elapsed had significance on distrust only. Likewise, the mental fatigue definitions for unit clerks and nurses, these affect definition was also interesting. Although unit clerks showed increase in mental workload with time, for distrust they showed opposite significance. In order to reduce the mental workload, unit clerks try to get adapted to the system and make them comfortable by setting patient specific threshold parameters which reduces nuisance alarms and reduces their distrust level as well. On the other hand, nurses showed increased distrust level with time. Decrease in the mental workload decreases with time, make nurses indifferent about alarms. As a consequence, the distrust towards alarms increases for nurses.

Significance of Task Priority

Task priority was found to have significances on the mental workload for both types of participants and on affect for nurses only. Task prioritization is a very important factor for both nurses and unit clerks. Task analyses showed that both participants have other equally important tasks to complete other than only alarm monitoring. While doing

those tasks, if there is any alarm, then they need to prioritize their tasks and determine which tasks they have to take care of. This situation create mental stress and make them mentally fatigued. Unit clerks did not show any influence of task priority on the affect. Unit clerks do not need to move from their place as often. They do almost all of their work sitting on the central monitoring area. The task distribution in Figure 15 also support this previous statement showing only 4% allocation for 'Not at station' phase. Their sedentary job facilitate them in prioritizing their tasks and create no emotional stress on them. In the case of nurses, they need to pass very busy time by moving all around the ICU to carry out their duties. These staff experience difficulty in prioritize their tasks. Therefore, at a stage they become highly indifferent towards alarms and do not want to trust them anymore.

Significance of Extraversion

Extraversion was significant on all the definitions for both unit clerks and nurses which are shown in Figures H6 and H15 in Appendix H. Staff having a higher percentile for extraversion are more assertive and energetic in nature. They always try to communicate with others and show their energy in their tasks. As alarm monitoring is a task demanding a high level of concentration and steadiness, this kind of person gets fatigued more quickly, mentally and emotionally. On the other hand, these people always try to be enthusiastic about their tasks and therefore their distrust towards alarms was found to be low.

Significance of Openness

Openness does not have any significance on the mental workload for unit clerks. This factor has influence on the other three definitions; on the mental workload for nurses and on the affect for both participants. Openness is sometime renamed as intellect. Person with higher percentile of openness are highly intellectual and imaginative. They prefer to do experiments with activities. This kind of person chooses to take risks which can create mental and emotional loads on them. Caregivers, having this personality type, are highly risky for patients. Figures are attached in appendix H as Figures H7 and H16.

Significance of Agreeableness

Agreeableness is the measure of person's trusting and helpful natures. This factor had shown its significance on all the definitions (see Appendix H for Figures H10 and H19). For mental workload and affect, most of the relationships with agreeableness were found to be positive except with the SWAT and distrust measure for nurses. The higher the percentile for a person for agreeableness, the more s/he will be prone to mental or emotional fatigue. It can also be said that, the more the unit clerks are cooperative or helpful, they will try to convey information to nurses frequently and will be fatigued within a short time. In the case of nurses, a different scenario occurs. SWAT and distrust measures for nurses showed that the higher the percentile for agreeableness will be, the less the person will be mentally fatigued and experience less distrust towards alarms. As the nurses remain extremely busy for most of the time, they need support from other co-workers. People with agreeableness trust other people, help them and expect support from

them. Therefore, they do not get mentally fatigued easily and do not feel distrust towards alarms.

Significance of Conscientiousness

Conscientiousness means the tendency to be organized and dependable. Nurses showed the influence of this factor on mental workload and were found to have negative significance (see Figure H8 in Appendix H). Conscientiousness also was found to have negative significances on affect for both types of participants (see Figure H17 in Appendix H). Therefore, it can be explained that the more organized or dependable the person is, the less mental workload or affect s/he experiences.

Significance of Neuroticism

Neuroticism refers to the degree of emotional stability. The lower the percentile for neuroticism, the more the person is emotionally stable. Surprisingly, graphs show that nurses have negative trend for neuroticism with NASA TLX, boredom, and apathy measures. That means, an emotionally unstable person has less chance of getting fatigued. The staff who experience anxiety, easily try to be caring about all the possible dangers from the beginning of the shifts. Their experience and preventive stands make them less fatigued. Figures are attached in Appendix H as Figures H9 and H18 support this above discussion.

Objective 5- Significance of Alarm Fatigue on Staff Performance

The first approach of this study was to define alarm fatigue in terms of mental workload and affect. Alarm fatigue definitions and individual significance of influencing factors showed the associations of alarm fatigue with working condition and staff individuality measures. The second approach of this study attempted to observe the significance of alarm fatigue measures on staff performance. The following sections discuss the attempts that had taken to find the cause of staff performance deterioration.

Unsuccessful Attempts to Relate Alarm Fatigue to Staff Performance

Based on the conceptual model of this study, several types of analysis were performed in order to determine if alarm fatigue measures deteriorate alarm monitoring performance. Attempts to shorten the list of variables through correlation analyses was unsuccessful for alarm fatigue measures (see Table F2 in Appendix F). In addition, regression analyses yielded no significant relationship between alarm fatigue measures and the performance measures of response (yes/no), response time or number of ignored alarms (see Tables 20-22 in result section) for any of the participants except for the number of ignored alarms for nurses. Principal Component Analyses were also conducted with the alarm fatigue measures and combined measures. Table 23 and 32 again show violations in linearity and normality assumptions for linear regressions with components extracted from PCA (see Table G4 and G6 in Appendix G).

Once it had been determined that alarm fatigue measures are not the sole cause of staff performance deterioration, this study conducted further analyses on staff performance measures to create successful models. Two approaches were taken: one, which considered only working condition and individuality measures, and a second which combined alarm fatigue measures with working condition and individuality measures. Non-parametric regression analyses considering working conditions and staff individuality as independent variables for unit clerks yielded unacceptable model for response (Yes/No). These unsuccessful models are shown in Figure 17 and Figure 18 with pertinent R^2 values.

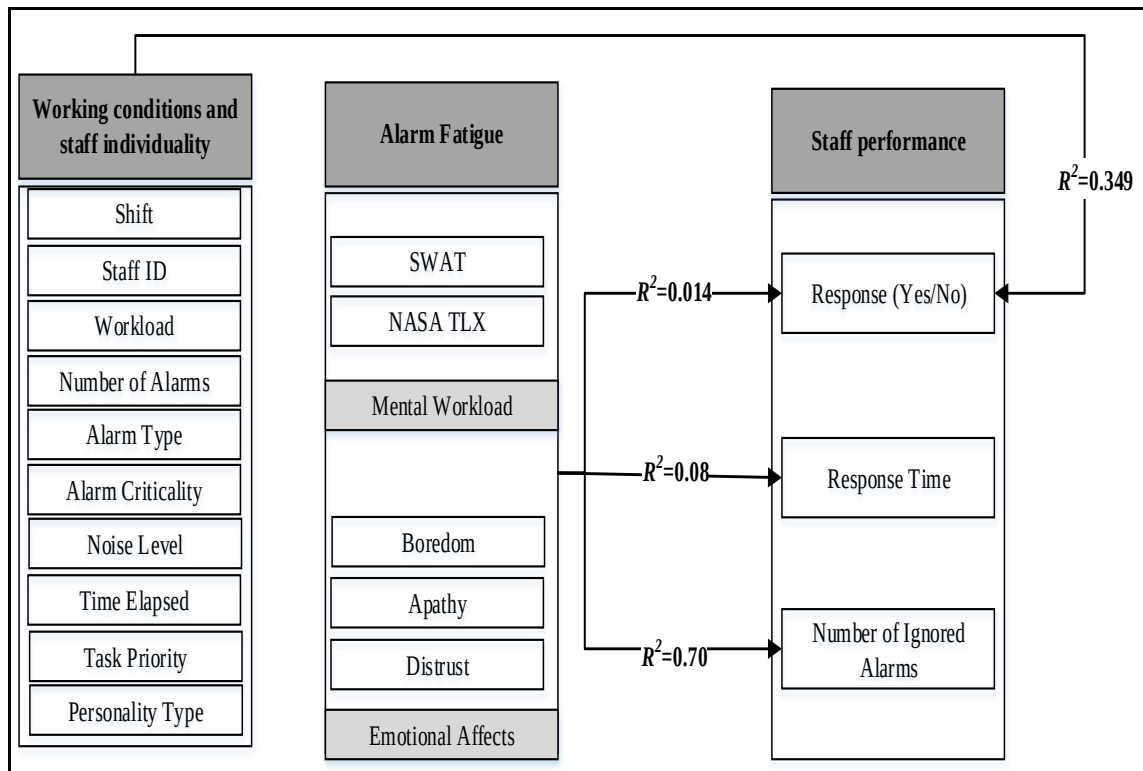


Figure 17: Unsuccessful models to associate alarm fatigue with staff performance for unit clerks

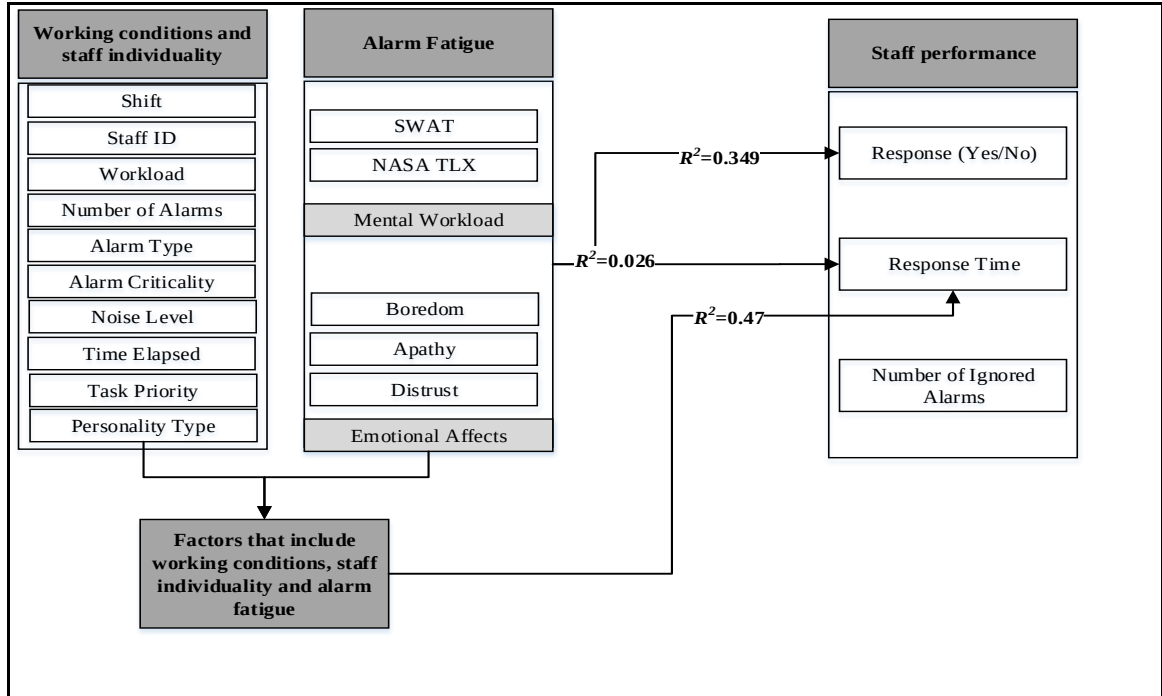


Figure 18: Unsuccessful models to associate alarm fatigue with staff performance for nurses

Successful Models which Define Factors Related to Staff Performance

In one respect, attempts to shorten the list of variables through correlation analyses was successful; in the case of combined factors and unit clerks, conscientiousness had correlation with apathy and therefore could be removed from the list of those combined variables (See Table F3 in Appendix F). With that reduced list of variables, non-parametric regression analyses using MARS on working conditions and individuality for unit clerks yielded acceptable results for response time and number of ignored alarms. The same analyses on a combination of alarm fatigue measures and working condition and individuality measures yielded successful models for all three performance measures. Models are shown in Figure 19 with R^2 values.

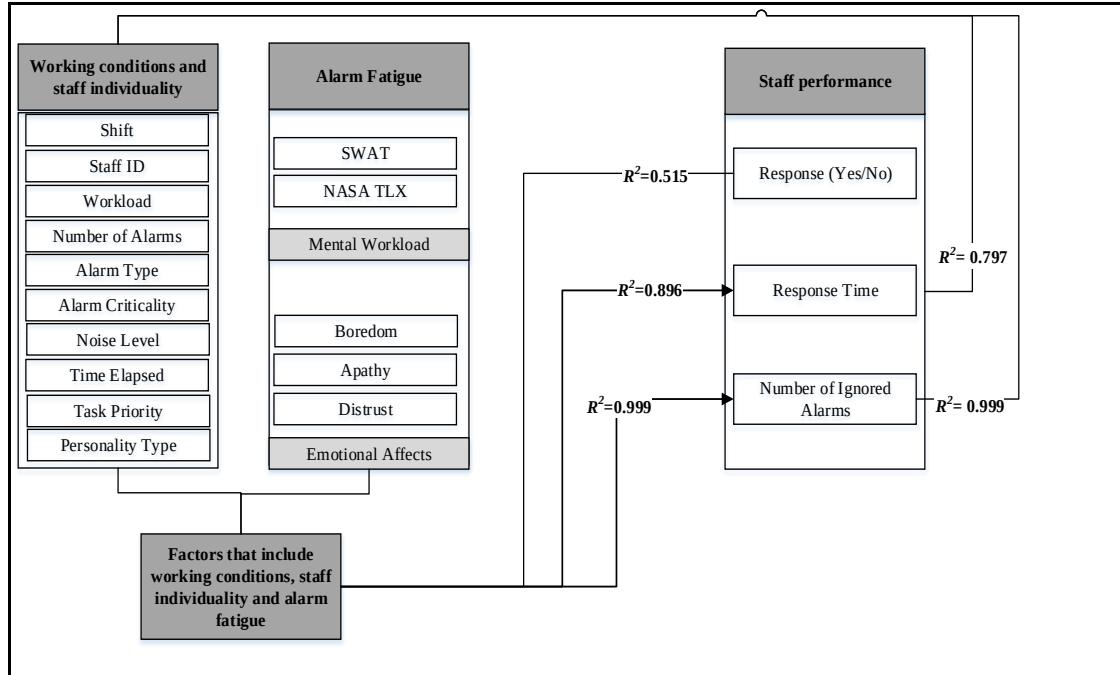


Figure 19: Successful models to associate alarm fatigue with staff performance for unit clerks

For nurses, non-parametric regression analyses on working condition and individuality measures yielded good results for all three performance measures. The same analyses were conducted using the combined alarm fatigue measures and working condition and individuality measures. This study found that using a combination of alarm fatigue measures and working condition and individuality measures produces a useful model for determining response (yes/no) and number of ignored alarms for nurses, but that a model considering only working conditions and individuality should be used for response time. This can be explained in that it is the nurses' duty to give care to patients, whether they are fatigued or not. Since alarms coming from patients' room can

make the patients tense and uncomfortable, nurses try to respond to all the alarms in the patient rooms as quickly as they can. Figure 20 show the successful models for nurses.

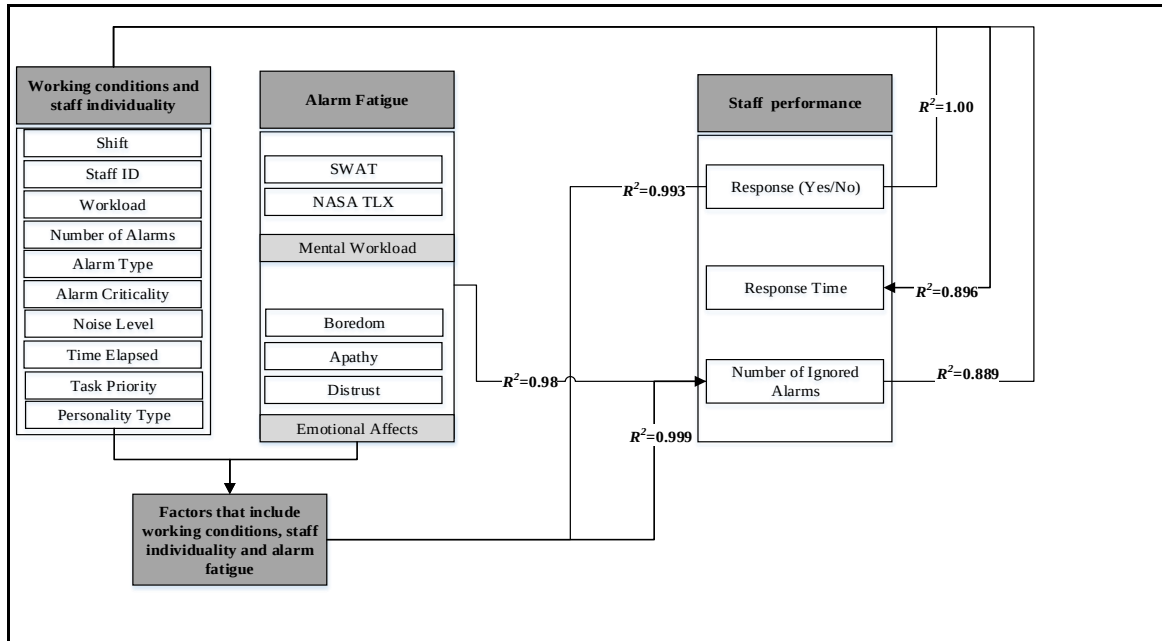


Figure 20: Successful models to associate alarm fatigue with staff performance for nurses

Until this study's attempt to define and measure alarm fatigue in terms of all influencing factors, most studies measured alarm fatigue as response time. However, response time is a performance measure, an influence of alarm fatigue, not alarm fatigue itself. In addition, previous studies did not differentiate between nurses and unit clerks. One of the most important findings of this is the fact that delayed response time for nurses does not seem to be caused by alarm fatigue, but rather by the working conditions and staff individuality. This is important for two reasons: 1) until now researchers have blamed delayed response time on alarm fatigue, which as shown in this study cannot be

claimed, and 2) because of this false assumption, what has been measured is, in fact, not alarm fatigue.

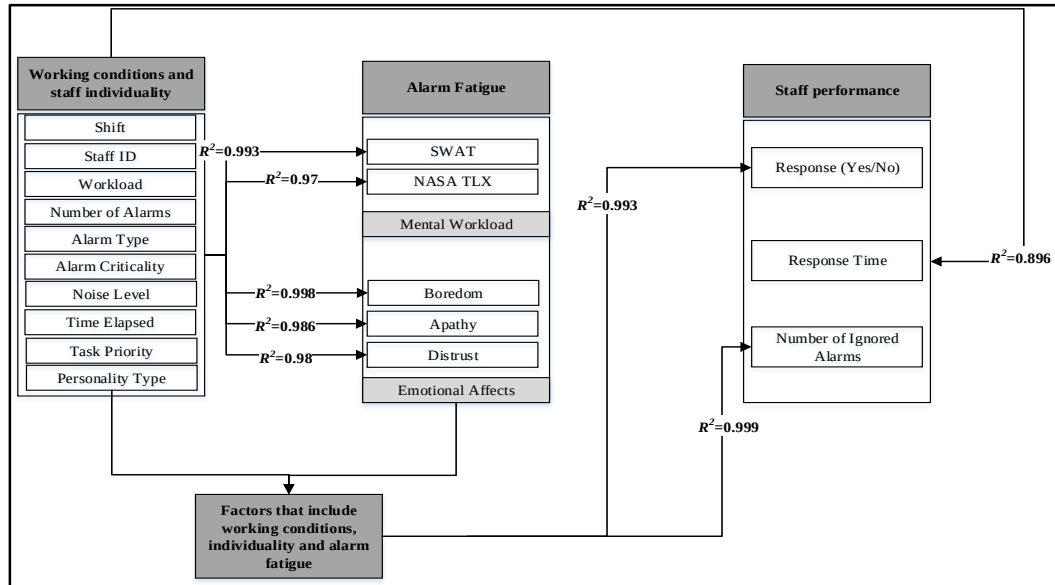


Figure 21: Proposed models on the association of alarm fatigue to staff performance for unit clerks

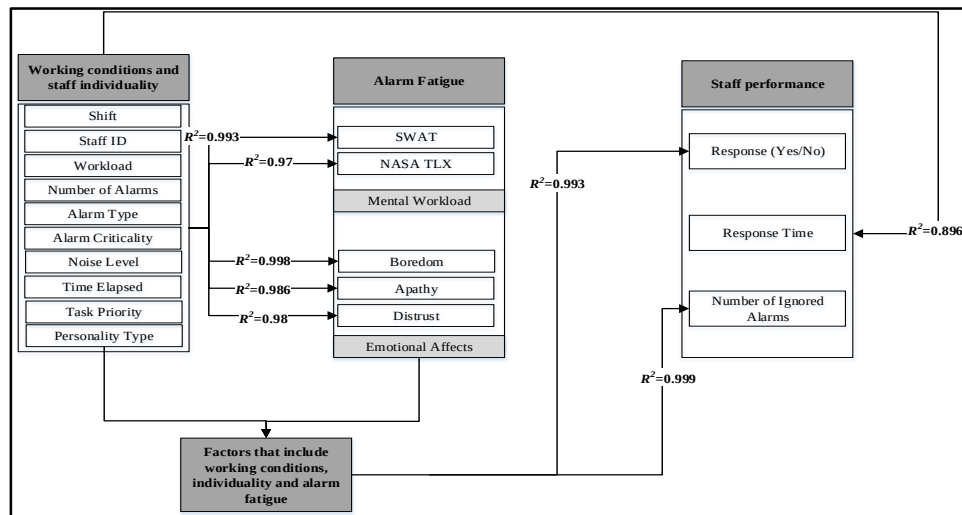


Figure 22: Proposed models on the association of alarm fatigue to staff performance for nurses

In order to find the best association of alarm fatigue measures with staff performance deterioration, this study proposes two successful models, one for unit clerks and one for nurses (see Figure 21 and Figure 22). Table 34 summarizes all the associations of different factors on staff performances (response time and number of ignored alarms). The association of those influencing factors are also explained.

Table 34: summary of the significances of influencing factors on staff performance

Combined factors of working conditions, staff individualities and alarm fatigue measures	Unit Clerk		Nurse	
	Response Time	Number of Ignored alarms	Response Time	Number of Ignored Alarms
Workload	-	-		-
Number of Alarms	+	+		
Time Elapsed	-	+	-	+
Noise Level	-		0	-
Extraversion	+		+	
Openness	+			0
Agreeableness			+	
Conscientiousness				
Neuroticism	-		-	
SWAT	+	+		-
NASA TLX	+	+		-
Boredom	0			
Apathy				+
Distrust	+	+		

+ Indicates positive association

- Indicates negative association

0 Indicates no change

Significance of Shift: Shift was found to have significance on response (Yes/No) and number of ignored alarms for unit clerks, which is obvious. From the significance of shift, explained on alarm fatigue measures, it has been shown that different shift had different levels of mental workload and affect on unit clerks and nurses. Although, for

nurses, the mental workload was higher during days compared to nights, surprisingly shift did not show its significance on their responses (Yes/No) and number of ignored alarms. This happens because of their responsibility to respond to alarms for patient safety.

Significance of Staff ID: Similar to the shift, staff ID also was found to have its influence on response (Yes/No) for unit clerks, but not for nurses. Willingness to respond to alarms depends on the experience of unit clerks. A novice unit clerk was observed to have responded to almost every alarms she could while in the case of other experienced clerks, they don't bother to respond to the nuisance alarms most of the time.

Significance of Workload: Workload was expected to find its positive influence on the staff performance. Independent analysis of workload with performance measures found that increasing workload results in increased number of ignored alarms for both unit clerks and nurses, and delayed response times to alarms for unit clerks only. Nurses do not show any influence of workload on response time as they tries to answer all the alarms going off in the patient's room to keep the rooms comfortable and noise free for patients. See Figures H21, H22, H30, and H32 in Appendix H.

Significance of Number of Alarms: Unit clerks are the people who actually experience most of the alarms in the ICU. A large number of alarms can make them fatigued and deteriorate their performance in terms of reduced response and increased

response times. Graphs (see Figures H23 and H29 in Appendix H) also show the influence of number of alarms on performance for unit clerks.

Significance of Alarm Type: Alarm type often makes a difference on staff response attitudes. Criticality of alarm type can be patient specific. Alarm type was found to have an influence on response and response time for unit clerks and response and number of ignored alarms for nurses.

Significance of Alarm Criticality: Staff experience a large number of alarms throughout the 12 hours long shift and unit clerks take care of a major portion of those alarms. Very few alarms are critical among those requiring any action. It was found that both the unit clerks and nurses become very conscious whenever there is a highly critical alarm, call light or phone call. All those noises requires quick response and therefore, response has the influence of alarm type for both types of participants. Response time for unit clerks was also found to have shown the influence of alarm criticality. The more critical the alarms, the faster the responses to these, which is obvious.

Significance of Time Elapsed: Unit clerks always need to sit in front of two telemetry monitors during their 12 hours shift. The constant beeps coming out of two monitors and continuous sedentary job make them stressed causing delayed response to alarms and increased number of ignored alarms. For nurses also, at the end of the shift high mental workload was found, which makes them ignore alarms frequently at a certain

level of fatigue or ignore quick response to alarms. These results can be supported by the Figures H26, H31, and H33.

Significance of Task Priority: Task analyses showed the importance of task prioritization for the staff at ICU. Several times a situation arrives where staff needs to prioritize their tasks. Thus task priority was found to have influence on response (yes/no) for both participants and response time for unit clerks only. While the staff remains busy doing a task, s/he can miss responding to an alarm going off at the same time.

Significance of Noise Level: Noise level is another important issue found in the ICU influencing staff performance which are shown in Figures H27 and H35. High noise level created from phone calls, call lights, and highly critical alarms disrupt patient comfort and is not acceptable in an acute care area for long time. Therefore, the louder the alarm, the faster the response which eventually decreases the number of ignored alarms.

Significance of Extraversion: Staff having a higher percentile for extraversion are energetic and talkative. These people like to have conversations with other staff in the facility and show their energy in all the tasks. As a result, they can get fatigued faster and make delayed responses to alarms (Figure H 28).

Significance of Agreeableness: Agreeableness is the personality traits of being helpful or cooperative to others. This factor has shown its influence on the response time

for nurses. The more helpful the staff is the more increased response time s/he shows in responding to alarms. Generally, one nurse is recruited for two patients. Often nurses need support from other coworkers. Furthermore, nurses need to provide support to the staff from other departments (physiotherapy, respiratory therapy etc.) as well. Therefore, the person showing more agreeableness i.e., more cooperative to others, sometimes have delays in responding to their own patient's alarms (Figure H24)

Significance of Neuroticism: People having low scores for neuroticism have less anxiety and less depression, which should make them slower in responding to alarms. This can be explained in that the staff do not get stressed with workload, criticality of patients, and large number of alarms or high noise level. Therefore, they do not care about the alarms that much to respond to them quickly (see Figure H25).

Significance of Openness: Staff having high values for openness like to experiment with new things. They are the people who like to take stress as curiosity and make them involved in many complex tasks other than routine work. Therefore, this type of personality lead the staff toward delayed response. Figures H21 and H34 support this discussion.

Significance of SWAT Measures: SWAT measures show the level of mental fatigue for staff. Increased mental workload results in longer response time and higher number of ignored alarms for unit clerks. On the other hand nurses show opposite results. Increases in SWAT measure make the nurses respond quickly to alarms. Nurses are

directly connected to patient care. Their mental workload can be the result of critical condition of the patient other than the alarm fatigue. Therefore, with increased mental workload, they try to monitor the patient carefully and reduce alarms by taking quick active response to them. See figures H21 and H34.

Significance of NASA TLX Measures: NASA TLX measures show similar results as SWAT measures which validate the measures of alarm fatigue using two different mental workload measuring tools. This similarity validates the use of any one of those tools. Graphs are attached as Figures H21 and H34 in Appendix H.

Significance of Boredom: Boredom was found to have influence on the response (yes/no) and response time for unit clerks (see Figures H21 and H34). Boredom is the state of being indifferent from where people can come out if the other conditions around them changes. Unit clerks are responsible for monitoring a large number of alarms. Several times the monotony of this job makes them bored. Again, experience of critical alarms, phone calls, other nurses to have conversations, make them energetic toward monitoring alarms. Although boredom showed its influence, no change was observed in the case of individual influence testing.

Significance of Apathy: Apathy showed its significance on the number of ignored alarms for nurses (Figure H34). As discussed earlier, at the end of the shifts, nurses become highly indifferent about the alarms and do not want to respond to them anymore. Therefore, the number of ignored alarms increased for nurses with the increase in apathy

level. In reality, at the end of the shifts nurses have to go through sever meetings to exchange their responsibly to the nurses for the next shift. As a result, they don't hear the alarms and don't reply to them.

Significance of Distrust: Distrust has influence on response (yes/no), response time, and number of ignored alarms for unit clerks only. Unit clerks get most of the nuisance alarms because they have types of alarms which can result from patient movements, battery low and no signal conditions. These high number of nuisance alarms create distrust in unit clerks and influence their response (yes/no), increase response time and consequently increase number of ignored alarms. Figures H20 and H21 are attached in Appendix H.

CONCLUSIONS

The objective of this study was to create a definition of alarm fatigue based on all the possible influencing factors so that hospitals can determine which factors need to be taken care of in order to design a safer work environment for their staff. This research also intended to investigate the influence of alarm fatigue on staff performance. Until this point, all the previous study tried to measure staff performance based on the working conditions and blamed alarm fatigue as the reason of any deterioration in performance. Most of the previous studies recommended several methods to reduce false alarms and patient specific threshold parameters settings, whereas, this study identified the necessity of a quantitative definition of alarm fatigue so that all the possible factors and their influence on the alarm fatigue measure can be determined. This definition opens the scope of further research for reducing alarm fatigue in staff analyzing the influences of all possible factors that can cause alarm fatigue. In addition, identification of the influence of alarm fatigue on staff performance investigates the role of alarm fatigue on performance deterioration which will help the future research in this arena.

The study began with an assessment of staff attitude toward the existing alarm monitoring system in the Intensive Care Unit (ICU) at Bozeman Deaconess Hospital (BDH). The assessment confirmed the frequent occurrence of nuisance alarms at the chosen facility and their effect on the ICU staff, including disruption of patient care and increase in distrust toward the alarms, leading to the staff decision to turn the alarms off.

With the confirmation that alarm-related problems exist in the chosen facility, this study was able to address the five research questions at BDH. As a consequence, an extensive literature review and Hierarchical Task Analyses (HTA) gleaned all the possible alarm fatigue causing factors (working conditions and staff individuality). Working condition factors included shift (day/night), workload (staff to patient ratio), time elapsed, task priority, number of alarms, alarm type, alarm criticality, and noise level. Staff individuality factors included staff ID and personality type dimensions: extraversion, openness, agreeableness, neuroticism, and conscientiousness.

Data were then collected for both unit clerks and nurses to define and measure alarm fatigue and investigate its influence on staff performance by quantifying alarm fatigue causing factors (working conditions and staff individuality), alarm fatigue measures (mental workload and affect), and staff performances (response [yes/no], response time, and number of ignored alarms). Definitions of alarm fatigue using mental workload and affect showed dissimilarities in the case of unit clerks, but not for nurses. This result support the findings from HTA that fatigue for unit clerks and nurses need to be investigated separately. In the case of unit clerks, the definitions using mental workload included most of the variables from working conditions while the definition using affect showed almost all of the personality dimensions except neuroticism. Therefore, for unit clerk one single definition can be proposed using all of the measures considered in this study which includes all the significant factors that influence unit clerk's fatigue. On the other hand, for nurses, either mental workload or affect can be

used to define alarm fatigue. Besides, as almost all of the personality type dimensions showed influence for causing fatigue on both groups of participants, personality type test can be added to the staff recruitment policy. In addition, for any kind of technical failure, support should be made available for night shifts as well.

After defining and identifying measures of alarm fatigue, several analyses attempts were run to associate alarm fatigue and staff performance: 1) using alarm fatigue measures as independent variables, 2) using alarm fatigue causing factors as independent variables, and 3) combined factors (alarm fatigue measures along with alarm fatigue causing factors) as independent variables. These analyses revealed that the original conceptual model is invalid and that alarm fatigue alone is not the cause of deterioration in staff performance, as has been strongly purported until now. This study, however, has shown that performance deterioration is actually the result of a combination of alarm fatigue with working conditions and staff individuality, overall, and in the case of nurses and response time, alarm fatigue plays no role, only working conditions and staff individuality.

Multiple recommendations can be addressed based on some unique contributions from the findings of the study. The recommendations are summarized below:

- 1) The list of the variables, found from literature review as well as from Hierarchical Task Analyses, needs to be considered in the case of any alarm related study.

2) Alarm fatigue should be defined and measured separately for unit clerks and nurses. For nurses any of the two measures (mental workload or affect) can be used, but in the case of unit clerks, all the measures are required to give a complete definition for them.

3) The role of alarm fatigue in the deterioration of staff performance must, therefore, be reevaluated. Until now, the generally held view has been that alarm fatigue is the cause of performance deterioration. This study has shown that it is not only alarm fatigue but also working conditions contributes to performance deterioration. Hospitals which desire to design a safer environment for their staff and patients, must need to consider working conditions and staff individuality in addition to alarm fatigue.

4) Personality type has very important influences on alarm fatigue as well as staff performance. Therefore, personality type testing is highly recommended to be included in the staff recruitment policy.

Figures 23 and 24 summarize all the analyses attempts along with the proposed models to define alarm fatigue and identify the cause of performance deterioration for unit clerks and nurses, respectively. These figures also include the type of influences (positive/negative) for all the associations found through regression analyses.

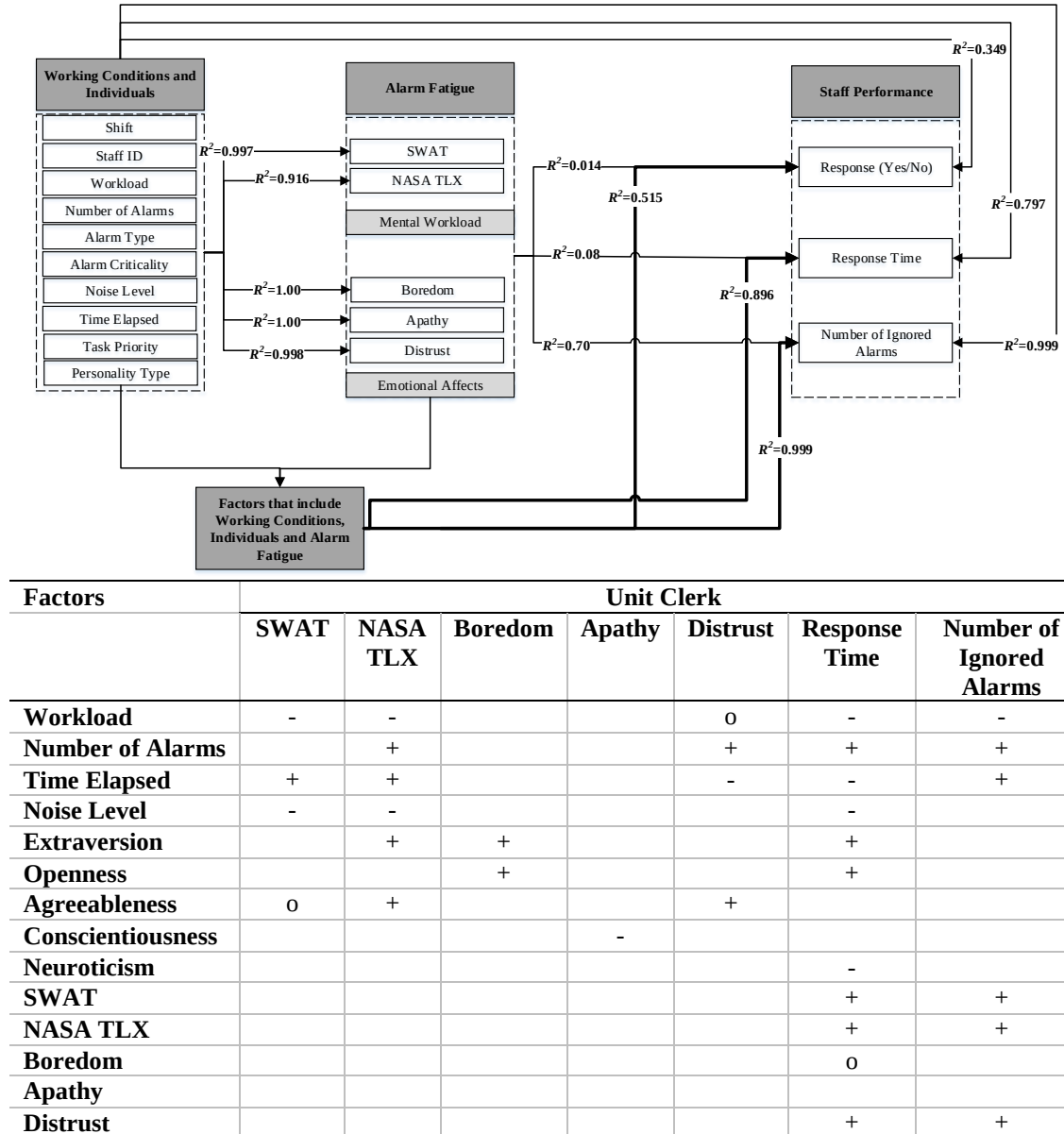


Figure 23: Conceptual model for unit clerks

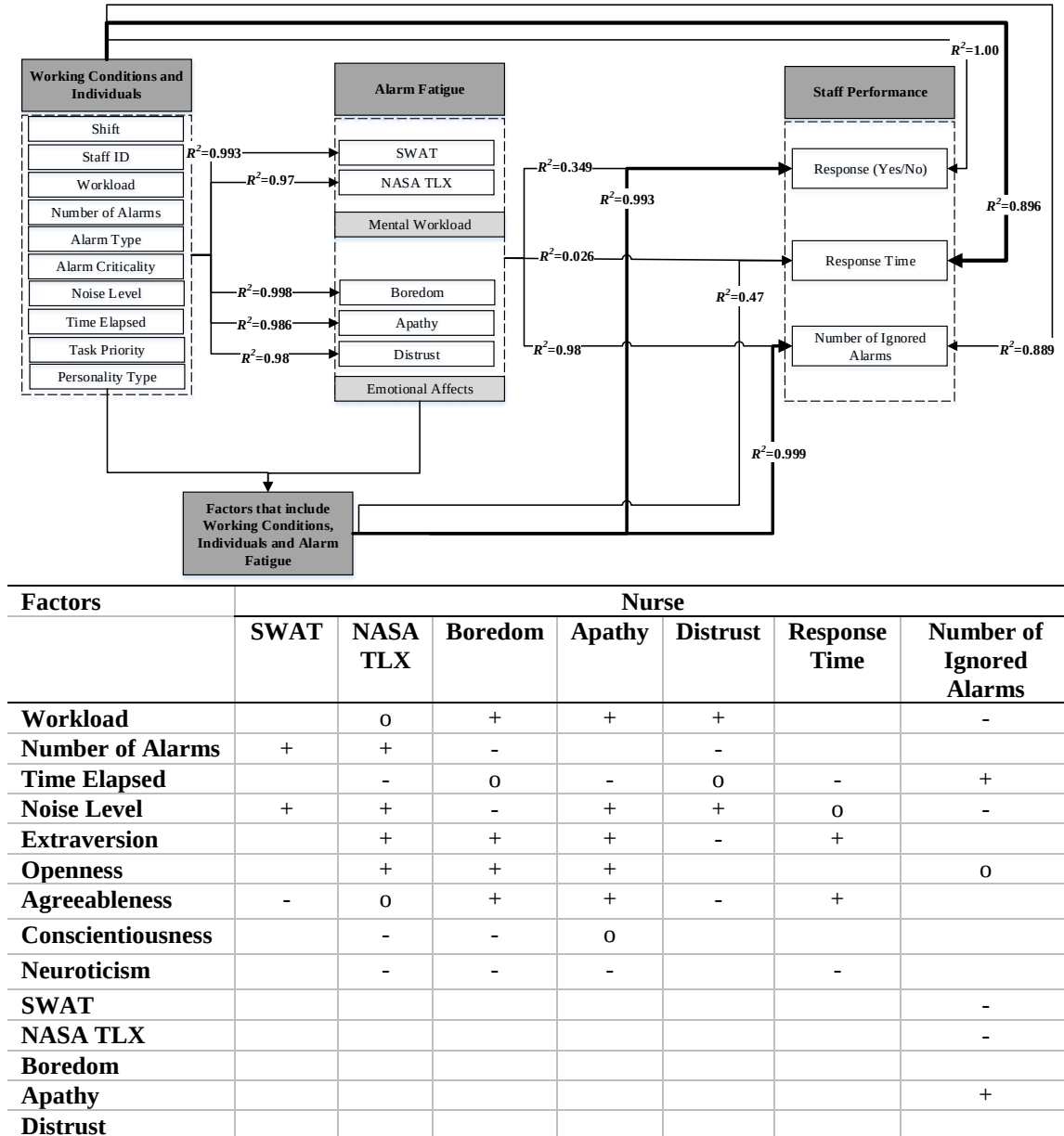


Figure 24: Conceptual model for nurses

The study was conducted in an acute care area. Therefore, several limitations were experienced during research work to ensure patient safety. Although the shifts are twelve hours long, only six-hour observation periods were designed for the study. The presence of researchers (two at a time) may have interrupted the usual alarm monitoring tasks. Additionally, with only two researchers observing two telemetry monitors as well as conducting two surveys (SWAT and Distrust), some alarms may have been missed. The study was designed for twelve hours slots, but due to difficulties experienced in data collection regarding unusual situations, only seven observational slots were used for analyses. Two noise meters were used in front of each telemetry monitor to collect data on noise level for every single alarm coming from these monitors. However, the noise level for the alarm in the patient rooms was also measured from central monitoring area and was, therefore, not accurate. Finally, this study was focused on only Intensive Care Unit as an acute care area. Other acute care areas in this hospital could be studied to validate the results.

This study opens up possibilities for future research into defining and measuring alarm fatigue and its influence on staff performance. The definitions of alarm fatigue proposed by this study can be validated by conducting further research on the acute care areas of the same hospital, as well as on several acute care departments of other hospitals. In addition, physiological measures can be used along with mental and affect to produce a complementary definition of alarm fatigue.

REFERENCES CITED

- Abdellah, F. G., & Levine, E. (1954). Work sampling applied to the study of nursing personnel. *Nursing Research*, 3(1), 11-16.
- Ackerman, P. L., & Heggstad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological Bulletin*, 121(2), 219-245.
- Ashton, J. (2011). "Alarm fatigue" blamed in hospital's deaths. Retrieved October 24, 2014, from <http://www.cbsnews.com/news/alarm-fatigue-blamed-in-hospital-deaths/>
- Association for the Advancement of Medical Instrumentation. (2011). *A siren call for action*. Arlington: Summit Publications: Clinical Alarms.
- Baillargeon, E. (2013). Alarm Fatigue: A Risk Assessment. Rhode Island College: Master of Science in Nursing. Retrieved October 24, 2014, from http://digitalcommons.ric.edu/school_of_nursing/23
- Barker Steege, L. M., & Nussbaum, M. A. (2012). Dimensions of Fatigue as Predictors of Performance: A Structural Equation Modeling Approach Among Registered Nurses. *IIE Transactions on Occupational Ergonomics and Human Factors*, 1(1), 16-30. doi:10.1080/21577323.2011.637153
- Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*, 44(1), 1-26. doi:10.1111/j.1744-6570.1991.tb00688.x
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment*, 9(1-2), 9-30. doi:10.1111/1468-2389.00160
- Barry, E. K., & McCarthy, P. (2001). The relationship between personality type and perceived hearing aid benefit. *The Hearing Journal*, 54(9), 41-46.
- Battiste, V., & Bortolussi, M. (1988). Transport Pilot Workload: A Comparison of Two Subjective Techniques. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 32(2), pp. 150-154. San Diego: Human Factors Society. doi:10.1177/154193128803200232
- Billingshurst, F., Morgan, B., & Arthur, H. M. (2003). Patient and Nurse-Related Implications of Remote Cardiac Telemetry. *Clinical Nursing Research*, 12(4), 356-370.

- Borowski, M., Gorges, M., Fried, R., Such, O., Wrede, C., & Imhoff, M. (2011). Medical Device Alarms. *Biomed Tech (Berl)*, 56(2), 73-83. doi:10.1515/BMT.2011.005
- Chambrin, M. C., Ravoux, P., Calvelo-Aros, D., Jaborska, A., Chopin, C., & Boniface, B. (1999). Multicentric Study of Monitoring Alarms in the Adult Intensive Care Unit (ICU): A descriptive Analysis. *Intensive Care Med*, 25(12), 1360-1366.
- Chamorro-Premuzic, T., Moutafi, J., & Furnham, A. (2005). The relationship between personality traits, subjectively-assessed and fluid intelligence. *Personality and Individual Differences*, 38(7), 1517-1528.
doi:<http://dx.doi.org/10.1016/j.paid.2004.09.018>
- Chien, S.-Y., Semnani-Azad, Z., Lewis, M., & Sycara, K. (2014). Towards the Development of an Inter-cultural Scale to Measure Trust in Automation. 35-46. Heraklion, Crete, Greece: Springer. Retrieved from <http://d-scholarship.pitt.edu/21820/>
- Clark, T., & David, Y. (2011). *Advancing the Safety of Healthcare Technology*. Retrieved from World Health Organization website:
http://www.who.int/medical_devices/global_forum/D05.pdf?ua=1
- Clinical Alarms Task Force. (2007). Impact Of Clinical Alarms on Patient Safety: A report from the American College of Clinical Engineering Healthcare Technology Foundation. *Journal of Clinical Engineering*, 32(1), 22-33. Retrieved from <http://thehtf.org/white%20paper.pdf>
- Coomber, S., Todd, C., Park, G., Baxter, P., Firth-Cozens, J., & Shore, S. (2002). Stress in UK intensive care unit doctors. *British Journal of Anaesthesia*, 89(6), 873-881.
- Crites, A. (2013). *Too much noise from hospital alarms poses risk for patients*. Retrieved from Washington Post website:
<http://www.washingtonpost.com/sf/feature/wp/2013/07/07/too-much-noise-from-hospital-alarms-poses-risk-for-patients/>
- Cvach, M. (2012). Monitor Alarm fatigue: An integrated review. *Biomed Instrum Technol*, 46(4), 268-277. doi:10.2345/0899-8205-46.4.268
- DCRC. (n.d.). *Guidelines for Apathy Evaluation Scale*. Retrieved October 25, 2014, from http://www.dementia-assessment.com.au/symptoms/AES_Guidelines.pdf
- Diaper, D., & Stanton, N. (2003). *The Handbook of Task Analysis for Human-Computer Interaction*. CRC Press.

- Dickson, W. M. (1978). Measuring Pharmacist Time Use: A Note on the Use of Fixed Interval Work Sampling. *American Journal of Hospital Pharmacy*, 35(10), 1241-1243.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41(1), 417-440. doi:10.1146/annurev.ps.41.020190.002221
- ECRI . (2013). Top 10 Health Technology Hazards for 2014. *Health Devices*, 42(11).
- ECRI. (1974). Hazard: Gaymar hypothermia machine. *Health Devices*, 229–230.
- ECRI. (2007). The hazards of alarm overload: Keeping excessive physiologic monitoring alarms from impeding care. *Health Devices*, 36(3), 73-83.
- ECRI. (2011). Top 10 Health Technology Hazards for 2012. *Heal Devices*, 40(11). Retrieved from https://www.ecri.org/Documents/secure/Health_Devices_Top_10_Hazards_2012.pdf
- ECRI. (2012). Top 10 Health Technology Hazards for 2013. *Health Devices*, 41(11). Retrieved from https://www.ecri.org/Documents/Secure/Health_Devices_Top_10_Hazards_2013.pdf
- Farmer, R., & Sundberg, N. D. (1986). Boredom proneness--The development and correlates of a new scale. *Journal of Personality Assessment*, 4-17. doi:10.1207/s15327752jpa5001_2
- Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations*, 46(3), 395–417. doi:10.1177/001872679304600305
- Franklin, C., Johnson, L. V., White, L., Franklin, C., & Smith-Olinde, L. (2013). The Relationship between Personality Type and Acceptable Noise Levels: A Pilot Study. *ISRN Otolaryngology*, 1-6. doi:10.1155/2013/902532
- Freyaldenhoven, M. C., Smiley, D. F., Muenchen, R. A., & Konrad, N. T. (2006). Acceptable noise level: reliability measures and comparison to preference for background sounds. *Journal of the American Academy of Audiology*, 17(9), 640-648.
- Geen, R. G. (1984). Preferred stimulation levels in introverts and extroverts: Effects on arousal and performance. *Journal of Personality and Social Psychology*, 46(6), 1303-1312. doi:http://dx.doi.org/10.1037/0022-3514.46.6.1303

- Graham, K. C., & Cvach, M. (2010). Monitor Alarm Fatigue: Standardizing use of physiological monitors and decreasing nuisance alarms. *Am J Crit Care*, 19(1), 28-34. doi:10.4037/ajcc2010651
- Hart, S. G. (2006). NASA-TASK Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), pp. 904-908. doi:10.1177/001872089203400405
- Hart, S. G., & Staveland, L. E. (1988). Development of Nasa-TLX (Task Load Index): Results of Empirical and Theoretical Research. In A. H. Peter, & M. Najmedin (Eds.), *Advances in Psychology* (Vol. 52, pp. 139-183). North Holland.
- Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of four subjective workload Rating Scales. *Human Factors: The Journal of Human Factors and Ergonomics Society*, 34(4), 429-439. doi:10.1177/001872089203400405
- Hoonakker, P., Carayon, P., Gurses, A., Brown, R., McGuire, K., Khunlertkit, A., & Walker, J. M. (2011). Measuring Workload of ICU Nurses with a Questionnaire Survey: The NASA Task Load Index (TLX). *IIE Transactions Healthcare System Engineering*, 1(2), 131-143. doi:10.1080/19488300.2011.609524
- Hsieh, C. J., Chu, H., Cheng, J. J., Shen, W. W., & Lin, C. C. (2012). Validation of apathy evaluation scale and assessment of severity of apathy in Alzheimer's disease. *Psychiatry and Clinical Neurosciences*, 227-234. doi:10.1111/j.1440-1819.2011.02315.x
- Hutchinson, K. M., Duffy, T. L., & Kelly, L. J. (2005). How personality types correlate with hearing aid outcome measures. *The Hearing Journal*, 58(7), 28-34. doi:10.1097/01.HJ.0000286418.68332.c1
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: a meta-analysis. *Journal of Applied Psychology*, 87(3), 530-541.
- Kaplan, R., & Saccuzzo, D. (2013). *Psychological Testing: Principle, Applications, and Issues* (8th ed.). Belmont: Cengage Learning.
- Kerr, J. H., & Hayes, B. (1983). An "Alarming" Situation in the Intensive Therapy Unit. *Intensive Care Medicine*, 9(3), 103-104.
- Kowalczyk, L. (2011, November 28). Suit Over Cardiac Monitor Settled. Retrieved from <http://www.bostonglobe.com/metro/2011/11/28/suit-over-cardiac-monitor-settled/GajKUFUgsjltsKMNWDpsuN/story.html>

- Lawless, S. (1994). Crying Wolf: False Alarms in a Pediatric Intensive Care Unit. *Critical Care Medicine*, 22(6), 981-985.
- Levy, M. M. (2004). Caring for the caregiver. *Critical Care Clinics*, 20(3), 541-547.
- Liptak, G. S., Super, D. M., Baker, N., & Roghmann, K. J. (1985). An analysis of waiting times in a pediatric emergency department. *Clinical Pediatrics (Phila)*, 24(4), 202-209.
- Luximon, A., & Goonetilleke, R. S. (1998). Continuous Subjective Workload Assessment Technique. *First World Congress on Ergonomics for Global Quality and Productivity*. Hong Kong. Retrieved from <http://repository.ust.hk/ir/bitstream/1783.1-1416/1/ergon01.pdf>
- Marin, R. S. (1991). Apathy: a neuropsychiatric syndrome. *J Neuropsychiatry Clin Neurosci*, 3(3), 243-254.
- Marin, R. S., Biedrzycki, R. C., & Firinciogullari, S. (1991). Reliability and Validity of the Apathy Evaluation Scale. *Psychiatry Research*, 38(2), 143-162.
- Mealer, M. L., Shelton, A. B., Rothbaum, B., & Moss, M. (2007). Increased Prevalence of Post-traumatic Stress Disorder Symptoms in Critical Care Nurses. *American Journal of Respiratory and Critical Care Medicine*, 175(7), 693-697. doi:10.1164/rccm.200606-735OC
- Meeusen, V. C., Brown-Mahoney, C., van Dam, K., van Zundert, A. A., & Knape, J. T. (2010). Personality dimension and their relationship with job satisfaction among dutch nurse anaesthetists. *Journal of Nursing Management*, 18(5), 573-581. doi:10.1111/j.1365-2834.2010.01066.x
- Mikulas, W. L., & Vodanovich, S. J. (1993). The essence of boredom. *The Psychological Record*, 43(1), 3-12.
- Mitka, M. (2013). Joint Commission Warns of Alarm Fatigue: Multitude of alarms from monitoring devices problematic. 2315-2316. *JAMA*, 309(22). doi:10.1001/jama.2013.6032
- Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the NEO-Personality Inventory and the Myers-Briggs Type Indicator. *European Journal of Personality*, 17(1), 79-94.
- Pishori, T. (2012). A Stochastic Simulation Model of Alarm Response Strategies on a Telemetry Floor. (Masters of Science). Paper 256. University of Connecticut. Retrieved from http://digitalcommons.uconn.edu/gs_theses/256/

- Reid, G. B., Potter, S. S., & Bressler, J. (1989). Subjective Workload Assessment Technique (SWAT): A user's guide. Wright Patterson Air Force Base, OH: Harry G. Armstrong Aerospace Medical Research Laboratory.
- Reid, R. A. (1975). A Work Sampling Study of Midlevel Health Professionals in a Rural Medical Clinic. *Medical Care*, 13(3), 241-249.
- Rothmann, S., & Coetzer, E. P. (2003). The Big Five Personality Dimensions and Job Performance. *Journal of Industrial psychology*, 29, 68-74.
doi:10.4102/sajip.v29i1.88
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, 53(1), 61-86. doi:10.1111/j.1464-0597.2004.00161.x
- Sendelbach, S., & Funk, M. (2013). Alarm fatigue: a patient safety concern. *AACN Advanced Critical Care*, 24(4), 378-386.
- Stafford, T. (1982). Whether Monitoring? *Critical Care Med*, 10(11), 792-795.
- Surani, S., Hesselbacher, S., Guntupalli, B., Surani, S., & Subramanian, S. (2014). Sleep Quality and Vigilance Differ Among Inpatient Nurses Based on the Unit Setting and Shift Worked. *Journal of Patient Safety*. doi:10.1097/pts.0000000000000089
- Tett, R. P., & Murphy, P. J. (2002). Personality and Situations in Co-worker Preference: Similarity and Complementary in Worker Compatibility. *Journal of Business and Psychology*, 17(2), 223-243. doi:10.1023/A:1019685515745
- The Joint Commission. (2013). *Sentinel Event Alert*. (50). Retrieved from http://www.jointcommission.org/assets/1/18/sea_50_alarms_4_5_13_final1.pdf
- The Joint Commission. (2014). Retrieved March 23, 2014, from Premier: https://legacy.premierinc.com/quality-safety/tools-services/safety/topics/patient_safety/index_3.jsp
- Tobacyk, J. J., Livingstone, M. M., & Robbins, J. E. (2008). Relationships Between Myers-Briggs Type Indicator Measure of Psychological Type and NEO measure of Big Five Personality Factors in Polish University Students: A Preliminary Cross-Cultural Comparison. *Psychol Rep*, 103(2), 588-590.
doi:10.2466/pr0.103.2.588-590
- Tsien, C. L., & Fackler, J. C. (1997). Poor Prognosis of Existing Monitors in the Intensive Care Unit. *Critical Care Medicine*, 25(4), 614-619.

- Wallis, L. (2010). Alarm Fatigue Linked to Patient's Death. *American Journal of Nursing*, 110(7), 16. doi:10.1097/01.NAJ.0000383917.98063.bd
- Weinger, M. B., & Smith, N. T. (1993). Vigilance, Alarms and integrated Monitoring systems. In *Anesthesia equipment: principles and applications* (pp. 350-384). Malvern, PA: Mosby Year Book.
- Welch, J. (2011). An Evidence-based Approach to Reduce Nuisance Alarms and Alarm Fatigue. *Biomedical Instrumentation & Technology: Alarm Systems*, 45(1), 46-52. doi:10.2345/0899-8205-45.s1.46
- Williams, P. M. (2001, April). Techniques for root cause analysis. *Proceedings of Baylor University Medical Center*, 14(2), pp. 154-157.
- Young, G., Zavelina, L., & Hooper, V. (2008). Assessment of workload using NASA Task Load Index in perianesthesia nursing. *Journal of Perianesthesia Nursing*, 23(2), 102-110. doi:10.1016/j.jopan.2008.01.008
- Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In S. R., *Handbook of intelligence* (pp. 581-610). New York, US: Cambridge University Press.

APPENDICES

APPENDIX A

APPROVALS FOR THE STUDY



INSTITUTIONAL REVIEW BOARD
For the Protection of Human Subjects
FWA 00000165

960 Technology Blvd. Room 127
 c/o Immunology & Infectious Diseases
 Montana State University
 Bozeman, MT 59718
 Telephone: 406-994-6783
 FAX: 406-994-4303
 E-mail: cherylj@montana.edu

Chair: Mark Quinn
 406-994-5721
 mquinn@montana.edu
Administrator:
 Cheryl Johnson
 406-994-6783
 cherylj@montana.edu

MEMORANDUM

TO: Shuchisnigdha Deb, FNU Pawan Kumar, David Claudio
FROM: Mark Quinn, Chair *Mark Quinn CH*
DATE: February 26, 2014
RE: "National Patient Safety Goal on Alarm Management" [SD022614-EX]

The above research, described in your submission of February 26, 2014, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

- ☐ (b) (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
- ☒ (b) (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.
- ☐ (b) (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
- ☐ (b) (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available, or if the information is recorded by the investigator in such a manner that the subjects cannot be identified, directly or through identifiers linked to the subjects.
- ☐ (b) (5) Research and demonstration projects, which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.
- ☐ (b) (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed, or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the FDA, or approved by the EPA, or the Food Safety and Inspection Service of the USDA.

Although review by the Institutional Review Board is not required for the above research, the Committee will be glad to review it. If you wish a review and committee approval, please submit 3 copies of the usual application form and it will be processed by expedited review.

SUBJECT CONSENT FORM

FOR PARTICIPATION IN NATIONAL PATIENT SAFETY GOAL ON ALARM MANAGEMENT AT BOZEMAN DEACONESS HOSPITAL

You are being invited to take part in a survey for patient safety goal that is intended to assess alarm fatigue at Intensive Care Unit (ICU) in Bozeman Deaconess Hospital (BDH) and propose possible solutions to mitigate alarm fatigue.

The purpose of this survey is to collect the initial responses from the nurses and other staffs in the ICU to assess the current state. The primary concern is to investigate whether there is a large number of nuisance alarms with a noisy environment that makes the nurses silence or ignore critical alarms due to sensory overload or fatigue.

Procedures: If you agree to take part in this survey, you will be asked to fill out a clinical alarm survey form proposed by ACCE Healthcare Technology Foundation. We will highly appreciate your time and complete honest responses on that.

Risks: No risk for the participants is associated with this survey.

Benefits: This survey is designed for further improvements to reduce the alarm fatigue, if there is any. The results will be beneficial for the hospital in the accomplishment of Joint Commission's requirement on "National Patient Safety Goals on Alarm Management".

Participation: Your participation will be completely voluntary.

Confidentiality: Your confidentiality will be fully protected. This survey will be completely anonymous.

Questions: Questions or complaints about the research should be directed to Shuchisnigdha Deb (336-402-8425), Dr. Durward K. Sobek (406-994-7140), Dr. Linda Torma (406-243-2599) or Dr. David Claudio (406-994-5943), Montana State University, Bozeman, MT 59715, USA.

You have certain rights as a participant in this research. Questions about these rights should be directed to Dr. Mark Quinn, Chair of the Human Subjects Committee, Montana State University – Bozeman, MT. Phone: 406-994-4707.

Your signature indicates that you have read this consent form and are agreed to participate in this survey.

Name of subject (Print): _____ Date: _____

Signature of subject: _____

Signature of Investigator: _____

*Approved by
V Reddy*

SUBJECT CONSENT FORM

FOR PARTICIPATION IN NATIONAL PATIENT SAFETY GOAL ON ALARM MANAGEMENT AT BOZEMAN DEACONESS HOSPITAL

You are being invited to take part in some surveys for patient safety goal that is intended to define alarm fatigue at Intensive Care Unit (ICU) in Bozeman Deaconess Hospital (BDH) and propose possible solutions to mitigate alarm fatigue.

The purpose of this survey is to collect the responses from the nurses and unit clerks in the ICU to assess their cognitive workload and emotional affects from a whole shift. The primary concern is to define alarm fatigue in terms of cognitive workload and emotional affect.

Procedures: If you agree to take part in these surveys, you will be asked to fill out some valid survey tools for cognitive workload, boredom, and apathy. We will highly appreciate your time and complete honest responses on that.

Risks: No risk for the participants is associated with this survey.

Benefits: This survey is designed for further improvements to reduce the alarm fatigue. The results will be beneficial for the hospital in the accomplishment of Joint Commission's requirement on "National Patient Safety Goals on Alarm Management".

Participation: Your participation will be completely voluntary.

Confidentiality: Your confidentiality will be fully protected. This survey will be completely anonymous.

Questions: Questions or complaints about the research should be directed to Shuchisnigdha Deb (336-402-8425) or Dr. David Claudio (406-994-5943), Montana State University, Bozeman, MT 59715, USA.

You have certain rights as a participant in this research. Questions about these rights should be directed to Dr. Mark Quinn, Chair of the Human Subjects Committee, Montana State University – Bozeman, MT. Phone: 406-994-4707.

Your signature indicates that you have read this consent form and are agreed to participate in this survey.

Name of subject (Print): _____ Date: _____

Signature of subject: _____

Signature of Investigator: _____

*Approved
Vgrosne
CNR
7-10-14*

APPENDIX B

CLINICAL ALARM SURVEY

CLINICAL ALARM SURVEY FORM



ACCE Healthcare Technology Foundation
 8200 Butler Pike
 Plymouth Meeting, PA 19462-1298
 Phone: 610-825-6067
 www.acce-htf.org

Clinical Alarms Survey

Please insert only one response for each question

A) Facility	Acute Care Hospital		Ambulatory Care Facility or Surgery Center		Sub Acute Care		
	Nursing Home		Assisted Living/Rehabilitation		Other (write in)		
B) Hospital department (if applicable)	ICU	Nursery	ER	Support			
	OR/Anes	General Floor	Progressive Care	Other (write in)			
C) Job Title	Physician	RN	LPN	Clinical manager			
	Nurses aide or Orderly	Respiratory therapy	Clinical Engineer	BMET			
	Paramedical e.g. Rad/Lab/Resp	Administrator/Non-clinical manager	Transport	Other (write in)			
D) Years Experience	0-3	3-6	6-11	11+			
E) Alarm-Related Information:			Strongly Agree	Agree	Neutral	Disagree	Strongly disagree
1.	The purpose of clinical alarms is to alert staff of an existing or potentially hazardous patient condition						
2.	Alarm sounds and/or visual displays should differentiate the priority of alarm						
3.	Alarm sounds and/or visual displays should be distinct based on the parameter or source (e.g. device)						
4.	Alarms should impact multiple senses (audible, visual, proprioceptive, etc.)						
5.	Nuisance alarms occur frequently						
6.	Nuisance alarms disrupt patient care						
7.	Nuisance alarms reduce trust in alarms and cause care givers to turn alarms off at times other than setup or procedural events						
8.	Properly setting alarm parameters and alerts is overly complex in existing devices						
9.	New (less than three years old) monitoring systems have solved most of the previous problems we experienced with clinical alarms						
10.	Since the implementation of the JCAHO Patient Safety Goal #6, now part of the overall JCAHO standards, patient adverse events related to clinical alarms have been reduced to an acceptable level.						

(Continued)

APPENDIX C

OBSERVATIONAL DATA COLLECTION TABLE

DATA COLLECTION TABLE

Number of Patients:				Number of Staff:				Name of the Observer:				Shift:	Time:	
	Alarm				Reading						Criticality	Response		
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			
Room #	HR	NBP	IV Pump	PI Vent	Asystole	Vent fib	Brady	Apnea	Incomplete	Noisy	High	Time	: Actionable	
	RESP	ABI	Bipap	Balloon Pump	high Occl	Pulse high	Air Temp	V tach	Failed	FALSE	Medium	Response Time	: Non-actionable	
	SPO2	EKG		Cpap	low PVC	Pulse low	leak	Incorrect R	Unplugged	Erratic	Low	Comment		
	Feed pump	SCD		Epidural	Leads-off	Missed beep		Phone	ST el	Incorrect	Comment			

APPENDIX D

ALARM FATIGUE MEASURING SCALES

BOREDOM PRONENESS SCALE

Please circle the option that best indicates your experience during the shift period.

1. It was easy for me to concentrate on my activities

Strongly Disagree Disagree Neutral Agree Strongly Agree

2. I found it easy to entertain myself

Strongly Disagree Disagree Neutral Agree Strongly Agree

3. Many things I had to do were repetitive and monotonous

Strongly Disagree Disagree Neutral Agree Strongly Agree

4. I got a kick out of most things I did

Strongly Disagree Disagree Neutral Agree Strongly Agree

5. In any situation I could usually find something to do or see to keep me interested

Strongly Disagree Disagree Neutral Agree Strongly Agree

6. It would be very hard for me to find a job that is exciting enough

Strongly Disagree Disagree Neutral Agree Strongly Agree

7. Many people would say that I am a creative or imaginative person

Strongly Disagree Disagree Neutral Agree Strongly Agree

8. Among my friends, I am the one who keeps doing something the longest

Strongly Disagree Disagree Neutral Agree Strongly Agree

9. Unless I am doing something exciting, even dangerous, I feel half-dead and dull

Strongly Disagree Disagree Neutral Agree Strongly Agree

10. When I was young, I was often in monotonous and tiresome situations

Strongly Disagree Disagree Neutral Agree Strongly Agree

APATHY EVALUATION SCALE

Please circle the option that best indicates your experience during the shift period

1. I get things done during the day

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

2. Getting things started on my own is important to me

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

3. I am interested in having new experiences

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

4. I put little effort into anything

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

5. I approach life with intensity

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

6. Seeing a job through to the end is important to me

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

7. I spend time doing things that interest me

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

8. Getting together with friends is important to me

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

9. I have initiative

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

10. I have motivation

Strongly Disagree **Neutral** **Agree** **Strongly Agree**

NASA-TLX MENTAL WORKLOAD RATING SCALE

Please place an "X" along each scale at the point that best indicates your experience during alarm monitoring.

Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Low | High

Physical Demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low | High

Temporal Demand: How much time pressure did you feel due to the rate or pace at which the alarms/events occurred? Was the pace slow and leisurely or rapid and frantic?

Low | High

Performance: How successful do you think you were in accomplishing the task of alarm monitoring and responding to alarms? How satisfied were you with your performance in accomplishing these goals?

Low | High

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low | High

Frustration: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content and complacent did you feel during the shift?

Low | High

SWAT RATING

Participant ID	Measures	1	2	3
	Physical			
	Mental			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental Time			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental Time			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental Time			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental			
	Time			
	Physical			
	Mental Time			
	Time			

DISTRUST RATING

Participant ID	Measures	Yes	I don't know	No
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			
	Can you rely on this alarm?			
	Is it easy to make decision on this alarm?			
	Do you think this alarm can be false?			

APPENDIX E

DESCRIPTIVE STATISTICS FOR NURSES

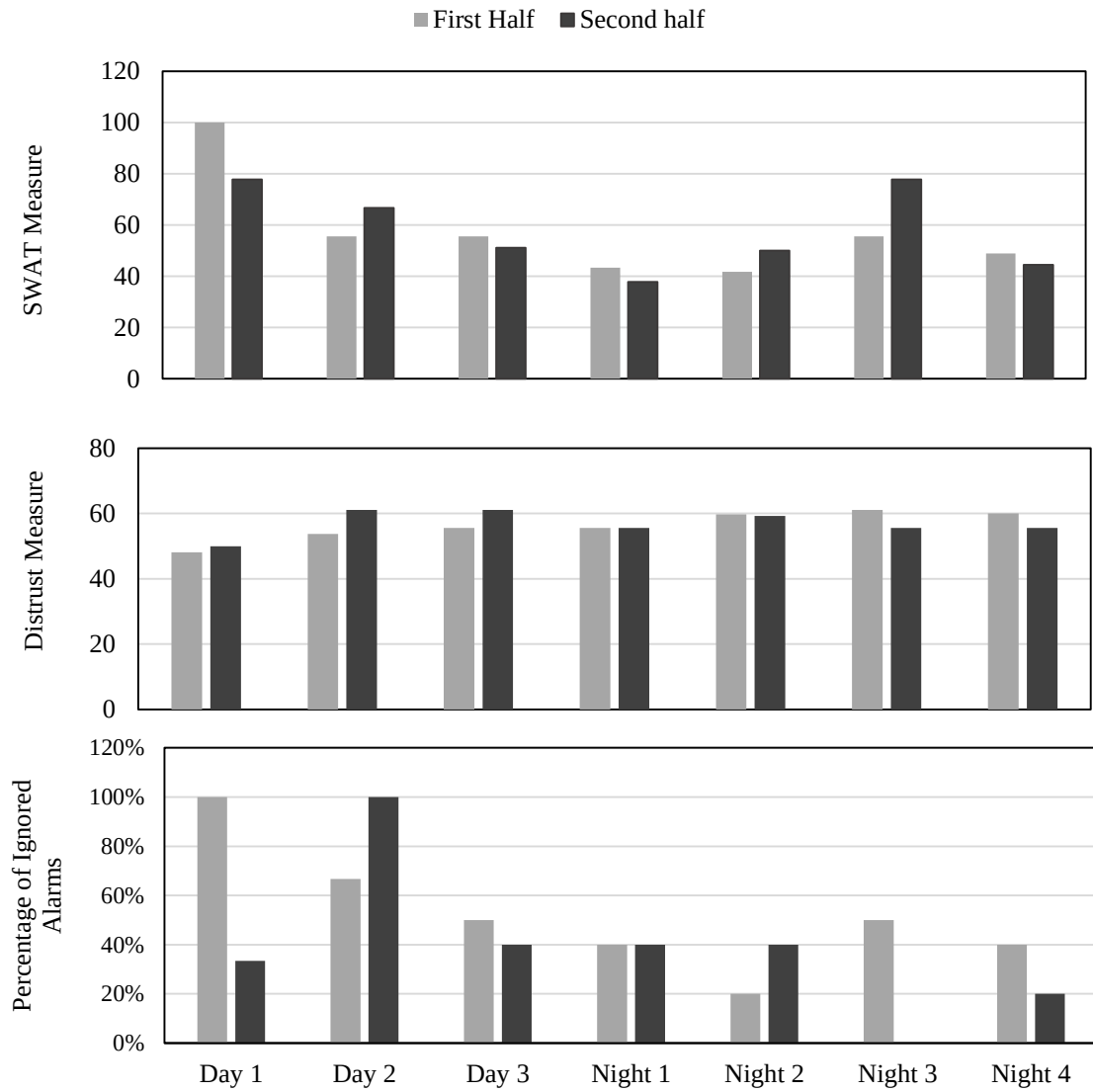


Figure E: Effect of alarm fatigue on alarm response for nurses

APPENDIX F

CORRELATIONAL ANALYSES

Table F1: Correlation analyses for nurse and unit clerk with all the factors of working conditions and staff individuality

	Unit Clerk													
	Shift	Staff ID	Number of Alarms	Ratio	Alarm Type	Alarm Criticality	Noise Level	Time Elapsed	Task Priority	Extraversion	Openness	Conscientiousness	Neuroticism	Agreeableness
Shift	1.00	.643	-.107	-.327	.031	.017	.001	-.118	-.117	-.277	-.001	-.618	.001	-.285
Staff ID	.643	1.000	-.156	-.277	.238	-.126	.010	.092	-.170	-.253	.193	-.289	.291	-.566
Number of Alarms	-.107	-.156	1.000	.309	-.039	-.029	.009	-.019	-.003	.150	.035	.165	.017	.260
Workload (Ratio)	-.327	-.277	.309	1.00	-.041	-.075	.049	-.105	-.012	.344	.106	.313	.011	.513
Alarm Type	.031	.238	-.039	-.041	1.000	-.369	.054	.002	-.046	-.108	.014	.002	.133	-.205
Alarm Criticality	.017	-.126	-.029	-.075	-.369	1.000	-.230	.032	-.099	.020	-.068	.068	.018	.084
Noise Level	.001	.010	.009	.049	.054	-.230	1.000	.008	.077	.071	.126	-.113	-.160	.027
Time Elapsed	-.118	.092	-.019	-.105	.002	.032	.008	1.000	-.110	-.098	-.043	.095	.164	-.069
Task Priority	-.117	-.170	-.003	-.012	-.046	-.099	.077	-.110	1.000	.109	.048	.044	-.120	.089
Extraversion	-.277	-.253	.150	.344	-.108	.020	.071	-.098	.109	1.000	.839	.207	-.580	.782
Openness	-.001	.193	.035	.106	.014	-.068	.126	-.043	.048	.839	1.000	-.071	-.626	.382
Conscientiousness	-.618	-.289	.165	.313	.002	.068	-.113	.095	.044	.207	-.071	1.000	.549	.172
Neuroticism	.001	.291	.017	.011	.133	.018	-.160	.164	-.120	-.580	-.626	.549	1.000	-.442
Agreeableness	-.285	-.566	.260	.513	-.205	.084	.027	-.069	.089	.782	.382	.172	-.442	1.000

Table F1: Correlation analyses for nurse and unit clerk with all the factors of working conditions and staff individuality (continued)

	Nurse													
	Shift	Staff ID	Number of Alarms	Ratio	Alarm Type	Alarm Criticality	Noise Level	Time Elapsed	Task Priority	Extraversion	Openness	Conscientiousness	Neuroticism	Agreeableness
Shift	1.00	-.075	-.403	-.29	.135	-.240	-.342	-.009	.118	.039	.067	-.036	-.130	.224
Staff ID	-.07	1.000	-.212	.108	.282	-.119	.054	-.158	-.368	-.207	.266	.150	-.160	-.137
Number of Alarms	-.40	-.212	1.00	.066	.330	-.162	.068	-.193	-.072	.324	-.094	-.015	.059	-.030
Ratio	-.28	.108	.066	1.00	.249	-.039	.374	.260	-.323	.147	.334	-.039	-.180	.337
Alarm Type	.135	.282	.330	.249	1.000	-.494	-.075	-.395	-.448	.382	.207	.137	-.349	.380
Alarm Criticality	-.24	-.119	-.162	-.04	-.494	1.000	.354	.227	.022	-.285	-.216	-.273	.371	-.387
Noise Level	-.34	.054	.068	.374	-.075	.354	1.000	.163	-.329	.013	-.079	-.235	.186	-.269
Time Elapsed	-.01	-.158	-.193	.260	-.395	.227	.163	1.000	.148	-.237	-.051	.036	.224	-.118
Task Priority	.118	-.368	-.072	-.32	-.448	.022	-.329	.148	1.000	-.056	-.109	-.197	.132	-.093
Extraversion	.039	-.207	.324	.147	.382	-.285	.013	-.237	-.056	1.000	.613	.273	-.690	.546
Openness	.067	.266	-.094	.334	.207	-.216	-.079	-.051	-.109	.613	1.000	.475	-.838	.630
Conscientiousness	-.04	.150	-.015	-.04	.137	-.273	-.235	.036	-.197	.273	.475	1.000	-.354	.474
Neuroticism	-.13	-.160	.059	-.18	-.349	.371	.186	.224	.132	-.690	-.838	-.354	1.000	-.624
Agreeableness	.224	-.137	-.030	.337	.380	-.387	-.269	-.118	-.093	.546	.630	.474	-.624	1.000

Table F2 Correlation analyses for nurse and unit clerk with all the alarm fatigue measures

	Unit Clerk				
SWAT Measures	1.000	.495	-.555	-.443	.311
NATA Measures	.495	1.000	-.001	-.112	.432
Boredom	-.555	-.001	1.000	.732	.342
Apathy	-.443	-.112	.732	1.000	-.019
Distrust	.311	.432	.342	-.019	1.000
	Nurse				
SWAT Measures	1.000	.443	-.395	-.090	.016
NATA Measures	.443	1.000	-.048	.389	-.144
Boredom	-.395	-.048	1.000	.486	-.085
Apathy	-.090	.389	.486	1.000	-.179
Distrust	.016	-.144	-.085	-.179	1.000

Table F3: Correlation Analyses between staff performances and working conditions with alarm fatigue measures for unit clerk

	SW	NT	B	AP	D	SH	ID	NA	WL	AT	AC	NL	TE	TP	E	O	C	N	AG
SW	1.00	.495	-.555	-.443	.311	.018	.008	.192	.245	.012	.065	-.137	.214	-.089	-.223	-.434	.533	.723	.093
NT	.495	1.000	-.001	-.112	.432	.004	-.187	.250	.432	-.118	.075	-.056	.012	.011	.417	.082	.294	.100	.690
B	-.555	-.001	1.000	.732	.342	.554	-.011	-.132	-.221	-.146	.067	.054	-.355	.070	.201	.231	-.712	-.653	.204
AP	-.443	-.112	.732	1.000	-.019	.554	.215	-.144	-.179	-.029	-.035	.066	-.131	-.047	-.249	-.117	-.904	-.385	-.073
D	.311	.432	.342	-.019	1.000	.487	-.139	.140	.038	-.156	.177	-.034	-.195	.025	.076	-.095	.022	.031	.298
SH	.018	.004	.554	.554	.487	1.000	.643	-.107	-.327	.031	.017	.001	-.118	-.117	-.277	-.001	-.618	.001	-.285
ID	.008	-.187	-.011	.215	-.139	.643	1.000	-.156	-.277	.238	-.126	.010	.092	-.170	-.253	.193	-.289	.291	-.566
NA	.192	.250	-.132	-.144	.140	-.107	-.156	1.000	.309	-.039	-.029	.009	-.019	-.003	.150	.035	.165	.017	.260
WL	.245	.432	-.221	-.179	.038	-.327	-.277	.309	1.000	-.041	-.075	.049	-.105	-.012	.344	.106	.313	.011	.513
AT	.012	-.118	-.146	-.029	-.156	.031	.238	-.039	-.041	1.000	-.369	.054	.002	-.046	-.108	.014	.002	.133	-.205
AC	.065	.075	.067	-.035	.177	.017	-.126	-.029	-.075	-.369	1.000	-.230	.032	-.099	.020	-.068	.068	.018	.084
NL	-.137	-.056	.054	.066	-.034	.001	.010	.009	.049	.054	-.230	1.000	.008	.077	.071	.126	-.113	-.160	.027
TE	.214	.012	-.355	-.131	-.195	-.118	.092	-.019	-.105	.002	.032	.008	1.000	-.110	-.098	-.043	.095	.164	-.069
TP	-.089	.011	.070	-.047	.025	-.117	-.170	-.003	-.012	-.046	-.099	.077	-.110	1.000	.109	.048	.044	-.120	.089
E	-.223	.417	.201	-.249	.076	-.277	-.253	.150	.344	-.108	.020	.071	-.098	.109	1.000	.839	.207	-.580	.782
O	-.434	.082	.231	-.117	-.095	-.001	.193	.035	.106	.014	-.068	.126	-.043	.048	.839	1.000	-.071	-.626	.382
C	.533	.294	-.712	-.904	.022	-.618	-.289	.165	.313	.002	.068	-.113	.095	.044	.207	-.071	1.000	.549	.172
N	.723	.100	-.653	-.385	.031	.001	.291	.017	.011	.133	.018	-.160	.164	-.120	-.580	-.626	.549	1.00	-.442
AG	.093	.690	.204	-.073	.298	-.285	-.566	.260	.513	-.205	.084	.027	-.069	.089	.782	.382	.172	-.442	1.000

SW: SWAT

ID: Staff ID

TE: Time Elapsed

AP: Apathy

AT: Alarm Type

O: Openness

NT: NASA TLX

NA: Number of Alarms

TP: Task Priority

D: Distrust

AC: Alarm Criticality

C: Conscientiousness

B: Boredom

WL: Workload

E: Extraversion

SH: Shift

NL: Noise Level

N: Neuroticism

Ag: Agreeableness

Table F4: Correlation Analyses between staff performances and combined factors from working conditions and alarm fatigue measures for nurse

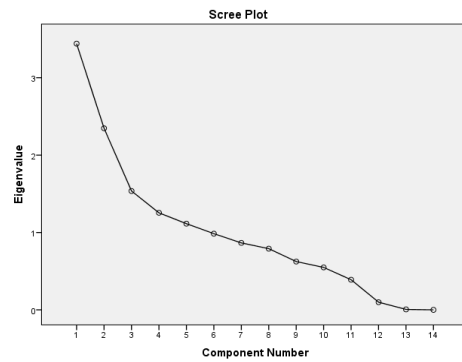
	SW	NT	B	AP	D	SH	ID	NA	WL	AT	AC	NL	TE	TP	E	O	C	N	AG
SW	1.000	.443	-.395	-.090	.016	-.546	-.092	.479	.087	.021	.340	.420	-.136	-.166	-.024	-.354	-.274	.418	-.352
NT	.443	1.00	-.048	.389	-.144	-.420	-.187	.415	-.011	-.098	.212	.095	-.164	.185	.467	.125	-.095	-.195	.043
B	-.395	-.048	1.00	.486	-.085	.373	-.259	-.359	.068	-.042	.003	-.112	.019	.089	.117	.191	-.251	-.405	.425
AP	-.090	.389	.486	1.00	-.179	-.080	-.443	-.059	.098	-.374	.192	.063	-.066	.330	.397	.457	.002	-.399	.457
D	.016	-.144	-.085	-.179	1.00	.166	.249	-.327	.067	.008	.080	.099	.021	-.034	-.111	.008	-.175	.026	-.113
SH	-.546	-.420	.373	-.080	.166	1.00	-.075	-.403	-.288	.135	-.240	-.342	-.009	.118	.039	.067	-.036	-.130	.224
ID	-.092	-.187	-.259	-.443	.249	-.075	1.00	-.212	.108	.282	-.119	.054	-.158	-.368	-.207	.266	.150	-.160	-.137
NA	.479	.415	-.359	-.059	-.327	-.403	-.212	1.00	.066	.330	-.162	.068	-.193	-.072	.324	-.094	-.015	.059	-.030
WL	.087	-.011	.068	.098	.067	-.288	.108	.066	1.00	.249	-.039	.374	.260	-.323	.147	.334	-.039	-.180	.337
AT	.021	-.098	-.042	-.374	.008	.135	.282	.330	.249	1.000	-.494	-.075	-.395	-.448	.382	.207	.137	-.349	.380
AC	.340	.212	.003	.192	.080	-.240	-.119	-.162	-.039	-.494	1.00	.354	.227	.022	-.285	-.216	-.273	.371	-.387
NL	.420	.095	-.112	.063	.099	-.342	.054	.068	.374	-.075	.354	1.00	.163	-.329	.013	-.079	-.235	.186	-.269
TE	-.136	-.164	.019	-.066	.021	-.009	-.158	-.193	.260	-.395	.227	.163	1.000	.148	-.237	-.051	.036	.224	-.118
TP	-.166	.185	.089	.330	-.034	.118	-.368	-.072	-.323	-.448	.022	-.329	.148	1.000	-.056	-.109	-.197	.132	-.093
E	-.024	.467	.117	.397	-.111	.039	-.207	.324	.147	.382	-.285	.013	-.237	-.056	1.00	.613	.273	-.690	.546
O	-.354	.125	.191	.457	.008	.067	.266	-.094	.334	.207	-.216	-.079	-.051	-.109	.613	1.000	.475	-.838	.630
C	-.274	-.095	-.251	.002	-.175	-.036	.150	-.015	-.039	.137	-.273	-.235	.036	-.197	.273	.475	1.000	-.354	.474
N	.418	-.195	-.405	-.399	.026	-.130	-.160	.059	-.180	-.349	.371	.186	.224	.132	-.690	-.838	-.354	1.000	-.624
AG	-.352	.043	.425	.457	-.113	.224	-.137	-.030	.337	.380	-.387	-.269	-.118	-.093	.546	.630	.474	-.624	1.00

APPENDIX G

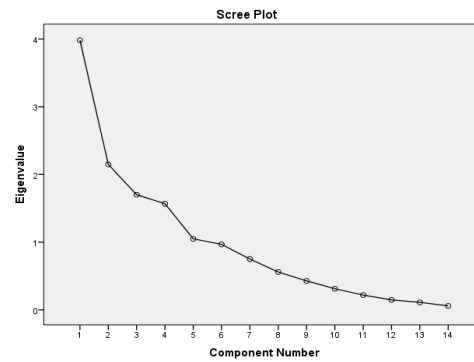
PRINCIPAL COMPONENT ANALYSES (PCA)

Table G1: Initial eigenvalues from PCA with the variables from working conditions

Component	Unit Clerk			Nurse		
	Eigenvalue	% of Variance	Cumulative %	Eigenvalue	% of Variance	Cumulative %
1	3.441	24.578	24.578	3.982	28.444	28.444
2	2.347	16.764	41.342	2.149	15.351	43.795
3	1.536	10.972	52.315	1.700	12.144	55.939
4	1.254	8.957	61.272	1.568	11.203	67.142
5	1.114	7.957	69.229	1.048	7.489	74.631
6	.985	7.034	76.263	.967	6.910	81.541
7	.866	6.182	82.445	.751	5.362	86.903
8	.791	5.650	88.095	.559	3.996	90.898
9	.624	4.459	92.555	.427	3.051	93.949
10	.548	3.918	96.472	.312	2.228	96.178
11	.389	2.781	99.253	.218	1.557	97.735
12	.099	.706	99.959	.147	1.051	98.786
13	.006	.041	100.000	.111	.795	99.581
14	-1.07E-013	-1.51E-013	100.000	.059	.419	100.000



(a) Unit Clerk



(b) Nurse

Figure G1: Extraction of components from PCA for working conditions

Table G2: Component matrix for unit clerk from PCA with working conditions and staff individuality

	Component				
	1	2	3	4	5
Openness	.956	-.025	.144	.261	-.073
Extraversion	.921	.168	-.028	-.006	.166
Neuroticism	-.738	.422	.086	.269	.126
Agreeableness	.611	.116	-.206	-.210	.403
Conscientiousness	-.113	.831	.009	-.053	.231
Shift	-.068	-.730	-.005	.488	-.062
Time Elapsed	.093	.562	-.018	.229	.462
Alarm Criticality	.016	.011	-.820	.096	-.120
Alarm Type	-.068	.126	.713	.199	.028
Noise Level	.131	-.092	.467	-.251	-.051
Staff ID	-.017	-.252	.304	.742	-.205
Task Priority	-.018	-.027	.164	-.649	.147
Number of Alarms	.000	-.033	.025	.143	.733
Workload	.158	.203	.083	-.007	.717

Table G3: Component matrix for nurses from PCA with working conditions and staff individuality

	Component				
	1	2	3	4	5
Openness	.926	.132	.113	-.102	.021
Neuroticism	-.825	-.048	.136	.011	-.091
Conscientiousness	.740	-.289	.113	-.249	-.200
Agreeableness	.725	.157	-.142	.225	.250
Extraversion	.659	.131	-.275	.410	-.164
Workload	.220	.850	.053	.007	.007
Noise Level	-.201	.676	.163	-.084	-.289
Alarm Type	.067	.267	.843	-.140	.124
Time Elapsed	.041	.316	.692	.140	.147
Alarm Criticality	-.221	.124	.589	-.032	-.259
Staff ID	.072	.014	-.138	.892	.018
Task Priority	-.008	-.467	.345	.533	.089
Shift	-.065	-.125	-.153	.083	.875
Number of Alarms	-.069	.032	-.511	.383	-.643

Table G4: Eigenvalues from PCA with alarm fatigue measures

Component	Unit Clerk			Nurse		
	Total Eigenvalues	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.263	45.269	45.269	1.714	34.284	34.284
2	1.662	33.247	78.517	1.593	31.851	66.135
3	.584	11.688	90.204	.916	18.312	84.447
4	.399	7.978	98.183	.437	8.741	93.188
5	.091	1.817	100.000	.341	6.812	100.000

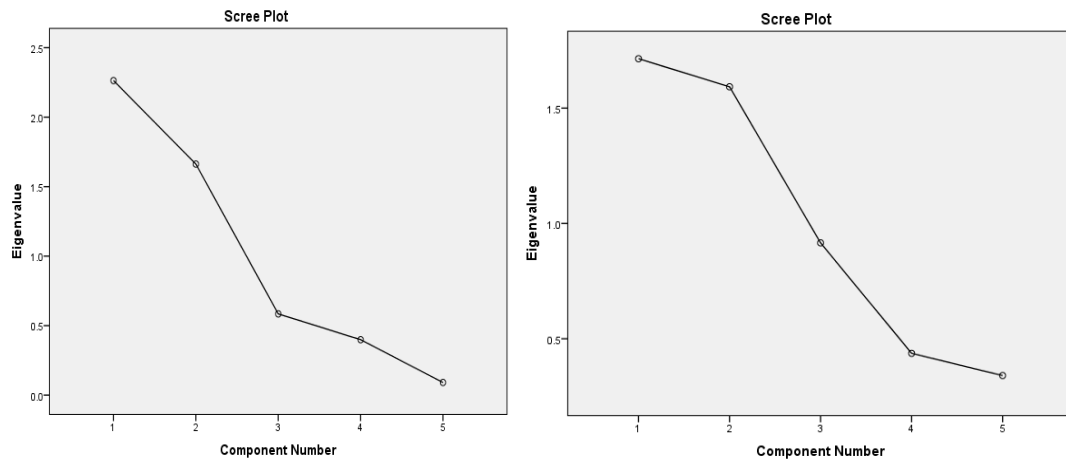


Figure G2: Extraction of components from PCA with alarm fatigue measures

Table G5: Component matrices from PCA with alarm fatigue measures

	Components			
	Unit Clerk		Nurse	
	1	2	1	2
Boredom	0.964	0.221	0.881	-0.085
Apathy	0.858	-0.019	-0.437	-0.064
SWAT	-0.656	0.554	-0.033	0.879
Distrust	0.230	0.854	0.583	0.682
NASA TLX	-0.108	0.819	0.612	-0.613

Table G6: Eigenvalues from PCA with combined factors

Component	Unit Clerk			Nurse		
	Total Eigenvalues	% of Variance	Cumulative %	Total Eigenvalues	% of Variance	Cumulative %
1	4.185	22.025	22.025	4.350	22.897	22.897
2	3.745	19.713	41.738	2.944	15.493	38.390
3	2.297	12.090	53.829	2.708	14.252	52.641
4	1.605	8.448	62.276	1.964	10.335	62.977
5	1.340	7.052	69.328	1.391	7.320	70.297
6	1.102	5.801	75.129	1.260	6.631	76.928
7	.982	5.170	80.299	.914	4.812	81.741
8	.853	4.488	84.787	.749	3.940	85.681
9	.814	4.286	89.074	.610	3.212	88.892
10	.721	3.793	92.867	.520	2.736	91.628
11	.553	2.910	95.776	.368	1.939	93.566
12	.395	2.077	97.853	.313	1.648	95.214
13	.209	1.099	98.952	.237	1.246	96.460
14	.121	.638	99.590	.235	1.236	97.696
15	.069	.362	99.952	.204	1.073	98.769
16	.007	.036	99.988	.091	.477	99.245
17	.002	.012	100.000	.079	.414	99.660
18	3.637E-05	.000	100.000	.034	.180	99.840
19	-7.676E-15	-4.0E-14	100.000	.030	.160	100.000

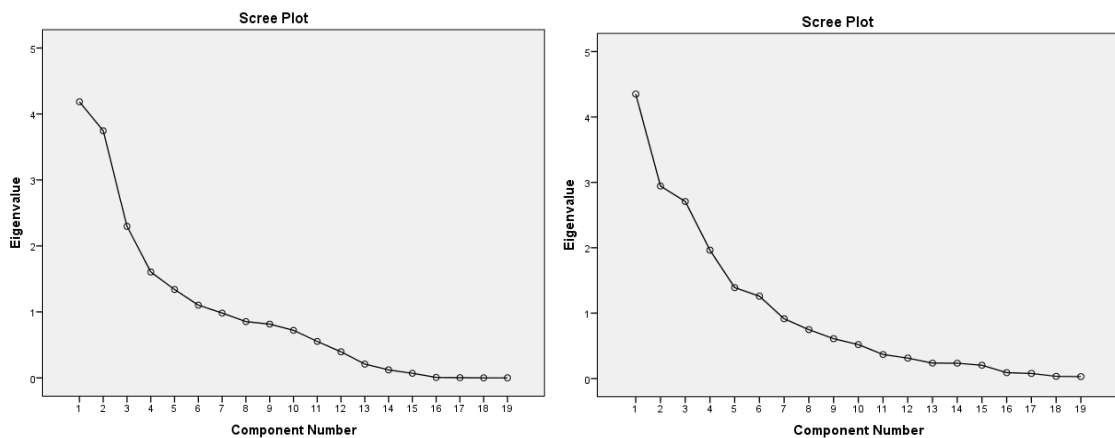


Figure G3: Extraction of components from PCA with combined factors

Table G7: Component matrices from PCA with combined factors for unit clerks

	Component					
	1	2	3	4	5	6
Shift	-.635	-.205	.610	.352	.132	.114
Staff ID	-.379	-.476	.044	.609	.361	.080
Number of Alarms	.315	.222	.172	.165	-.219	-.268
Workload	.488	.406	.070	.184	-.272	-.296
Alarm Type	-.031	-.252	-.257	.550	-.251	.090
Alarm Criticality	.066	.061	.313	-.535	.533	-.006
Noise Level	-.125	.122	-.204	.270	-.360	-.089
Time Elapsed	.206	-.240	-.149	.071	.366	-.472
Task Priority	.012	.184	-.123	-.147	-.362	.568
Extraversion	.168	.889	-.179	.225	.269	.116
Openness	-.165	.654	-.351	.430	.438	.154
Neuroticism	.531	-.753	.252	.108	-.017	.083
Agreeableness	.313	.868	.206	.015	-.040	-.144
SWAT	.689	-.327	.507	.183	-.012	-.026
NASA TLX	.446	.440	.561	.256	.004	-.093
Boredom	-.779	.447	.314	-.093	-.065	.069
Apathy	-.820	.037	.250	-.058	-.218	-.350
Distrust	0.032	0.215	0.798	0.051	-0.04	0.307

Table G8: Component matrices from PCA with combined factors for nurses

	Component					
	1	2	3	4	5	6
Openness	.912	-.043	.064	.124	-.271	.198
Neuroticism	-.899	.067	.156	.021	.045	-.154
Extraversion	.745	.297	-.258	-.013	.050	-.151
Agreeableness	.713	-.271	-.222	.145	-.010	-.254
Apathy	.698	.144	.484	-.017	.279	-.206
NASA TLX	.399	.780	.176	-.168	.104	-.072
SWAT	-.292	.771	-.077	.184	.138	-.010
Shift	.011	-.700	-.210	-.283	.305	.060
Number of Alarms	-.099	.616	-.439	.019	-.067	-.475
Alarm Type	.112	-.014	-.912	.166	.051	.035
Alarm Criticality	-.131	.291	.632	.161	.173	.198
Time Elapsed	-.191	-.348	.515	.445	-.227	-.283
Workload	.236	-.034	-.098	.853	.042	-.076
Noise Level	-.074	.327	.136	.661	.195	.160
Task Priority	.005	-.033	.458	-.544	.099	-.248
Conscientiousness	.386	-.095	.021	-.068	-.805	-.092
Boredom	.382	-.464	.066	.092	.636	-.115
Staff ID	.046	-.008	-.165	.060	-.391	.738
Distrust	.003	-.079	.056	.018	.178	.696

APPENDIX H

SIGNIFICANCE OF INDIVIDUAL FACTORS

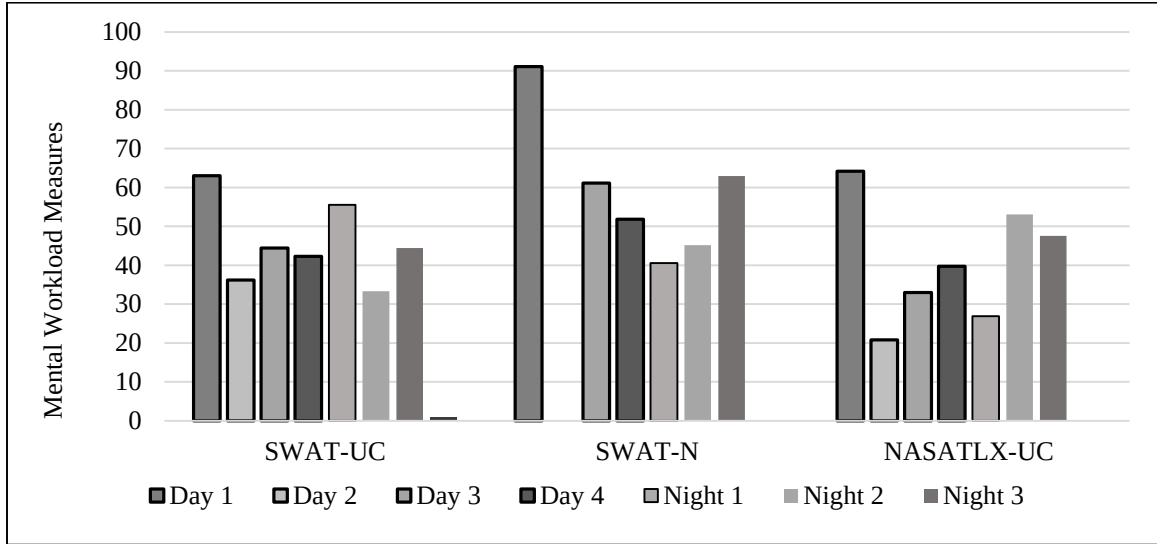


Figure H1: Influence of shift on mental workload measures for both unit clerk and nurse

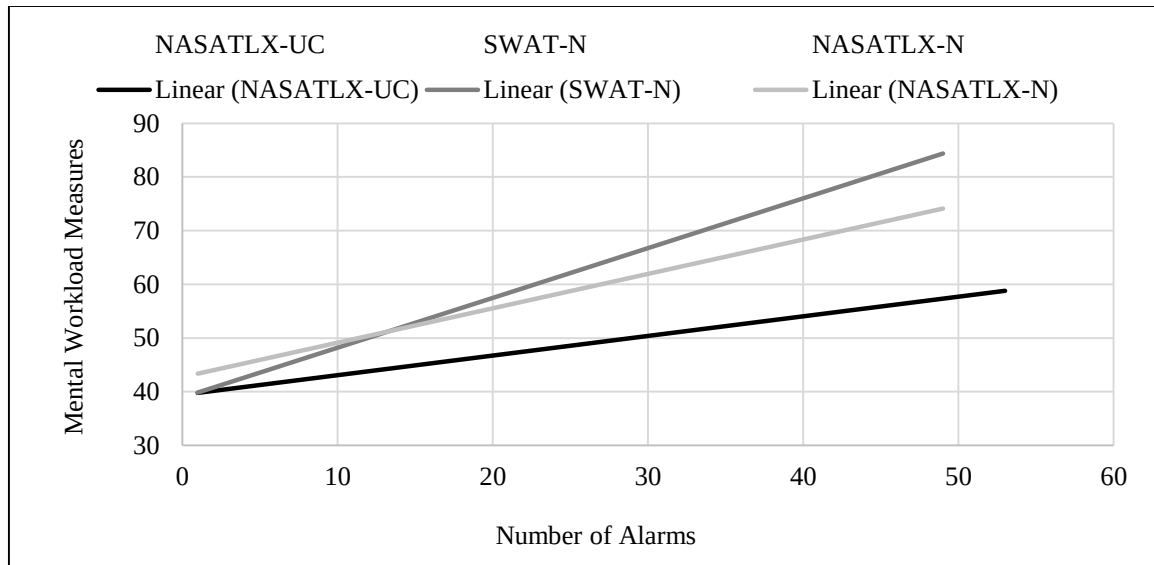


Figure H2: Change in mental workload measures for different number of alarms

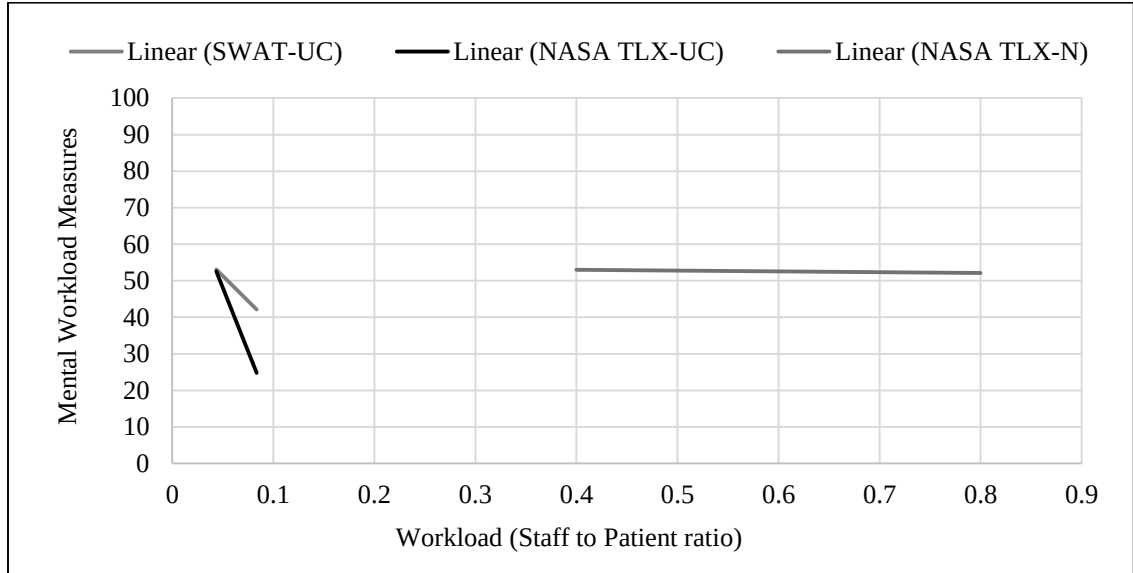


Figure H3: Change in mental workload measures for different workloads (staff to patient ratio)

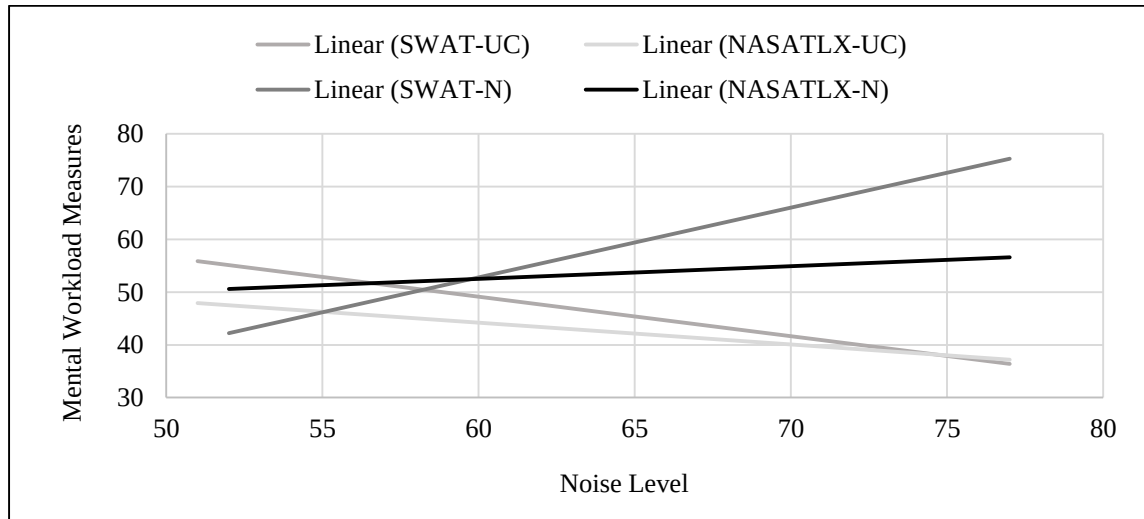


Figure H4: Change in mental workload scores for different noise levels

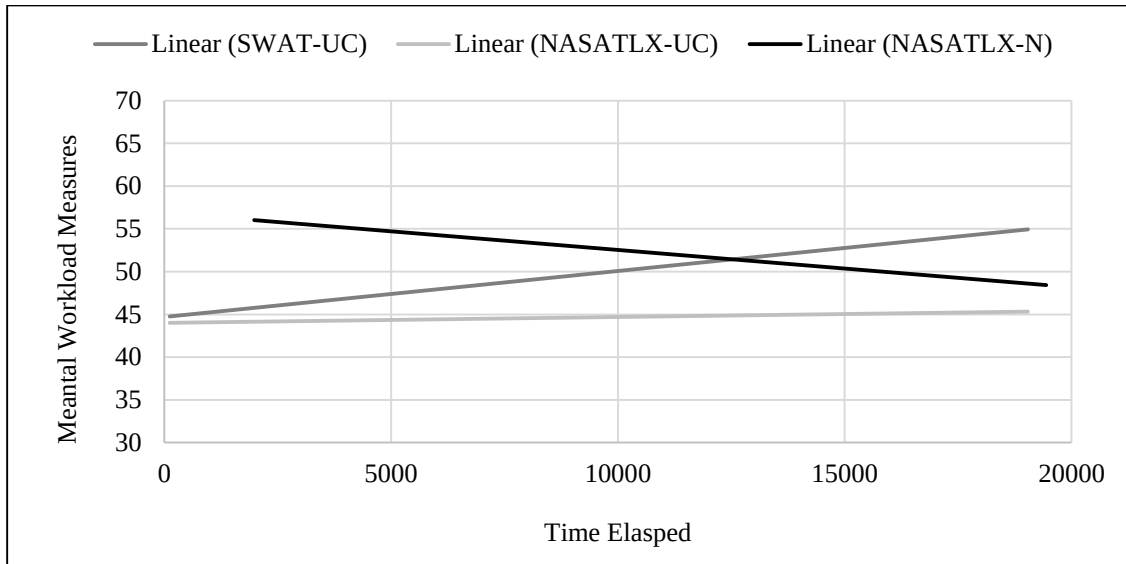


Figure H5: Change in mental workload measures with time elapsed since the start of the shift

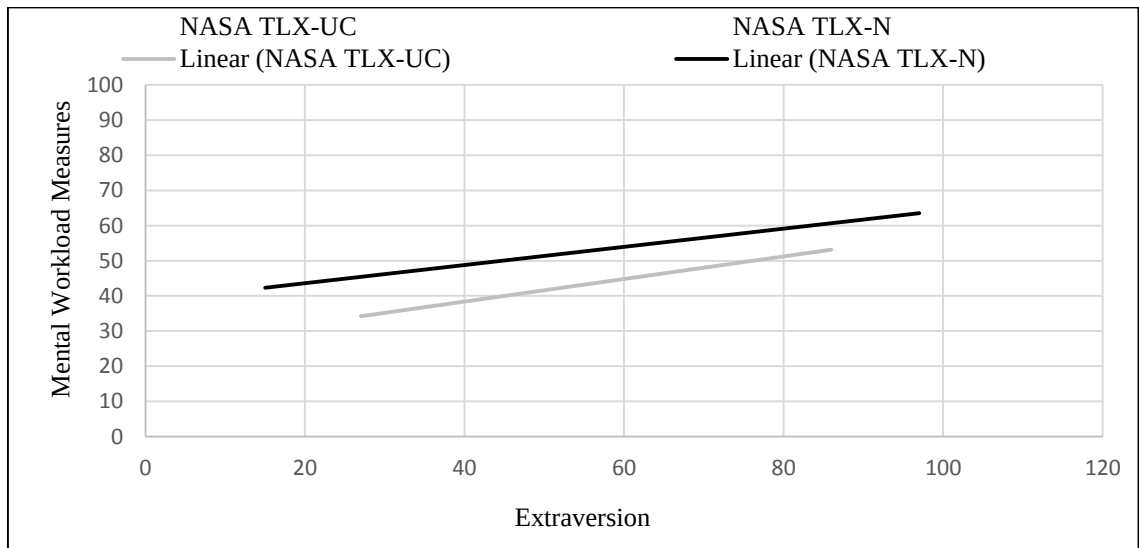


Figure H6: Extraversion vs NASA TLX score for both unit clerk and nurse

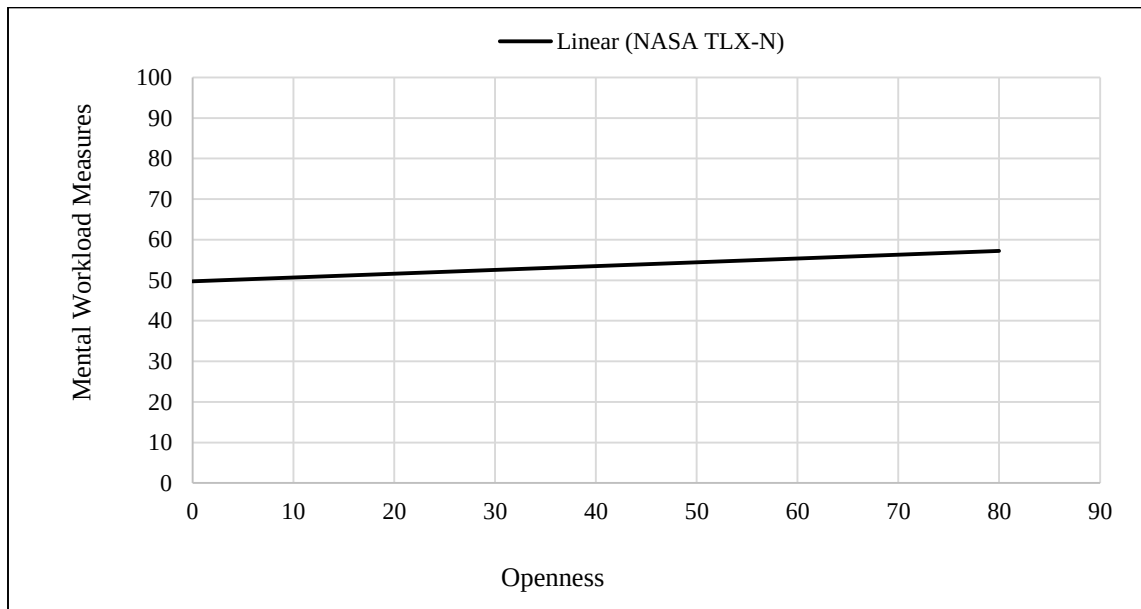


Figure H7: Change in mental workload measures for different percentiles of openness

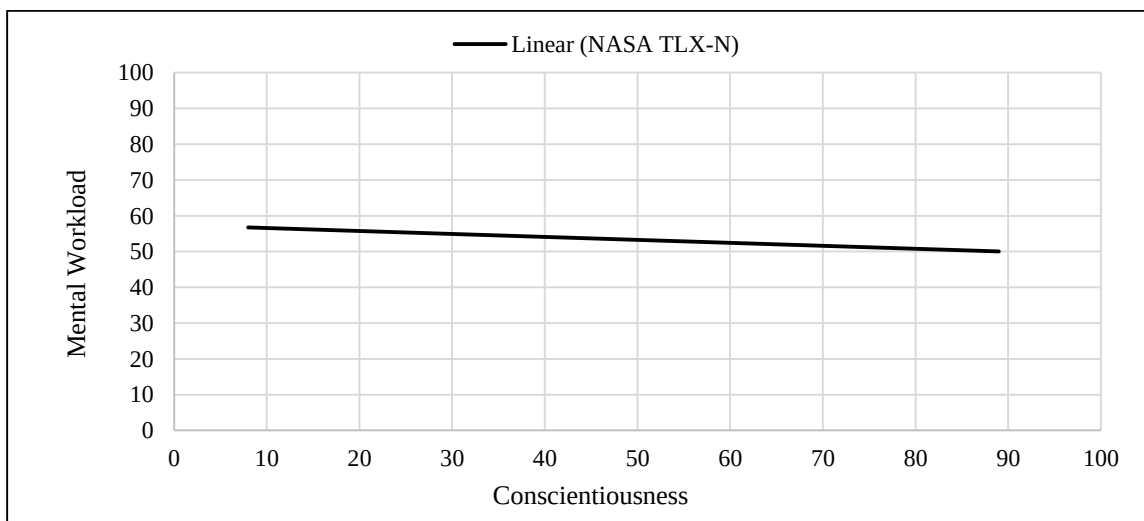


Figure H8: Changes in mental workload measures for different percentiles of conscientiousness

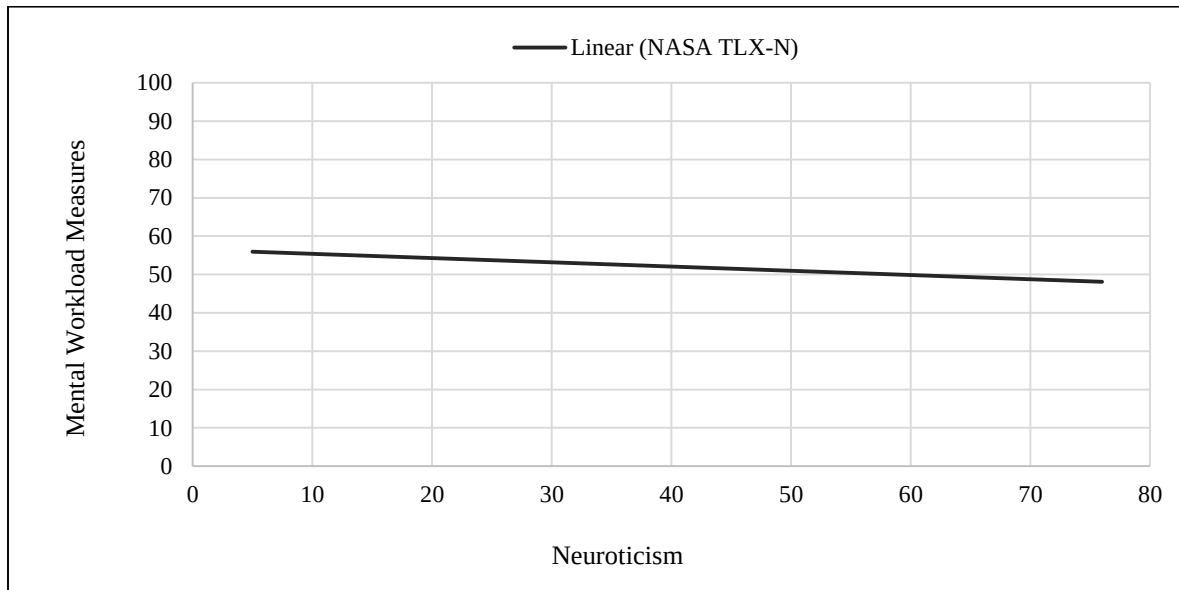


Figure H9: Changes in mental workload measures for different percentiles of neuroticism

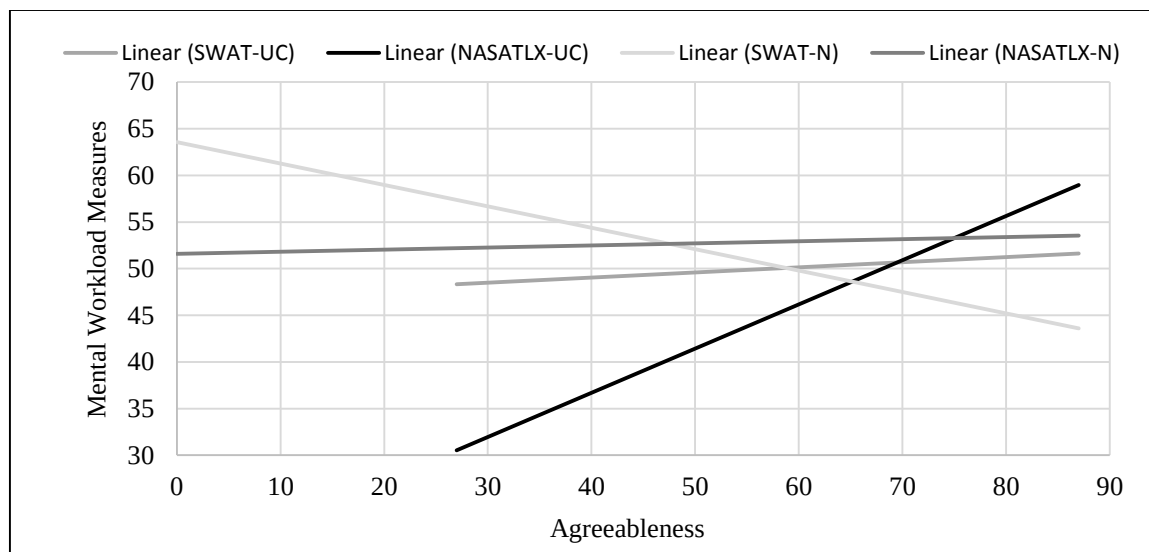


Figure H10: Change in mental workload scores for different scores of agreeableness

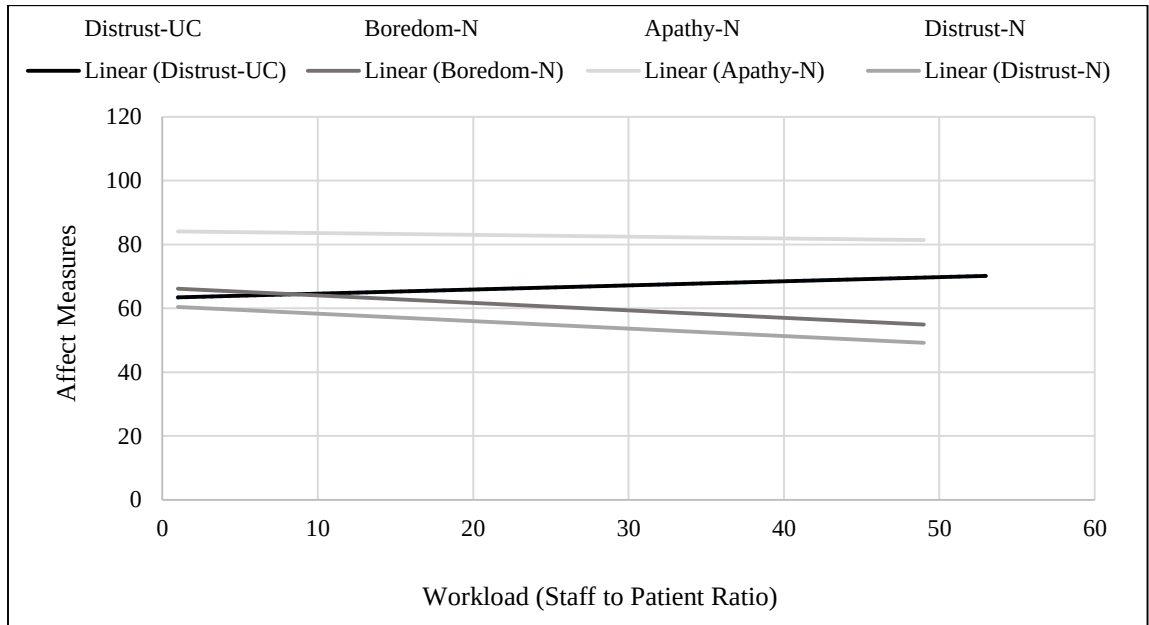


Figure H11: Change in affect measures for different levels of workload

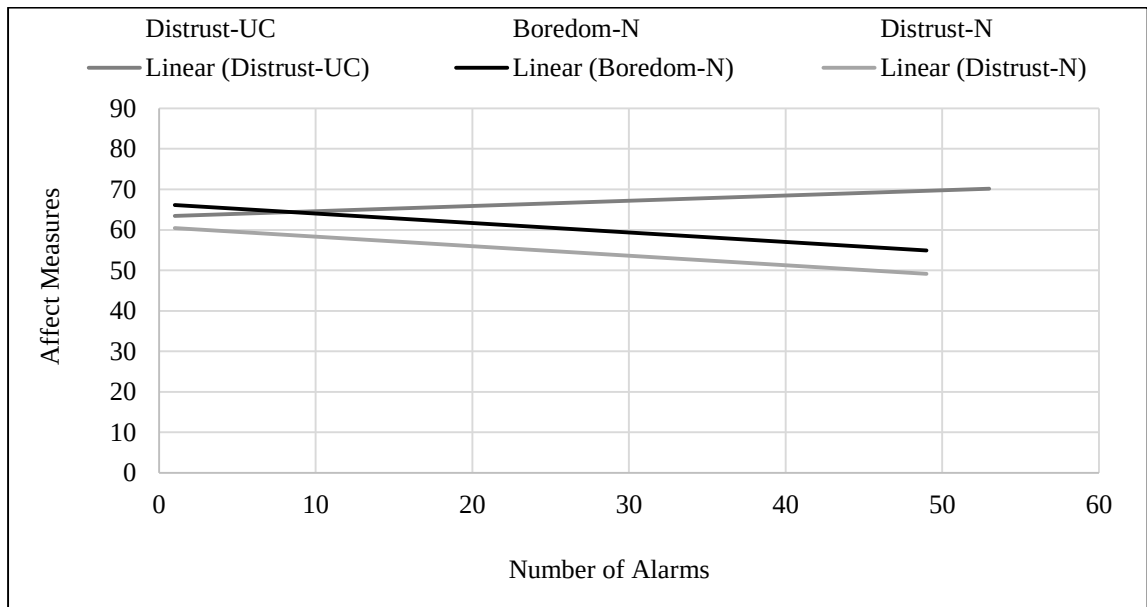


Figure H12: Change in affect measures for the change in number of alarms

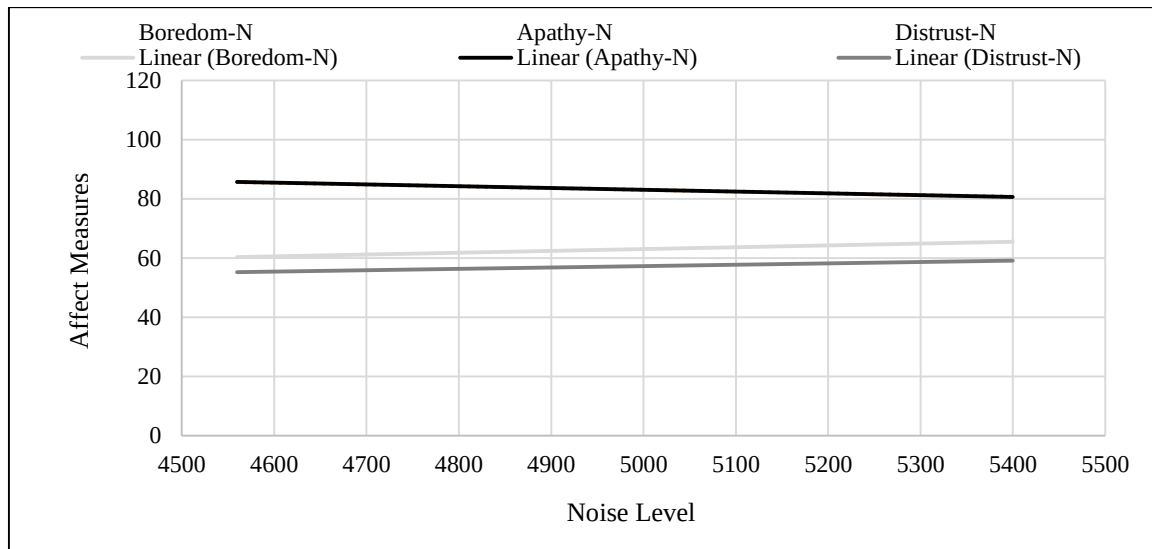


Figure H13: Changes in affect measures for different noise levels

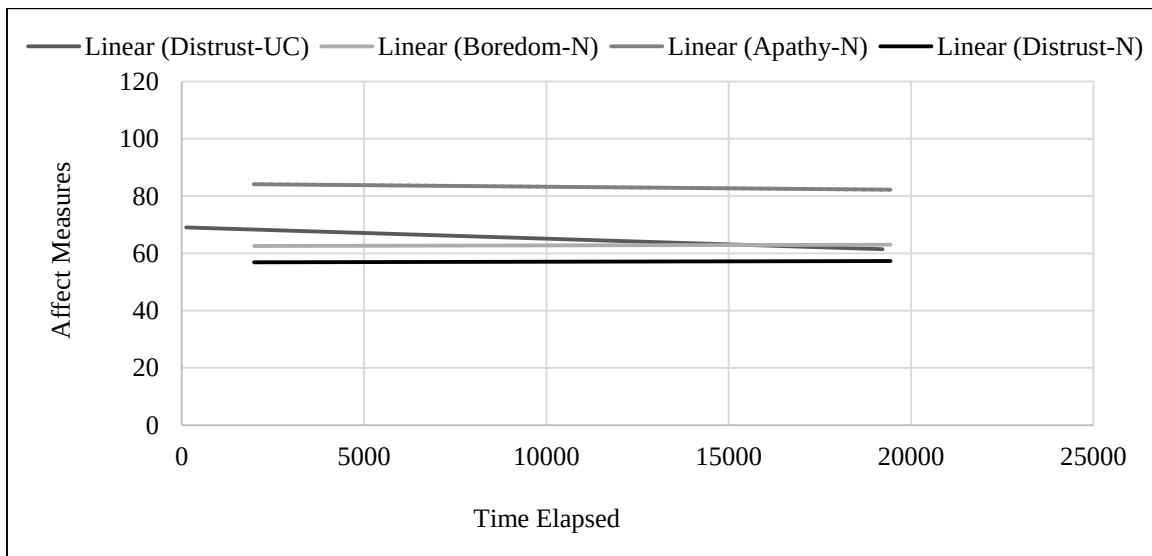


Figure H14: Changes in affect measures with time elapsed since the start of the shift

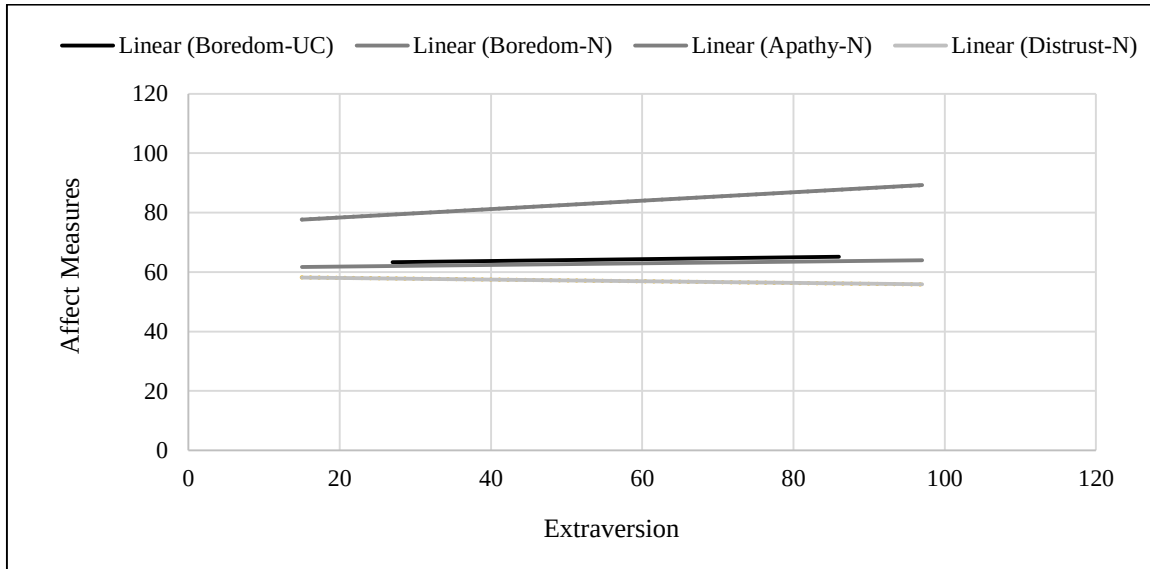


Figure H15: Change in affect measures for different percentiles of extraversion

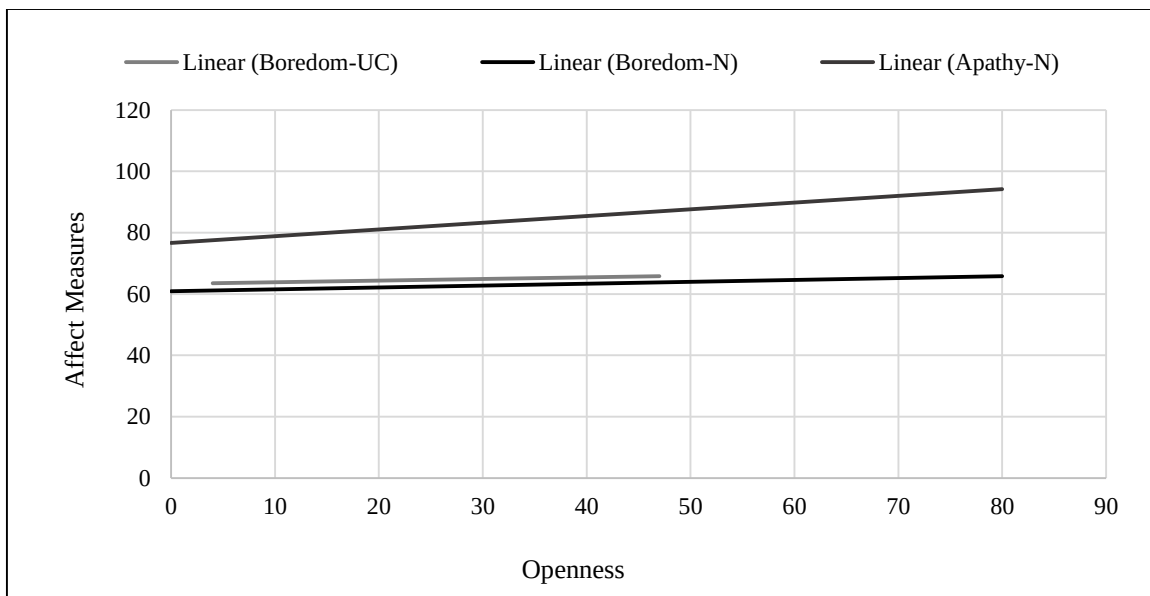


Figure H16: Change in affect measures for different percentiles of openness

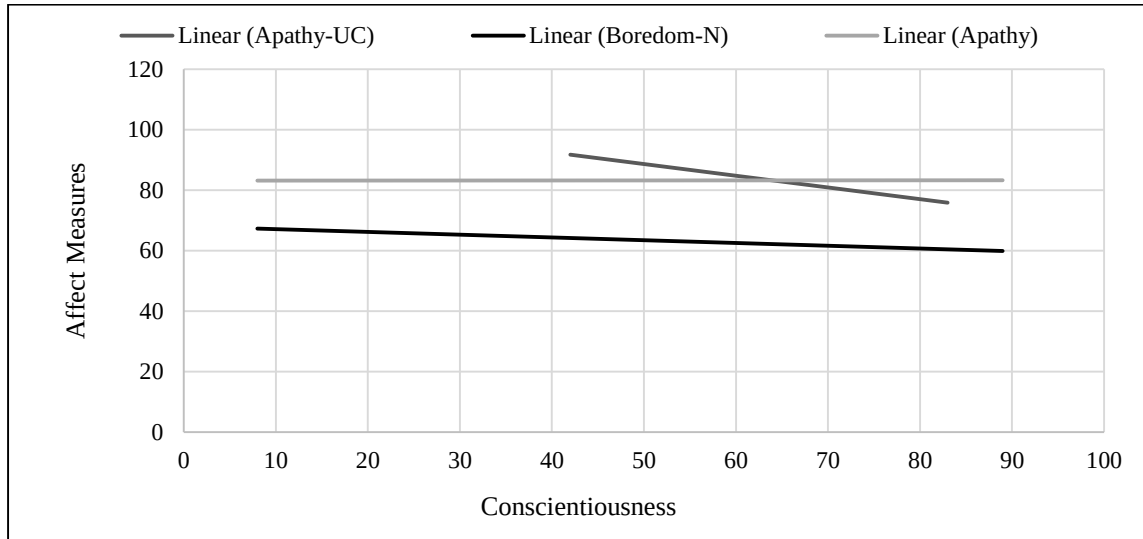


Figure H17: Change in affect measures for different percentiles of conscientiousness

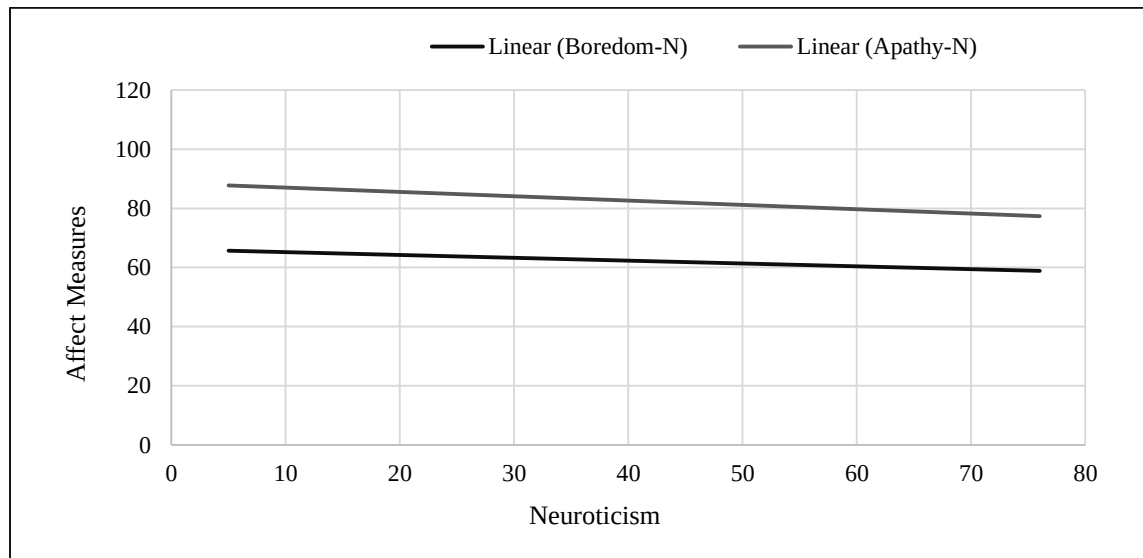


Figure H18: Change in affect measures for different percentiles of neuroticism

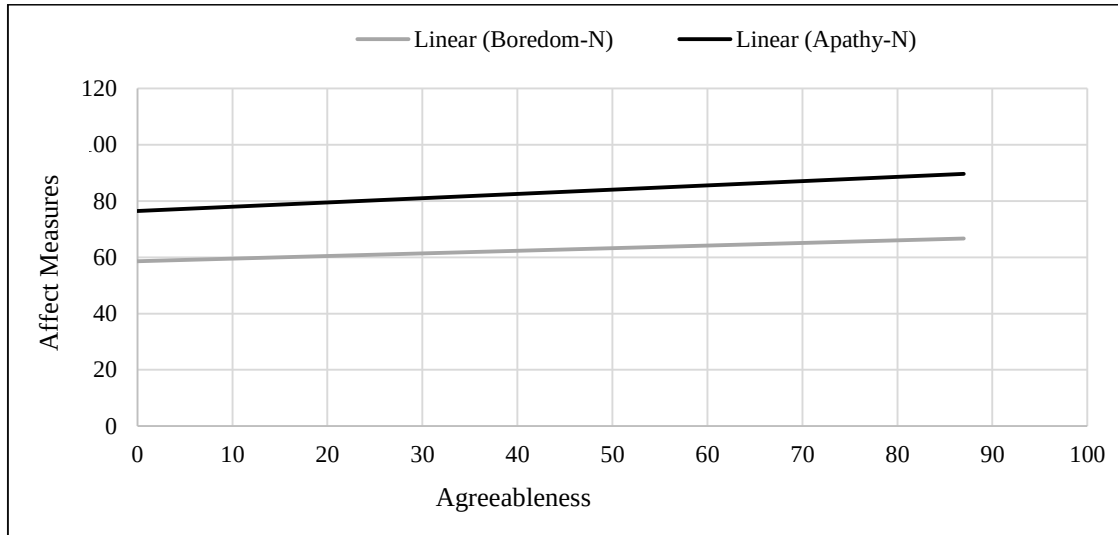


Figure H19: Change in affect measures for different percentiles of agreeableness

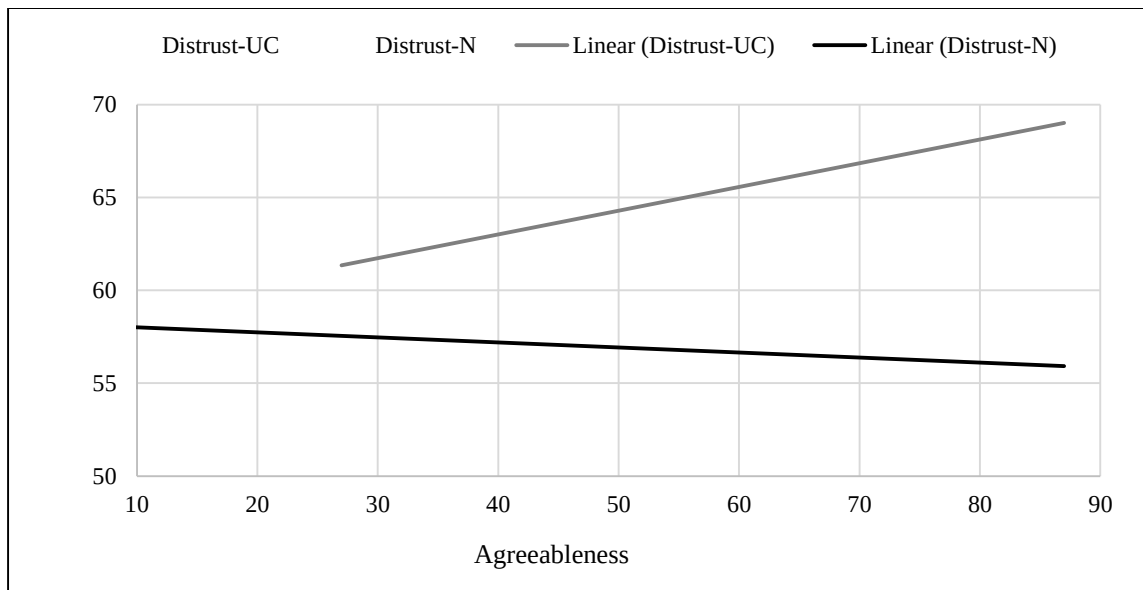


Figure H20: Change in affect measures for different percentiles of agreeableness

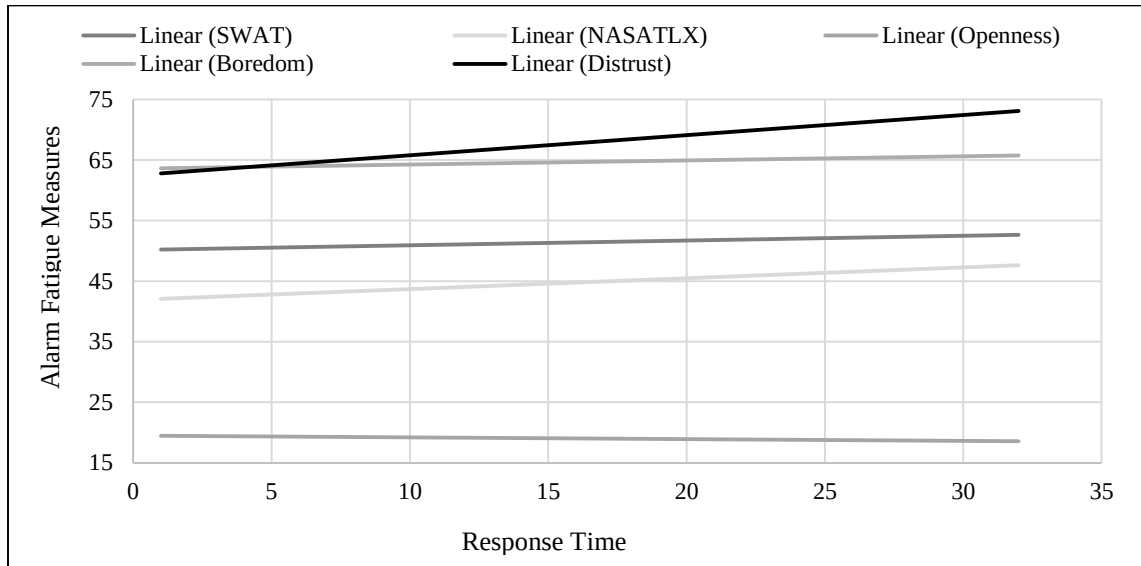


Figure H21: Change in mental workload and affect for different response time for unit clerks

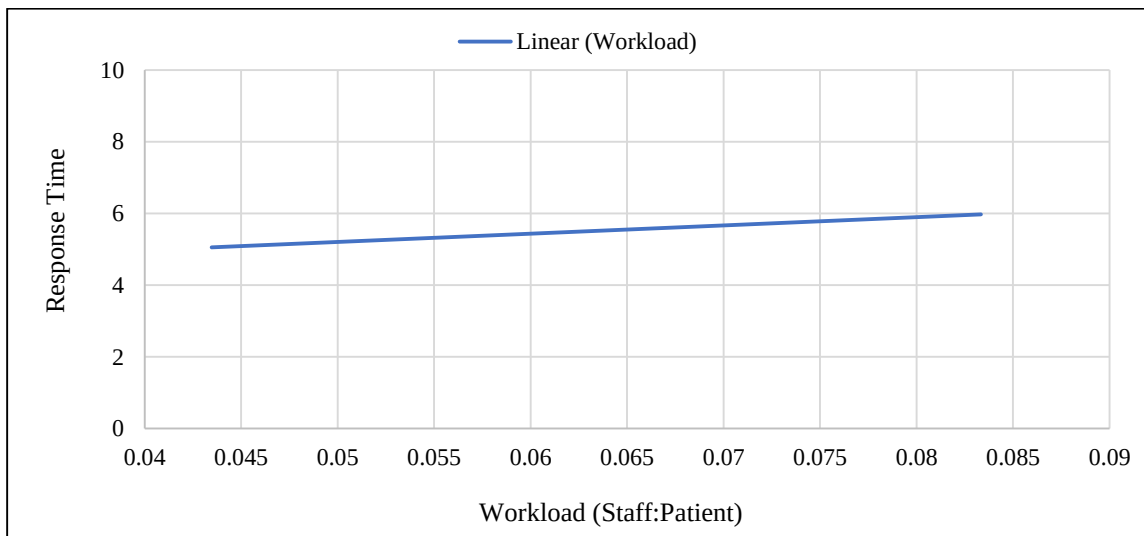


Figure H22: Workload vs response time for unit clerks

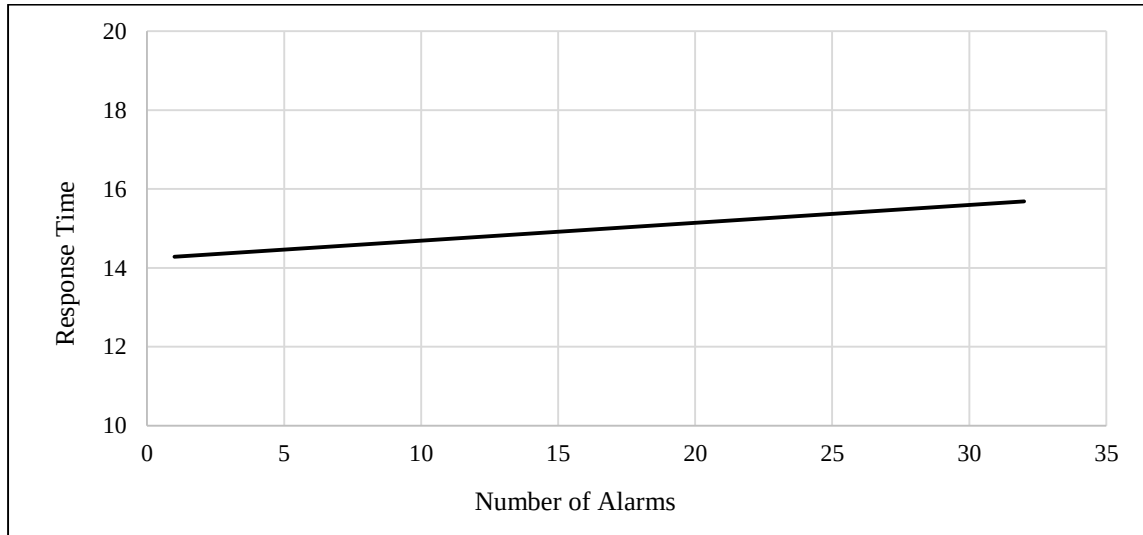


Figure H23: Number of alarms vs response time for unit clerk

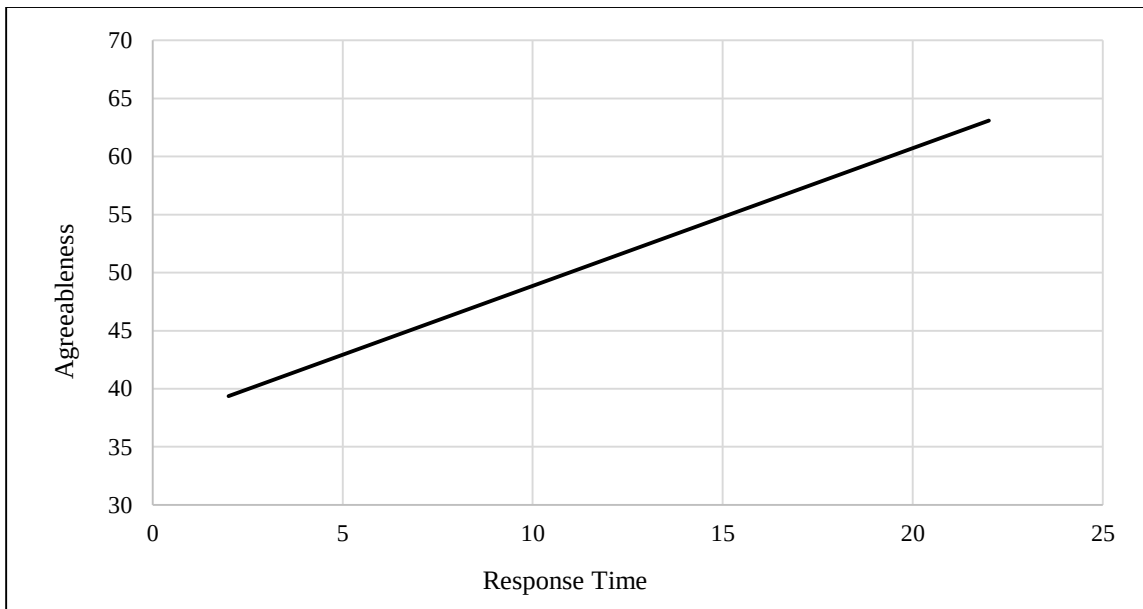


Figure H24: Response time vs agreeableness for nurses

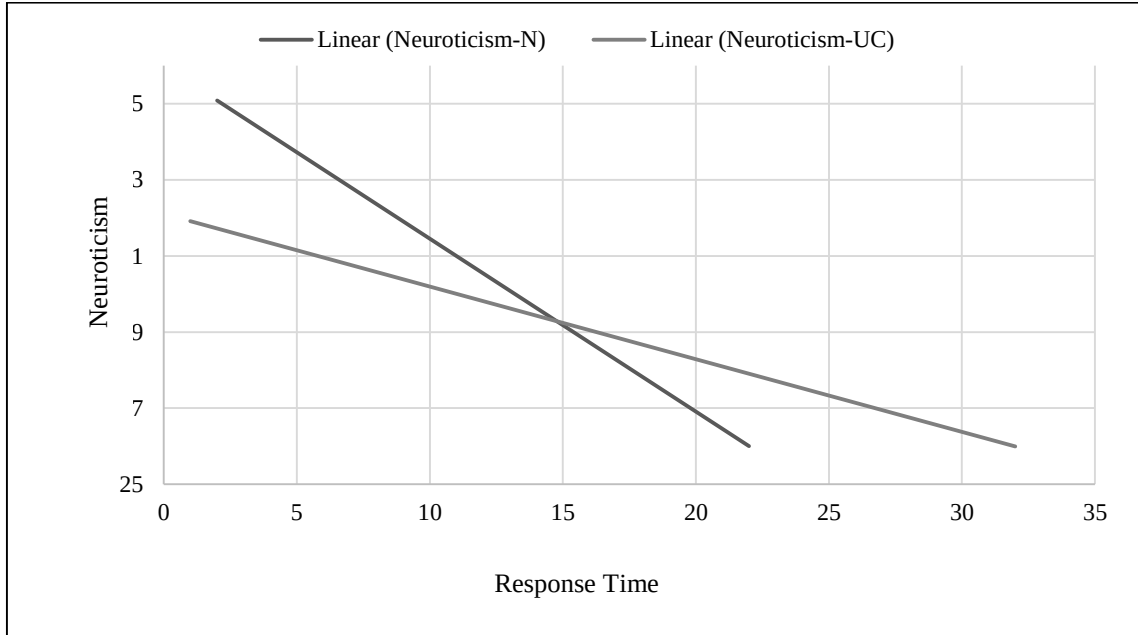


Figure H25: Neuroticism vs response time

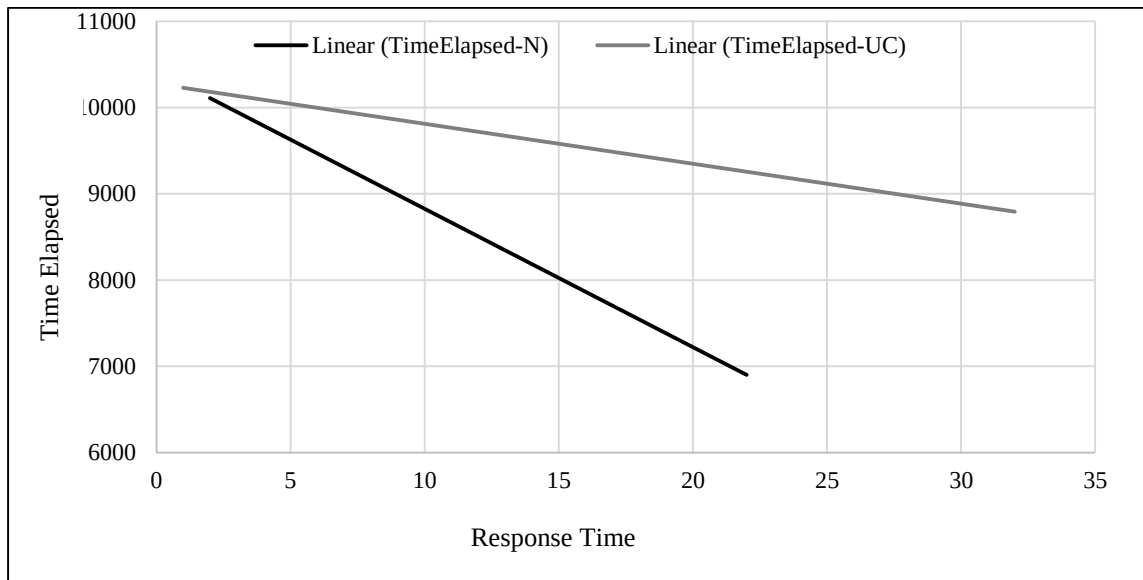


Figure H26: Time elapsed vs response time

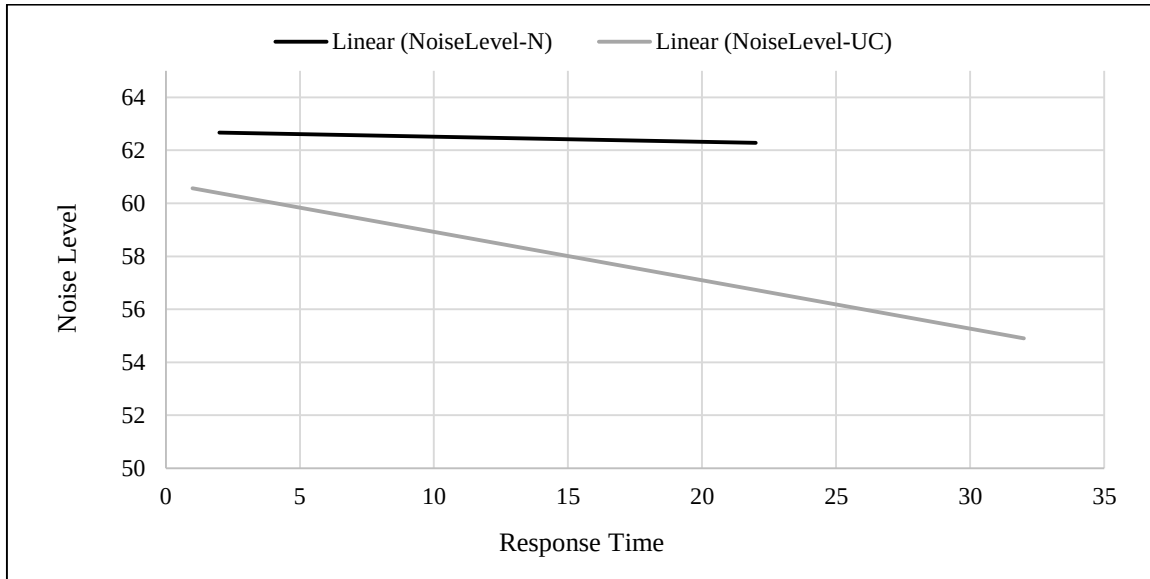


Figure H27: Noise level vs response time

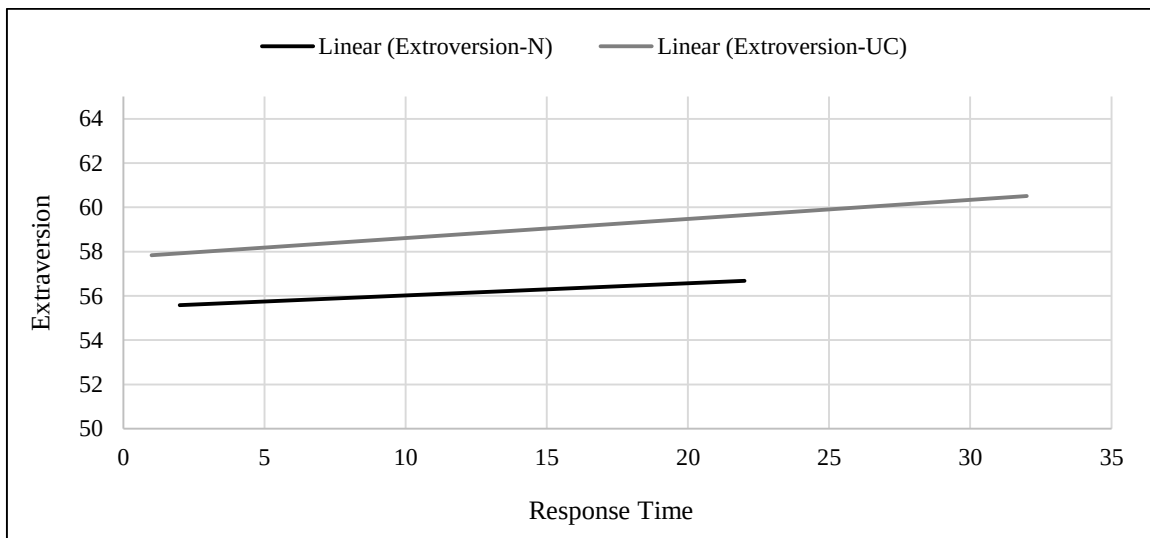


Figure H28: Extraversion vs response time

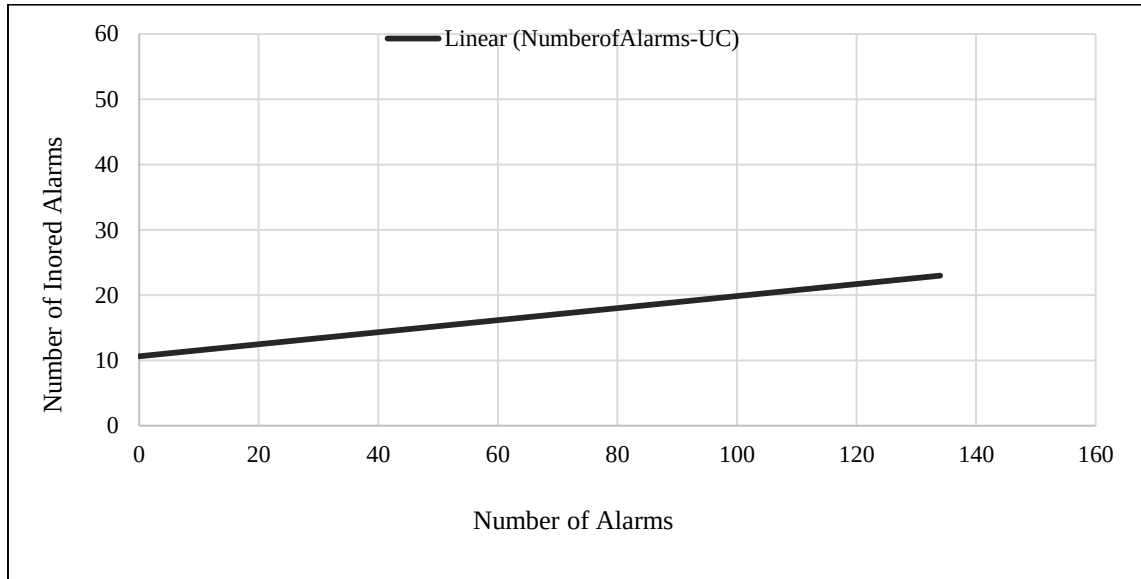


Figure H29: Number of ignored alarms vs number of alarms

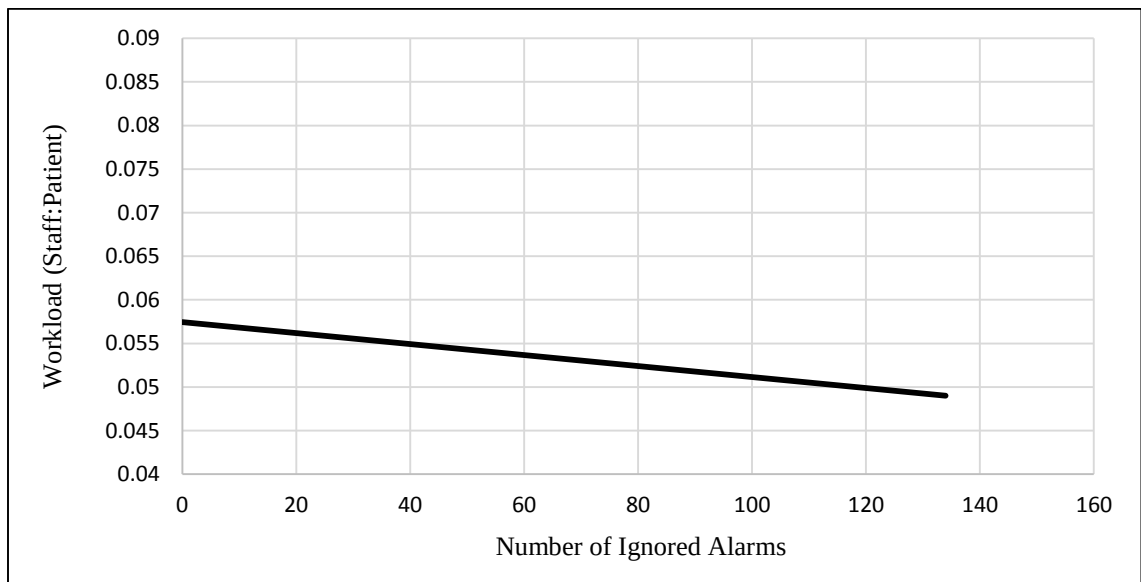


Figure H30: Number of ignored alarms vs workload for unit clerks

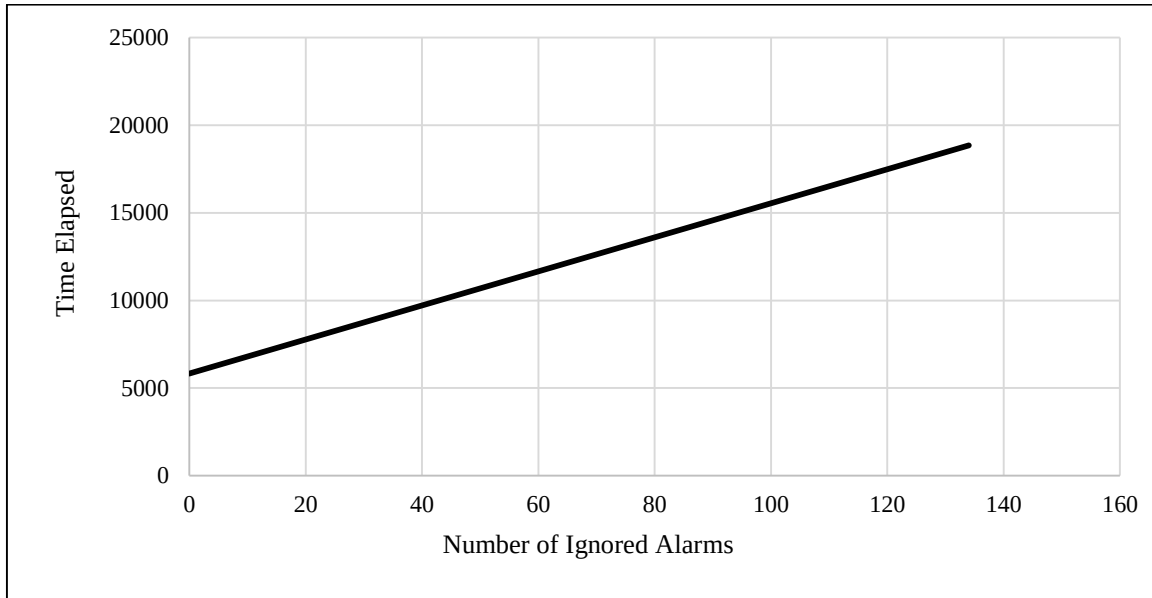


Figure H31: Number of ignored alarms vs time elapsed for unit clerks

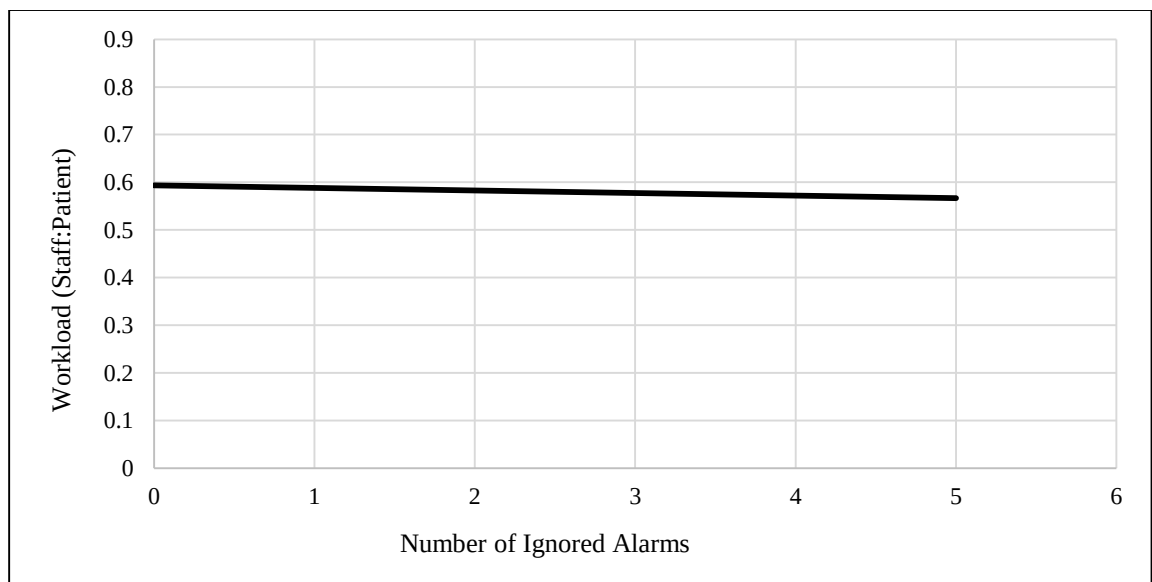


Figure H32: Number of ignored alarms vs workload for nurses



Figure H33: Number of ignored alarms vs time elapsed for nurses

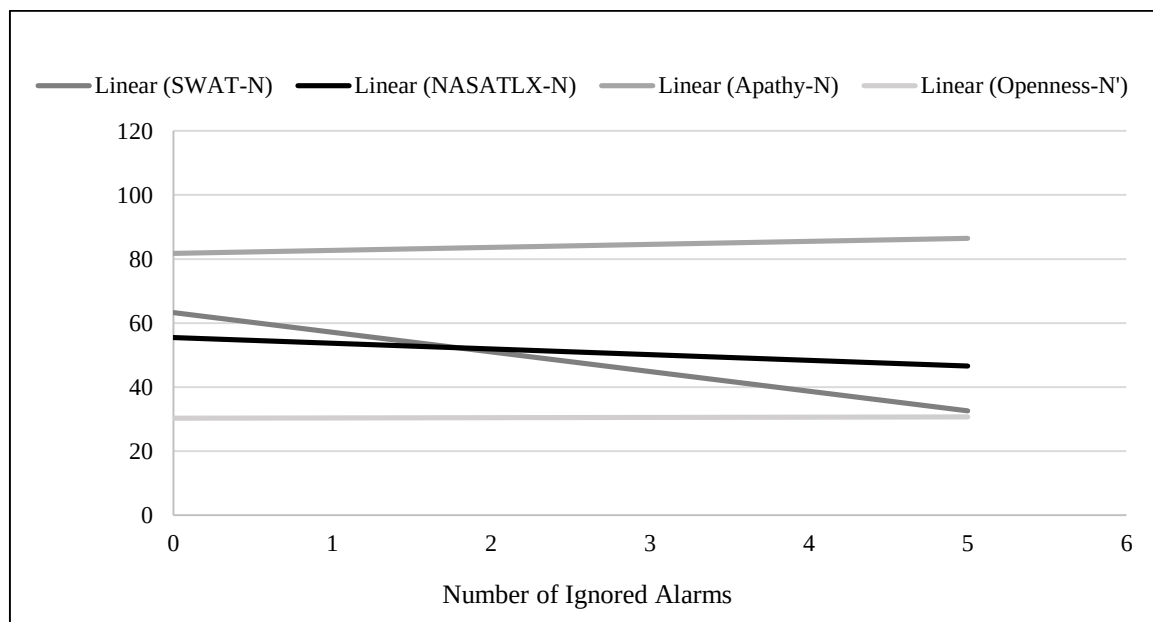


Figure H34: Change in mental workload, affect and openness with number of ignored for nurses

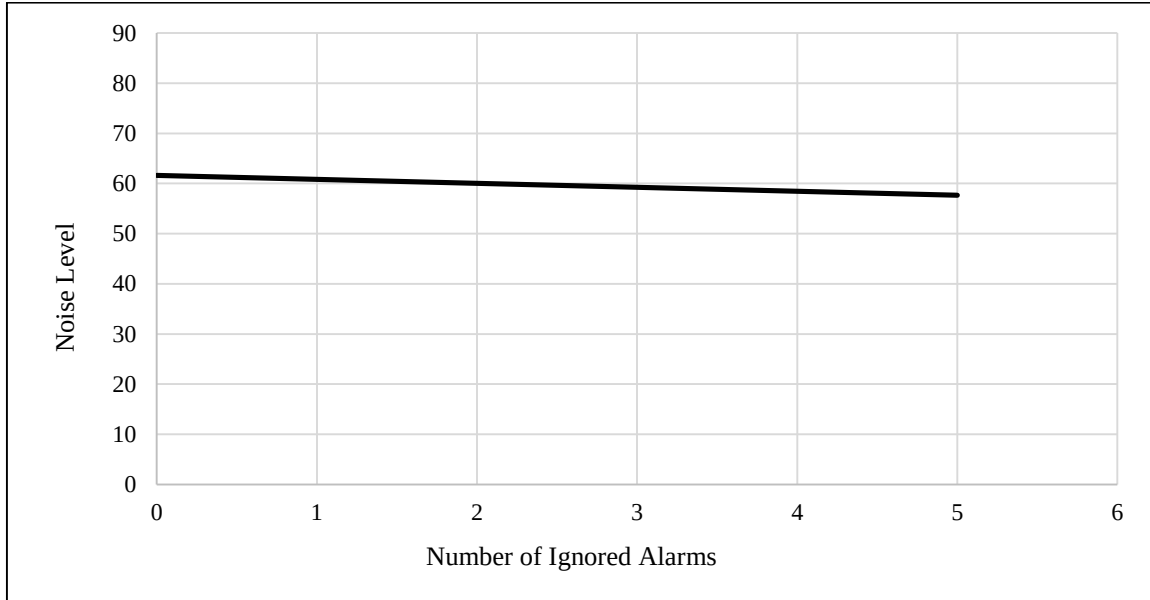


Figure H35: Number of ignored alarms vs noise level for nurses