OBJECTIVE MEASUREMENT OF SPECTACLE WEAR TIME WITH A TEMPERATURE SENSOR DATA LOGGER

By

Matthew Lentsch, B.S.

THESIS

In partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

PHYSIOLOGICAL OPTICS

Presented to the Graduate Faculty of the

College of Optometry
University of Houston

May, 2017

Approved:

____
Heather Anderson, O.D., Ph.D. (Chair)

____
Jason Marsack, Ph.D.

____
Vallabh Das, Ph.D.

.

Committee in Charge

Acknowledgements
I want to thank Dr. Heather Anderson for her direction over the past 3 years working on this project. Without her consistent guidance and incredible patience, I would have never completed this program.

I would also like to thank Drs. Jason Marsack and Vallabh Das for their input and feedback throughout this process.

I also need to acknowledge the invaluable contributions of Hope Queener (development of *Temperature Log Viewer 1.2 & 2.0*) and Chris Kuether (design and production of the silicone mounts) to the completion of this work.

Abstract

Purpose Many practitioners and researchers desire to objectively quantify spectacle wear time. The purpose of this study was to evaluate the Smartbutton Data Logger Temperature Recorder for monitoring spectacle wear.

Methods Fifty adults (32 female, 18 male) wore a thermosensor on their spectacles for 2 weeks for each of 2 mount types while keeping wear diaries. Temperatures during reported spectacle wear (ON) were compared to temperatures during non-wear (OFF). The success of two methods to approximate wear time was evaluated by percent error with respect to subject reported wear time. The first filtered temperatures, defining wear time from temperature ranges determined from group or individual mean temperatures calculated during subject-reported ON times. The second utilized examiners interpreting temperature versus time plots to identify spectacle wear.

Results Group mean ON ($31.8 \pm 0.6\ ^{0}$C) and OFF ($24.7 \pm 1.5\ ^{0}$C) temperatures differed significantly ($p<0.001$), female ON temperatures averaged 1^{0}C higher than males ($p=0.04$), and there was no significant difference in temperature between mounts ($p=0.18$) by repeated measures ANOVA. Median percent error and first and third quartiles (Q1, Q3) of each approximation technique was: group mean filtering = 8% (Q1 3%, Q3 18%), individual mean filtering = 7% (Q1 4%, Q3 19%), examiner 1 = 6% (Q1 2%, Q3 14%), examiner 2 = 7% (Q1 3%, Q3 12%). Evaluation of the most detailed diaries (8 subjects) revealed that brief wear intervals and leaving spectacles in a warm, parked car resulted in higher percent error in approximating wear time.

Conclusions The SmartButton is a promising device to monitor spectacle compliance in patients with all approximation methods evaluated providing less than 10% median percent error in wear time.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

The ability to monitor spectacle wear objectively is a desirable tool for many practitioners and researchers in vision care. Eye care practitioners often prescribe spectacles to be worn full-time for patients with certain conditions, such as amblyopia or latent hyperopia with the goal of improving visual acuity over the long-term; however, the patients may not demonstrate good compliance if they do not perceive an immediate improvement in clarity of vision (as contrasted with patients with myopic refractive error). Spectacle wear alone has been shown to improve visual acuity in patients with strabismic or combined strabismic and anisometropic amblyopia,[1] and thus spectacle non-compliance could negatively impact treatment success. Other researchers are interested in determining predictive factors that lead to non-compliance in children[2-5] and could use a tool that objectively monitors spectacle wear to determine which children are non-compliant. Specifically for this work, monitoring spectacle wear in patients with Down syndrome is of interest for work being conducted to optimize spectacle prescribing strategies for these patients.

More than 400,000 people in the United States currently have Down Syndrome (DS) and many experience poor visual acuity even when corrected with spectacles.[6-8] The source of reduced visual acuity in patients with DS is currently unknown; however a recent study found that patients with DS have elevated higher order aberrations when compared to controls without DS.[7] The presence of elevated aberrations can impact visual performance and may not be best compensated by standard refraction techniques.[7] Current studies are now exploring computerized methods of determining spectacle prescriptions based on the higher order aberrations of an individual in hopes of optimizing refractive error correction and thus improving best corrected visual acuity in patients with DS.

Due to cognitive deficits and the inability of some patients to participate in a visual acuity measure or communicate which spectacle prescription they prefer, we must first develop alternative metrics of spectacle performance for evaluating these novel prescribing techniques. In a spectacle dispensing trial, subjects with DS will be provided 2-3 pairs of spectacles given over separate time periods, 1 based on standard refraction techniques (subjective refraction, retinoscopy, or auto-refraction) and 1-2 based on the novel computerized techniques. One quantitative measure of subjective preference could be voluntary spectacle wear time in one prescription versus another. However, there is not a technique readily available that can quantify spectacle wear time objectively.

The current technique of quantifying spectacle wear time is through the use of questionnaires or wear time diaries taken by subjects or parents of subjects. This technique has been used in recent clinical trials in which researchers sought to compare spectacle and contact lens wear time.[3] However, other studies have shown that adult subjects vastly overestimate spectacle wear time on questionnaires.[9] Given that individuals with DS would likely be unable to complete a wear time questionnaire, the parent or guardian of these participants would need to provide this information. It can be inferred that that parents would likely provide an even more inaccurate estimate of wear time for children that spend part of their day away from the parent at school or other activities.

Researchers in Europe have sought the use of various thermosensors to monitor treatments for ocular and refractive conditions. These thermosensors have been mounted on both spectacles and ocular occlusion patches as a means of monitoring spectacle wear time or eyepatch wear time.[10, 11] While the results of their studies are encouraging, the devices that were validated for spectacle wear monitoring are not readily available in the United States. Maconachie et al. and Januschowski et al. both reported agreement between thermosensors and subject reported spectacle wear time, as reported by the subjects in a diary. [10, 11] However, their samples were limited to 3 and 4 subjects and the devices only sampled every 10 and 15 minutes, despite the capability to sample more frequently. Januschowski et al. also reported less accuracy in predicting spectacle wear time in environmental conditions between 33-37° Celsius (C) which could be problematic for practitioners or researchers deploying the device in warm climates. [10]

The purpose of this study was to validate the *ACR Systems Smartbutton Data Logger Temperature Sensor* as an objective device that can be mounted to spectacles to approximate wear time. This study utilized a large sample size of 50 adult subjects without disabilities to validate the data logger as a means of approximating spectacle wear time prior to considering the future use of this device in patients with Down syndrome. Subjects were encouraged to wear their spectacles in a variety of environmental conditions to determine if any conditions resulted in decreased accuracy of approximated spectacle wear time.

Prior to deploying the thermosensor to our subject population of adults without disabilities, we first needed to evaluate the capabilities of the device within a controlled laboratory setting. The specifics that were tested included: 1) the variability in sampled temperatures between multiple devices to ensure all devices record the same temperature (e.g. quality control evaluation of the devices to reliably record temperature); 2) time course of the device to reach peak or minimum temperatures when moved from extreme hot to extreme cold; 3) impact of proximity to the body on the temperature recorded by comparing the differences in temperature data from sensors exposed to a variety of matched environments when worn on spectacles versus in a pocket or on a stationary surface; and 4) determination of whether or not one side of the device is more sensitive to temperature changes than the other such that a standardized method of device orientation is necessary when mounting the device onto subjects' spectacles.

From the preliminary studies, we learned: 1) all devices record essentially the same temperature; 2) devices reach peak or minimum temperatures relatively quickly (within 15 minutes); 3) the device does not reach internal body temperature when worn on spectacles and thus further studies are needed in cold climates to determine the range of temperatures recorded during spectacle wear in a variety of environmental conditions; 4) we will need to evaluate the pattern of the temperature versus time data trace rather than temperature data alone given the range of temperatures recorded during spectacle wear being impacted by the environment; 5) the logo side of the device is more sensitive to temperature detection and thus we will deploy the device to subjects with the logo side facing the head in an attempt to minimize environmental impacts. These preliminary results also influenced the design of subject logs for the main study, such as the decision to include information about subject hair covering or not covering the device during spectacle wear. The results of these preliminary studies are presented in greater detail in the appendix.

After the preliminary studies were performed, the device was deployed to a population of 50 adult subjects without developmental disabilities. These subjects wore the data logger mounted to their spectacles in one of two ways, each for 2 weeks. While wearing the data logger, subjects were required to keep spectacle wear time logs which accounted for the location of their spectacles for the entire week, specifically whether or not the spectacles were being worn. The data logger sampled the temperature every 5 minutes and stored the data for one week at a time. After four weeks of data collection, the temperature data was analyzed to assess the accuracy of different methods of spectacle wear time approximation.

When analyzing the data, we first sought to determine if spectacle wear time could be accurately approximated based solely on the temperatures alone by filtering data for specific ranges of temperatures, or if we needed to visually assess the pattern of the temperature versus time data to successfully identify spectacle wear. In addition, we sought to: 1) identify problematic subject wearing patterns and environmental conditions that could affect the accuracy of this device, 2) assess whether inspection of temperature versus time data allowed accurate prediction of the actual spectacle wear ON and OFF transitions times, and 3) determine the sampling rate that would provide acceptable accuracy in spectacle wear time estimates, but also allow for the longest duration of data collection before the thermosensor reached its storage capacity. The findings of this study will be described in detail in the following chapter.

The ultimate goal of this study is to provide a new device for objectively monitoring spectacle wear time. This new approach can then be used in our intended population of subjects with Down syndrome to monitor their spectacle wearing patterns with various experimentally-derived spectacle prescriptions as described above. The approach in this study can also be applied to other studies in which researchers are concerned with compliance of prescribed spectacle wear, such as for amblyopia treatment or to reveal predictive factors as to why children choose not to wear their spectacles, among many other applications.

Chapter 2: Objective Measurement of Spectacle Wear Time with a Temperature Sensor Data Logger

Introduction

Assessing spectacle compliance is an important factor in the management of some patients, or the evaluation of treatment success for various ocular conditions. For example, spectacle wear alone has been shown to improve visual acuity in patients with strabismic or combined strabismic and anisometropic amblyopia,[1] and thus spectacle non-compliance could negatively impact treatment success. The technique may also be useful in measuring spectacle wear compliance to reveal predictive factors as to why children choose not to wear their spectacles.[2-5] Currently, there is not an accepted best method to objectively quantify spectacle wear time for patient care or research. Investigators may utilize questionnaires administered to subjects or parents of subjects to quantify wear time; however, this methodology has limitations in that it has been shown that adult subjects over-estimate spectacle wear time significantly on questionnaires.[3, 9, 12] Another strategy is to evaluate compliance based on whether subjects wear their spectacles to follow-up appointments, but this too is not a precise measure of day to day compliance.[5, 13] Objective techniques are necessary when, for instance, dealing with a subject that is forgetful, or parents of a young subject that may not always be with the subject to document wearing patterns.

A thermosensor is a device that autonomously records temperature data from the surrounding environment at set intervals. Multiple researchers have sought to use a thermosensor to objectively measure spectacle and/or eye patch wear time (Occlusion

Dose Monitoring).[9, 10, 14] One study used the thermosensor as the gold standard when comparing a newly developed questionnaire about spectacle wear[9]. Although this study showed promise in utilizing a thermosensor on spectacles, it is first necessary to validate the thermosensor as a means to monitor spectacle wear. Maconachie et al. used a dose monitoring system, previously validated for ocular occlusion monitoring, mounted on spectacles to monitor spectacle wear.[11] This system utilizes a temperature gradient between the front and back of the device. Januschowski et al. sought to validate an orthodontic thermosensor (TheraMon® microsensor), mounted to both spectacles and eye patches, to monitor spectacle wear time.[10] The results of both of these studies showed agreement between the sensor and subject reported wear time as recorded by the subject in a diary; however, their samples were limited to 3 and 4 subjects, the devices only sampled every 10 and 15 minutes despite the ability to sample at greater intervals, and these particular thermosensors are not commercially available outside of Europe.[10, 11] Although their findings are promising, Januschowski et al. also found that the orthodontic thermosensor showed a lack of accuracy in predicting spectacle wear time in environmental conditions between 33 – 37^0 Celsius (C), signifying potential difficulties in warmer climates.

The present study sought to determine the ability of a commercially available temperature sensor to accurately monitor spectacle wear in a large sample of subjects. In this study, we evaluated two methods for mounting the thermosensor to the spectacles to determine whether this factor influences accuracy of spectacle wear measurements derived from the temperature data, or subject tolerance to wear the thermosensor. We also evaluated two strategies to approximate spectacle wear from temperature data: filtering data based on temperature ranges to identify spectacle wear versus examiner inspection of temperature versus time plots to identify spectacle wear. We aimed to recruit a large sample of adult subjects in an effort to collect data obtained during a variety of wear patterns and environmental conditions to help determine conditions that could potentially negatively impact accurate assessment of spectacle wear time. The overall goal of this study is to form recommendations regarding deployment of the device to objectively monitor spectacle wear for use in future research studies and eliminate the need for patient spectacle wear diaries.

Methods

Subjects

This research was approved by the University of Houston Committee for the Protection of Human Subjects and adhered to the tenets of the Declaration of Helsinki. Informed consent was collected from all subjects. Adult participants (18 years or older) were recruited from the students, faculty, and staff at the University of Houston College of Optometry, as well as adults from the community who were family, friends, or acquaintances of the investigator conducting data collection. All subjects were required to own spectacles and commit to part-time wear for a minimum of 10 hours per week throughout the duration of the study. Full-time wear was discouraged due to the desire to eliminate predictable wear patterns from the dataset, and thus potential participants who were dependent upon full-time spectacle wear (e.g. did not have contact lens corrections or secondary spectacles as an alternative refractive correction option) were not included.

Description of Thermosensor Device

The thermosensor used for this study was the *Smartbutton Data Logger Temperature Recorder (ACR Systems Inc., British Columbia, Canada).* The sensor is a commercially available, low-voltage, battery powered Class B Digital device that passively obtains and stores temperature readings at fixed time intervals (1 to 255 minutes).The sensor specifications are as follows: 17.35 mm diameter x 5.89 mm height, 4 grams, 3.0 volt lithium battery (approximated battery life of 10 years at a 20 minute sample rate), temperature operating range -40° C to 85° C, logs up to 2048 consecutive time-stamped temperature measurements in internal memory in 0.5° increments. The internal memory is retrieved through a USB interface connected to a computer with *Trendreader for Smartbutton Software (ACR Systems Inc., British Columbia, Canada)*; data are not transmitted wirelessly from the device. The software extracts the data into a spreadsheet with temperature recordings and the corresponding time-stamp, as well as a graph of temperature versus time.

Mounting Strategies

Temperature sensors were mounted to subjects' spectacles in two different ways. One mount type used a double-sided adhesive circle, 15 mm in diameter, to adhere the sensor directly to the temple of the subject's spectacles behind their ear (Figure 2.1A). The alternate mount type was a custom designed silicone sleeve which held the sensor either above or below the temple of the spectacle frame behind the subject's ear (Figure 2.1B). Only one subject wore the silicone mount below the temple due to spectacles with straight temples that accommodated this option. Each mount type was worn for two weeks by every subject with mount type order randomized.

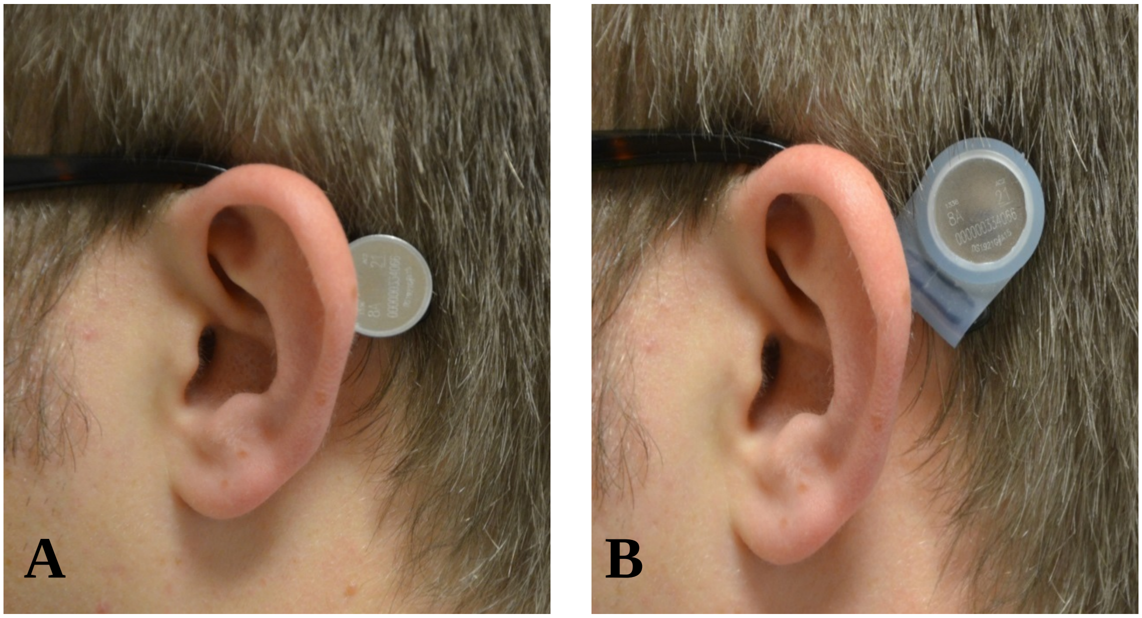


Figure 2.1. Two strategies were used to mount the sensor to the subject's spectacles. One mount used a double sided adhesive circle to adhere the sensor to the temple of the spectacle frame behind the subject's ear (A). The alternate mount used a custom designed silicone sleeve which held the sensor either above (shown here in B) or below the temple of the spectacle frame behind the subject's ear.

Data Collection

Subjects were requested to wear their spectacles with the sensor for a minimum of 10 hours per week and to complete daily logs (accounting for all 168 hours of the week) with specific times (hour and minute) that spectacle ON/OFF transitions occurred. These logs were returned to study investigators weekly. Subjects were encouraged to document additional details in the logs which may impact temperature readings, such as whether or not the subject's hair covered the sensor during wear and other environmental details corresponding to the location of the glasses during specific log times (e.g. wearing glasses while driving, activities indoors, activities outdoors, leaving the glasses in a parked car, etc.). Subjects were encouraged to use note taking programs on their mobile device to assist in the documentation of exact times that they put their spectacles on or off, and to transfer this information daily to the log sheets. The sensors were set to record time-stamped temperatures every five minutes for one week (seven days and two hours). At the end of each week, subjects returned to the lab for a follow-up visit. At each follow-up visit, subjects submitted their log from the previous week, data from the sensor was downloaded and the sensor was reset. Using Microsoft Excel, a cell was populated with each time stamped temperature recording and an adjacent cell was marked based on the subject log report of spectacles 'ON' or 'OFF'. Temperature recordings were taken every 5 minutes, however subjects' 'ON'/'OFF' transitions often did not match the recordings to the minute. The merging of data was standardized to not include more wear time than was reported by subject logs. For example, if a subject reported that spectacles were 'ON' at 1:42, but the sensor recorded temperature at 1:40 and 1:45; when merging the data, spectacles were not marked as 'ON' until the 1:45 data point.

At the end of each two week period, subjects completed a survey about the mount type worn for the past two weeks. The survey was scored 1-5 (1=worst, 5=best). Subjects were asked to rate the comfort of the mount, the cosmetic acceptability of the mount, and the perceived fragility of the mount (5 = not fragile). The survey also included a question asking whether the subject would be willing to continue wearing this type of mount, which was answered *yes* or *no*.

Data Analysis

Subject Log Confidence

Subject log data were manually entered into the downloaded time-stamped temperature data using Microsoft Excel to create a file of subject reported spectacle ON/OFF times associated with temperature readings. Subject logs were graded for confidence (high, medium, low) in the accuracy of logged data. The confidence grading was standardized as follows. Characteristics of high confidence logs included: the subject accounted for all of the time in the week, time markings were specific and clear (eg. AM/PM), and environmental details were provided. Medium confidence meant one of the characteristics of high confidence was missing consistently throughout the log. Low confidence meant that two or more of the characteristics of high confidence were missing throughout the log, or subjects reported poor log quality and/or filled in their logs at the time of the follow-up visit.

Establishing a Reference to Compare Temperature Data

The temperatures associated with subject-reported wear time recorded in the diary served as ‘truth’ in this experiment (Appendix Figure A6). Mean and standard deviation of the temperatures recorded during the time intervals the subject reported the spectacle ON versus OFF were calculated for each subject individually and the entire group combined. These temperature ranges were then used to filter individual subject data to predict wear times as described below.

Approximating Spectacle Wear by Temperature Filtering

First, the ability to approximate spectacle wear was tested based purely upon the temperatures recorded by the thermosensor. Four methods of data filtering were attempted with the goal of determining an appropriate temperature range to capture ON temperatures without overlapping the OFF temperatures. The first two filtering methods defined spectacle ON times using the entire group mean and standard deviations (SD) of ON temperatures for week 1 data. Spectacles were marked as ON when recorded temperatures fell within 1.0 average SD and 1.5 average SD around the mean. The other two filtering methods were individualized per subject and defined spectacle wear as times when recorded temperatures fell within 1.5 SD and 2.0 SD of that individual subject's mean spectacle ON temperature. Using these four filtering methods, spectacle wear was approximated across each week of study participation and compared to subject reported spectacle wear times.

Approximating Spectacle Wear by Examiner Inspection of Temperature Plots

Second, in an attempt to utilize information from the patterns of the temperature versus time plots rather than merely the actual temperature values themselves, spectacle ON time was approximated by two examiners visually inspecting the log data who were masked to the subject logs. Each examiner independently viewed temperature versus time plots using *Temperature Log Viewer 1.2,* custom software (Matlab, Natick MA) developed for this study to allow examiners to mark presumed spectacle ON and OFF times on the plots of the raw temperature data (Figure 2.2). The examiners reported that they determined what was believed to be spectacle ON time based on sharp increases in temperature, temperature range, and the frequency (noise) of the data trace. The software then calculated spectacle ON time based on examiner markings and compared this to subject reported ON time.

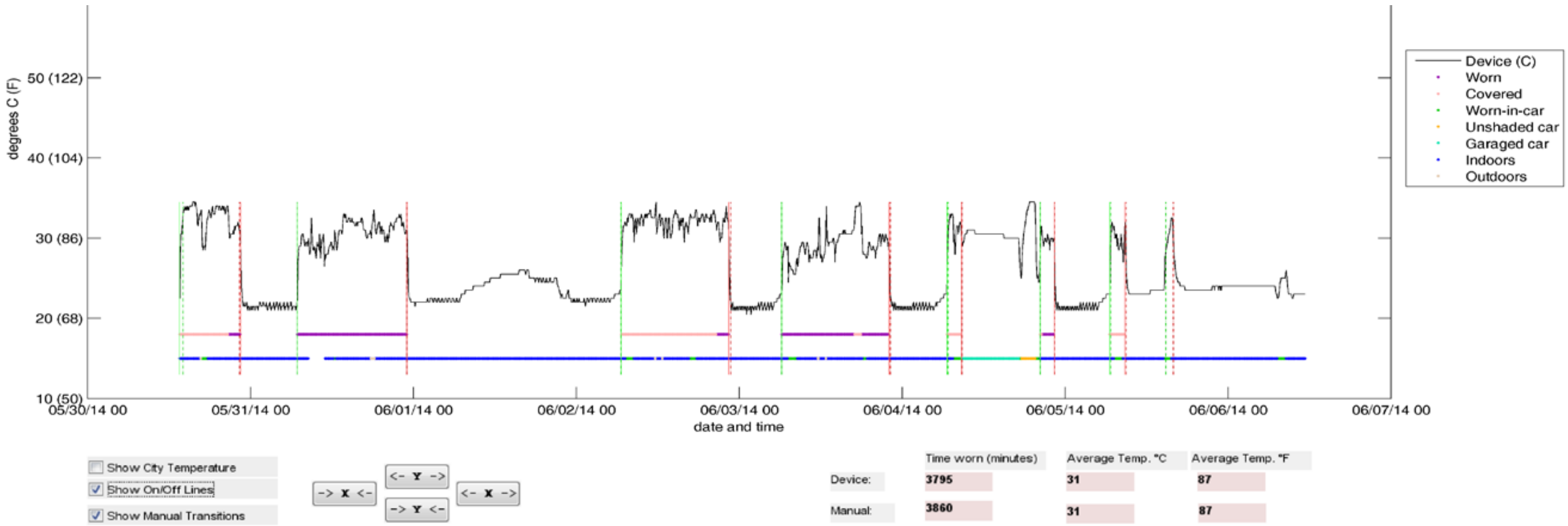


Figure 2.2. A sample of a completed plot using *Temperature Log Viewer 1.2.* During evaluation of the plot, examiners can only view the solid black line depicting temperature versus time. The green (on) and red (off) dashed lines indicate the manual markings placed by the examiner to approximate transitions of spectacles on and off. All other markings and numbers appear once the examiner submits their final markings. The pink and purple solid horizontal lines indicate subject reported spectacle on time and whether or not the sensor was covered by their hair. The blue, yellow, and green solid horizontal lines indicate the location (indoors, outdoors, or in car) of the sensor. Manual time worn (bottom right side) is calculated once the examiner submits their markings. Device time worn indicates subject reported wear time and is visible upon completion of the examiner markings.

Accuracy of Approximation Methods

Percent error was calculated to determine the accuracy of each method of spectacle wear time approximation for all 50 subjects for week 1 data using the following formula:

*Percent Error = ((Approximated ON Time – Subject Reported ON Time)/Subject Reported ON time) * 100*

The absolute value of percent error was grouped into three categories (≤10%, >10 to ≤20%, and >20%) for each wear approximation method and chi-square analysis performed to determine whether the distribution of errors differed between methods. Percent error versus subject reported spectacle wear time was also plotted for each method and regression analysis was performed to determine if percent error was linearly related to the duration of spectacle wear.

Survey Data Analysis

Mean and SD were calculated for the first three survey questions and paired t-tests used to compare the adhesive mount type to the silicone mount type. Yes/No responses to the last survey question for both mount types were tallied.

Wear Pattern and Environmental Condition Assessment

A subset of data from subjects with good quality logs for all four weeks was further descriptively analyzed to identify potential sources of high percent error with the various wear time approximation methods. Both subject wear patterns (total wear time, number of on/off transitions, average duration of continuous wear time) and environmental details were compared across subjects in an effort to determine sources of poor spectacle wear time approximation.

Analysis of Examiner Ability to Predict Actual Periods of Wear Time

Temperature Log Viewer 1.2 compares total wear times as approximated by examiners versus subject wear logs, but it does not evaluate whether examiner assessment of ON and OFF transitions corresponds to the actual transitions reported by subjects. To assess whether wear intervals identified by examiners matched true wear intervals reported by subjects, descriptive analysis was further performed on week 1 data from the subset of subjects with good quality logs syncing the intervals of wear marked by examiners in the spreadsheet of subject reported wear. The percentage of agreement between examiner and subject report was quantified for both ON and OFF spectacle wear.

Analysis of Sampling Interval

An additional aim of this study was to identify the minimum sampling rate necessary to approximate spectacle wear time, given that sampling rate will impact total storage time, and therefore the need to visit the laboratory to download data. The sensor in this study was set for a high density sampling rate (every 5 minutes) which yielded 7 days of data. To determine whether a coarser sampling rate would negatively impact percent error, the functionality of the plot marking software was updated (*Temperature Log Viewer 2.0)* which allowed the examiner to select the sampling rate for which to view the temperature versus time data, allowing coarser representation of the original set of temperature versus time data. Examiner 1 re-graded the plots for week 1 data of the eight subjects with good quality logs for all weeks, using three additional sampling intervals (every 20 minutes, 30 minutes, and 60 minutes). Sampling at these intervals would allow the device to be deployed for the following periods of time: 5 minute sampling = 1 week, 20 minute = 4 weeks, 30 minute = 6 weeks, and 60 minute = 12 weeks. Percent error for each sampling interval was compared to evaluate the impact of coarser temperature sampling. In addition, the percentage of agreement between actual subject reporting timing of spectacle wear and intervals marked by the examiner were evaluated for the various sampling intervals.

Results

Fifty-three subjects were enrolled in the study, three of whom withdrew prior to the conclusion of the study. Two subjects withdrew due to inability to schedule weekly follow-up appointments and the other due to skin irritation from the silicone mount against his bald head. One subject lost a sensor due to the adhesive mount detaching from the spectacle frame, but a new sensor was dispensed and the subject completed an additional week of data collection. In total, 50 subjects (32 female, 18 male) completed data collection and are included in the analysis below. Data were collected between May and August in Houston, Texas. The outdoor temperature range during this period was 22 to 37 degrees Celsius (73 – 103 Fahrenheit).

Confidence in subject reported log data was graded as high, medium, or low; as described above. Log confidence for all 50 subjects in week 1 was as follows; 18 high, 28 medium, and 4 low. Log confidence varied within most subjects over the four week study period, and generally decreased. The number of logs with high confidence across weeks 1 - 4 were 18, 15, 12, and 17, respectively. Therefore, we present the week 1 data for all 50 subjects when log confidence was highest, but included only the subjects with high log confidence for all four weeks of data collection for the subsequent analysis. In total, eight subjects were graded to have high log confidence for all four weeks.

Temperatures extracted from data sensors were categorized as spectacles ‘ON’ or ‘OFF’ based on merging data logger time stamps with subject log reported wear activities.

Group mean ON & OFF temperatures for all subjects for week 1 are shown in Table 2.1. Data were analyzed by repeated measures ANOVA with two between group factors (gender, mount types). Group mean ON ($31.8 \pm 0.6\ ^{0}C$) and OFF ($24.7 \pm 1.5\ ^{0}C$) temperatures differed significantly ($p<0.001$), as did male and female mean temperatures, with females averaging 1^{0} C higher than males ($p=0.04$). There was not a significant difference in mean temperatures between mount types ($p=0.18$).

Mean ± SD data logger temperatures (^{0} C) categorized by subject wear logs					
	Group (n = 50)	**Male (n = 18)**	**Female (n = 32)**	**Adhesive (n = 23)**	**Silicone (n = 27)**
ON	31.8 ± 0.6	31.2 ± 1.7	32.2 ± 1.1	31.5 ± 1.6	32.0 ± 1.2
OFF	24.7 ± 1.5	24.2 ± 2.0	24.9 ± 2.1	24.3 ± 1.8	25.0 ± 2.2

Table 2.1. Repeated measures ANOVA (ON vs OFF) with two between group factors (gender, mount type) found that group mean ON & OFF temperatures differed significantly ($p < .0001$) and varied with gender ($p = 0.04$), but not mount type ($p = 0.18$).

Four temperature filtering strategies for approximating spectacle wear time were analyzed, as described in the methods above. The top performing group mean and individual mean strategies are reported here. The first method filtered based on the mean ON temperature for the entire group, marking spectacles as ON for temperatures between 28.4 to 35.2 ^{0}C (range determined from the group mean spectacle ON temperature ± 1.5 times the average SD of the group ON temperature). The next method filtered based on each individual subject's mean ON temperature ± 1.5 SD. Approximated spectacle wear time based on two masked examiners' visual inspection of the temperature plots is also reported here.

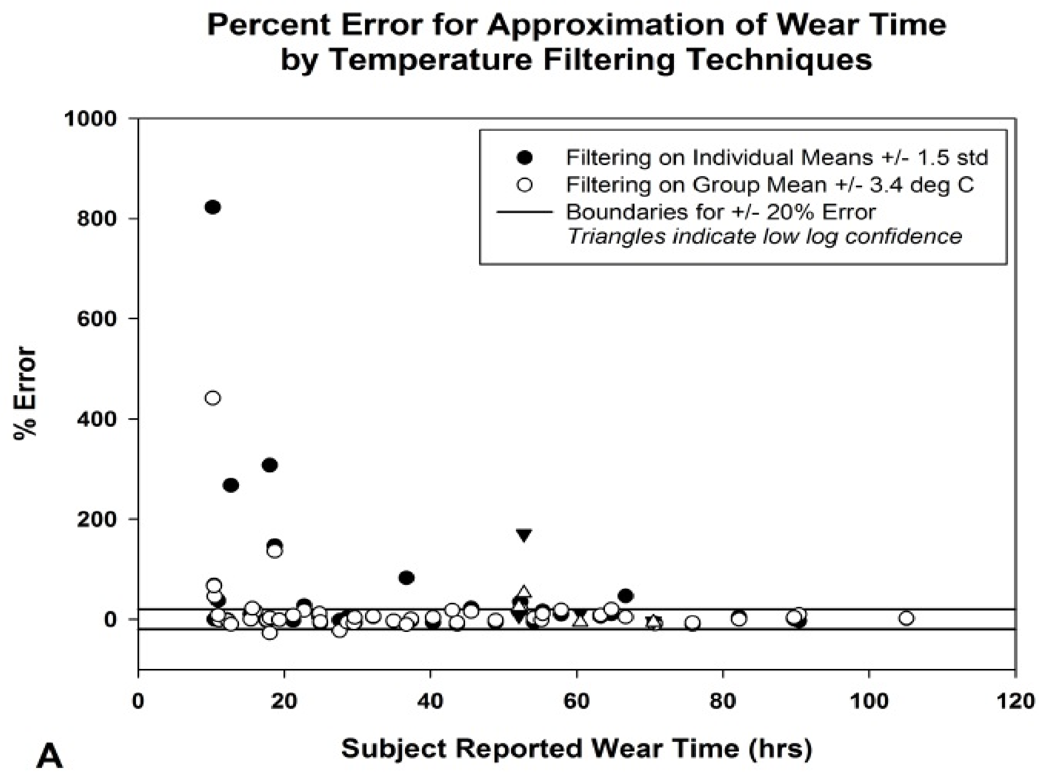
Accuracy of Approximation Methods

The distribution of the absolute value of percent error from week 1 for each wear approximation method is shown in Table 2.2. All methods had a similar distribution of percent error for the 50 subjects across the three categories (Table 2, Chi-square, p = 0.72). Although the masked examiners had somewhat better median percent error than the filtering methods upon visual comparison (6.1% & 6.9% versus 7.1% & 8.0%), there was no significant difference in median percent error across methods (Kruskal-Wallis, p = 0.63).

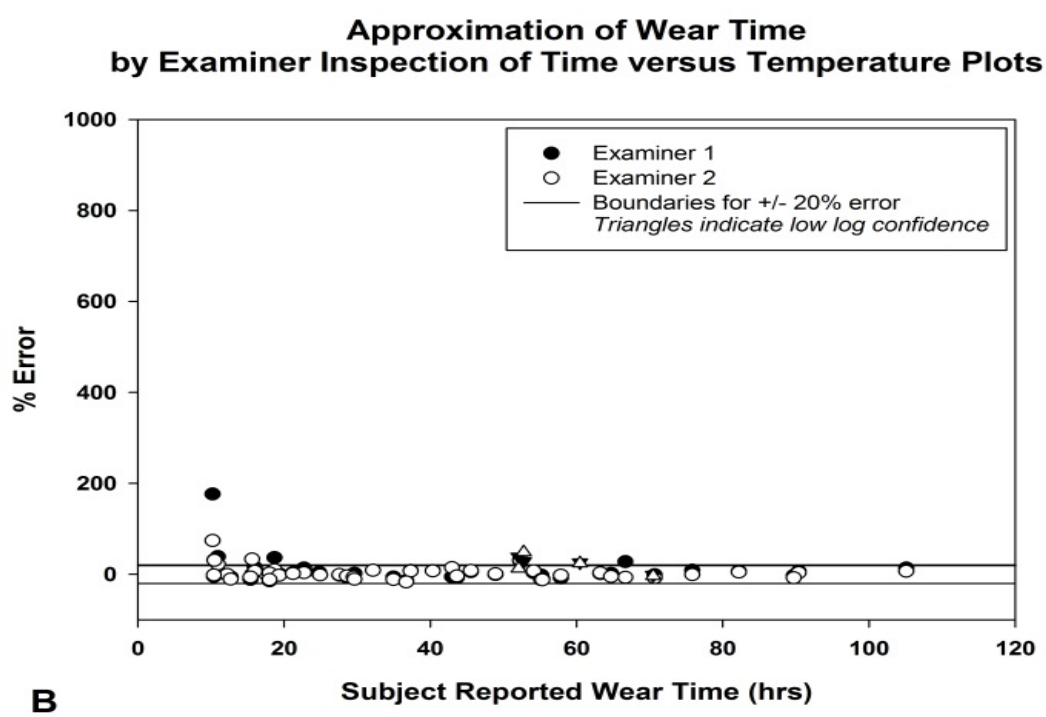
Classification of Absolute Percent Error in Approximated Wear Time				
	Group Mean ± 3.4 ° C	**Ind. Mean ± 1.5 SD**	**Examiner 1**	**Examiner 2**
Median Error (Quartile 1, Quartile 3)	8.0% error (3%, 18%)	7.1% error (4%, 19%)	6.1% error (2%, 14%)	6.9% error (3%, 12%)
≤ 10% Error	32 subjects (64%)	32 subjects (64%)	35 subjects (70%)	35 subjects (70%)
11 – 20 % Error	9 subjects (18%)	5 subjects (10%)	6 subjects (12%)	8 subjects (16%)
> 20% Error	9 subjects (18%)	13 subjects (26%)	9 subjects (18%)	7 subjects (14%)

Table 2.2. The distribution of classification of absolute percent error did not differ between methods (Chi-square, $p = 0.72$).

Figures 2.3A&B show signed percent error for individual subjects versus subject reported spectacle wear time for each of the wear time approximation methods. Positive values indicate overestimation of wear time and negative values indicate underestimation of wear time. Although the distribution of percent error did not differ between methods (Table 2), the overall magnitude of error for those subjects falling outside of ±20% error was much greater for the filtering methods. As can be seen in Figure 3, percent error by examiner approximation was less than 100% with one exception nearing 200%, whereas the filtering techniques had numerous instances of percent error exceeding 100% and even beyond 800%. It should also be noted that for all methods, the largest errors were positive errors, indicating an overestimation of wear time (range for all subjects, all methods = -27% to 822%). There was not a linear relationship between subject reported wear time and percent error for any of the approximation methods, although visual inspection does indicate a larger number of over-estimations occurred for subjects wearing their spectacles fewer than 20 hours in the week.


Percent Error for Approximation of Wear Time
by Temperature Filtering Techniques
Filtering on Individual Means +/- 1.5 std
Filtering on Group Mean +/- 3.4 deg C
Boundaries for +/- 20% Error
Triangles indicate low log confidence
% Error
Subject Reported Wear Time (hrs)

A


Approximation of Wear Time
by Examiner Inspection of Time versus Temperature Plots
Examiner 1
Examiner 2
Boundaries for +/- 20% error
Triangles indicate low log confidence
% Error
Subject Reported Wear Time (hrs)

B

Figure 2.3. There was no significant linear relationship between reported wear time and percent error for either filtering method to approximate wear time (3A) ($R^2 \leq 0.07$, $p \geq 0.06$), or either examiner approximating wear time from temperature versus time plots (3B) ($R^2 \leq 0.05$, $p \geq 0.12$). Note that although the percentage of subjects with approximations within 20% error was similar between temperature filtering and examiner approximation techniques, the temperature filtering techniques had a much larger range of percent errors in excess of 800%.

Survey Results

Survey results regarding subject perceptions of the two mount types are shown in Table 2.3. Subjects did not have a preference in terms of comfort (paired t-test, $p=0.15$); however, subjects preferred the cosmesis of the adhesive mount (paired t-test, $p=0.001$) and perceived the silicone mount as less fragile (paired t-test, $p<0.001$). As to whether or not they would wear the mount type again, 42/50 subjects reported *yes* for the adhesive mount, while 39/50 reported *yes* for the silicone mount.

Survey Results (Mean±SD)			
	Comfort	**Cosmesis**	**Fragility**
Adhesive	3.9±1.2	3.9±1.1*	3.8±1.0
Silicone	3.5±1.2	3.3±1.1	4.7±0.6^{+}

Table 2.3. Survey of mount types (1=worst, 5=best) found no significant difference in comfort (p=0.15), improved cosmesis with adhesive (*p=0.001), and less fragility with silicone ($^{+}$p<0.001).

Subgroup Results

Data from the eight subjects with good quality logs for all four weeks were further analyzed to identify potential sources of high percent error with the various approximation methods. Tables 2.4A&B present a summary of signed percent error, by both examiner approximation and temperature filtering, for these subjects for all weeks. Both the filtering methods and examiner approximation performed well with all techniques having less than 20% absolute error for a majority of the subjects across the four weeks of spectacle wear. Group and individual mean filtering provided less than 20% error 78% and 84% of the time respectively, while examiners 1 and 2 provided less than 20% error 81% and 84% of the time respectively. However, as was noted previously with the week 1 data, the magnitude of error was often much higher for the filtering methods for those instances where percent error exceeded 20%. It can be seen in Table 2.4 that subject 9 accounted for the majority of the higher magnitude error for all methods. When looking only at the examiner approximation, subjects 9 and 19 accounted for all of the instances of error greater than 20%.

% Error for Examiner Inspection of Temperature vs Time Plots for Subjects with High Confidence Logs										
	Subject #	*1*	*2*	*3*	*9*	*11*	*19*	*23*	*41*	*Median*
Week 1	Examiner 1	2%	4%	-1%	37%	9%	-8%	14%	2%	3%
	Examiner 2	18%	8%	-7%	10%	0%	-2%	7%	-5%	4%
Week 2	Examiner 1	2%	-1%	1%	32%	-1%	-9%	5%	2%	2%
	Examiner 2	0%	1%	1%	47%	2%	-64%	4%	1%	1%
Week 3	Examiner 1	1%	0%	1%	139%	3%	21%	7%	-4%	2%
	Examiner 2	3%	5%	8%	160%	3%	7%	9%	-7%	6%
Week 4	Examiner 1	5%	4%	-7%	22%	-4%	40%	1%	9%	5%
	Examiner 2	5%	4%	0%	-61%	-6%	54%	0%	14%	2%

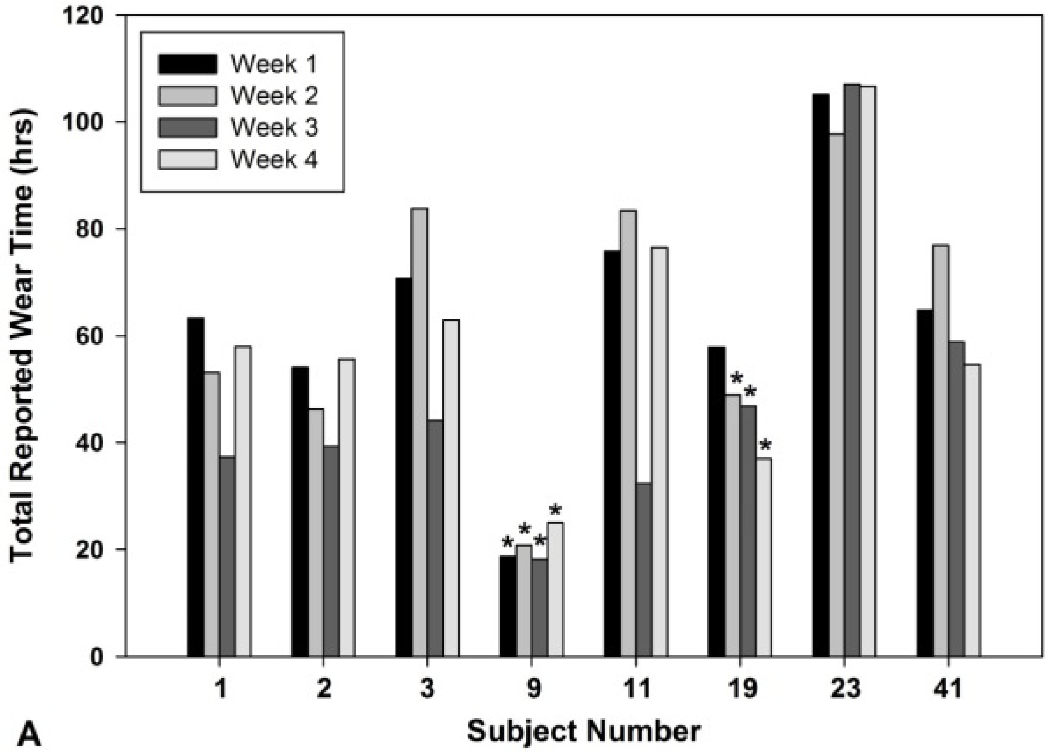
Table 2.4A Positive values indicate overestimation of wear time and negative indicate underestimation of wear time. Subjects with greater than 20% error are shaded in gray.

% Error for Filtering Methods										
		1	*2*	*3*	*9*	*11*	*19*	*23*	*41*	*Median*
Week 1	Individual Mean Filtering	5%	-5%	-8%	147%	-10%	7%	1%	11%	3%
	Group Mean Filtering	9%	3%	-10%	136%	-6%	15%	2%	20%	6%
Week 2	Individual Mean Filtering	0%	8%	-12%	179%	-12%	7%	-6%	-4%	-2%
	Group Mean Filtering	-4%	15%	0%	118%	-12%	11%	2%	-1%	1%
Week 3	Individual Mean Filtering	48%	-7%	-3%	328%	-5%	5%	-2%	-6%	-3%
	Group Mean Filtering	83%	0%	1%	233%	4%	21%	-4%	-10%	3%
Week 4	Individual Mean Filtering	-7%	-4%	-9%	76%	-9%	11%	-7%	3%	-6%
	Group Mean Filtering	5%	1%	-1%	93%	-9%	45%	-12%	-1%	0%

Table 2.4B. Positive values indicate overestimation of wear time and negative indicate underestimation of wear time. Subjects with greater than 20% error are shaded in gray. Filtering temperature ranges were based on individual subject mean ± 1.5 SD and group mean ± 1.5 average SD.

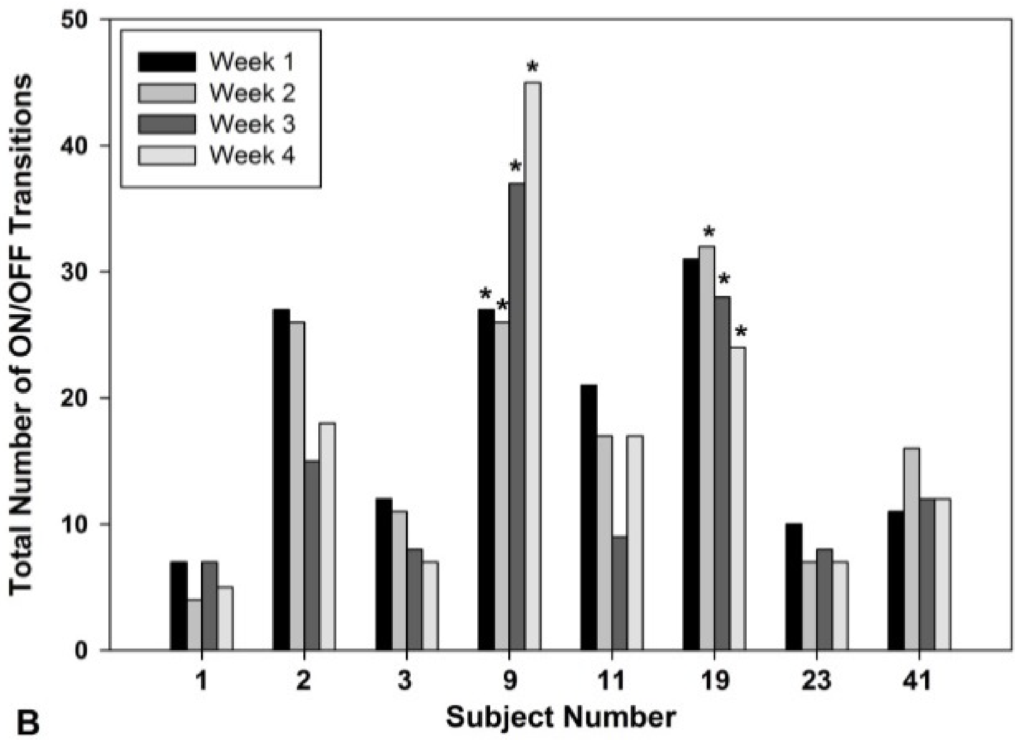
Wear Pattern and Environmental Condition Results

Subject spectacle wear patterns and environmental conditions were next inspected to identify possible sources for poor approximation of wear time for subjects 9 and 19. As can be seen in Figure 2.4, both subjects had similar wear patterns that differed from the group as a whole to include lowest total amount of wear time (2.4A), highest number of ON and OFF transitions (2.4B), and lowest durations of continuous spectacle wear (2.4C).



Total Reported Wear Time (hrs)
Week 1
Week 2
Week 3
Week 4
Subject Number
1
2
3
9
11
19
23
41

A



Total Number of ON/OFF Transitions
Week 1
Week 2
Week 3
Week 4
Subject Number
1
2
3
9
11
19
23
41

B

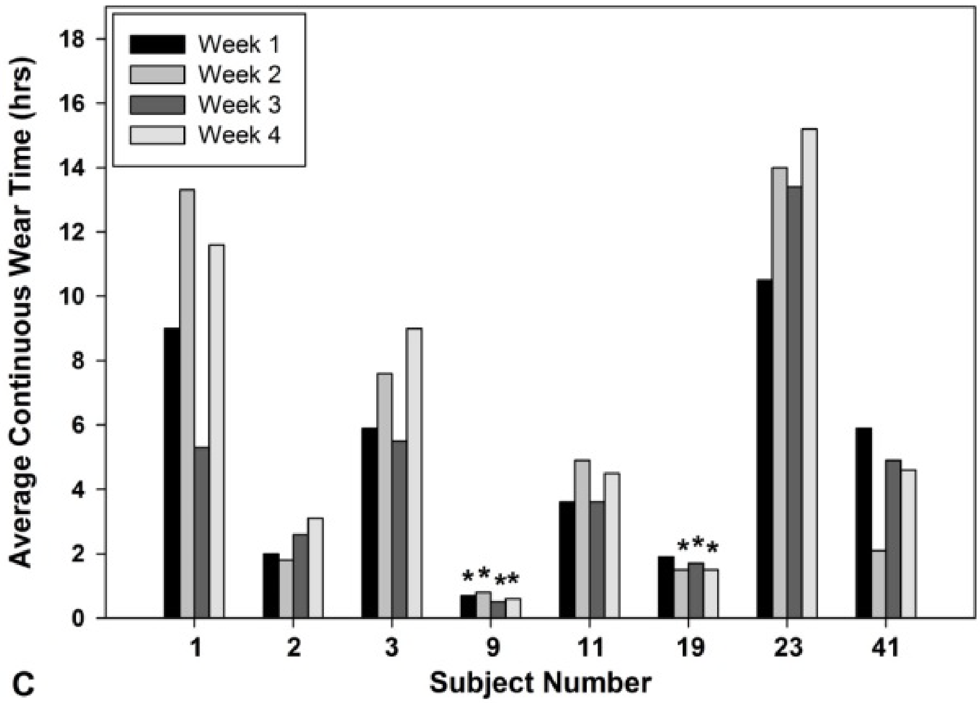


Figure 2.4. Asterisks indicate particular subjects and weeks with absolute percent error greater than 20% by one or both examiners. A trend can be seen in which the subjects with greater error consistently had less total wear time (A) with more frequent ON/OFF transitions (B), resulting in shorter durations of continuous wear time (C).

When evaluating the environmental details provided by each of the eight subjects on their written logs, there was a mix of indoor and outdoor activity throughout the study, with most of the data consisting of indoor wear and indoor storage of spectacles. However subject 9, who had the worst error in wear time approximation, primarily stored their spectacles in the car, only wearing their spectacles for short duration driving trips. In particular, the storage of the spectacles in an unshaded parked car overnight allowed the sensor to reach temperatures that mimicked body temperature, posing difficulty for both the examiner and filtering techniques. An example of one week of temperature data from subject 9 is shown in Figure 2.5, the pattern of which is in stark contrast to the subject in Figure 2.2 who primarily kept their spectacles indoors.

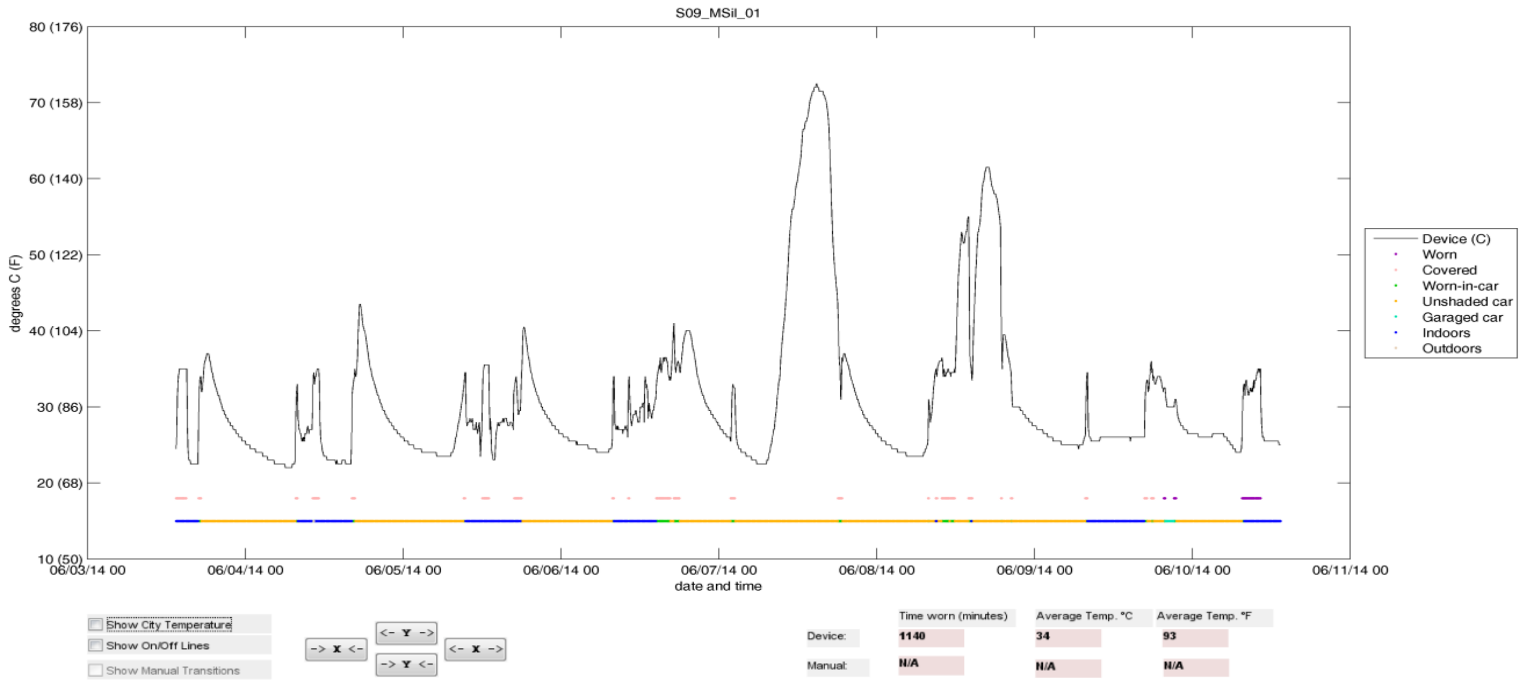


Figure 2.5. The temperature versus time plot for week 1 data of subject 9. The examiners' percent error for this plot were 37% and 10%, meaning both examiners overestimated wear time. It can be seen that for the majority of the week, the spectacles were left in a car that was parked in an unshaded area, often overnight in the warm Houston climate (yellow dots/line). Periods of spectacle ON time were very brief and only represented a small portion of the week (pink and purple dots/line). The combination of storage in an unshaded car can cause temperature ranges near body temperature.

Three of the eight subjects also reported storing their spectacles in a pocket or hanging them on their shirt throughout the study, a behavior that was suspect to mimic ON spectacle temperatures. However, this behavior was not associated with higher percent error. Upon inspecting the temperature data during these periods of pocket or shirt storage, a sharp decrease in temperature was observed when transferring spectacles from ON to a pocket or shirt. The average pocket/shirt temperature was several degrees lower and more variable ($28.2 \pm 3.5\ ^0C$) than the average ON temperature ($31.8 \pm 0.6\ ^0C$).

Examiner Ability to Predict Actual Periods of Wear Time

To investigate whether the examiners were identifying true periods of spectacle wear, the time stamps for the examiners' marked ON and OFF transitions were exported from *Temperature Log Viewer 2.0* into a Microsoft Excel sheet for comparison with subject reported transition times for the eight high quality log subjects for week 1. The overlaps in ON and OFF spectacle wear were compared and reported as a percentage of agreement in Table 2.5. On average, examiners correctly identified the true timing of spectacle wear 93% (examiner 1) and 91% (examiner 2) of the time and correctly identified no spectacle wear 92% (examiner 1) and 94% (examiner 2) of the time.

Subject	Examiner 1 ON Times	Examiner 1 OFF Times	Examiner 2 ON Times	Examiner 2 OFF Times
1	99%	98%	99%	89%
2	96%	97%	95%	94%
3	97%	99%	91%	99%
9	76%	93%	65%	95%
11	99%	92%	96%	97%
19	84%	96%	91%	96%
23	99%	68%	96%	82%
41	97%	98%	92%	98%
Average	93%	92%	91%	94%

Table 2.5. Percentage of agreement between actual timing of examiner marked ON and OFF periods compared to subject reported time periods.

Sampling Interval Results

Table 2.6 shows that percent error did not vary dramatically when decreasing the sampling rate. The 30 minute sampling interval was the only interval with all eight subjects having less than 20 percent error with small increases in percent error for some subjects, but a sizeable improvement in percent error for subject 9. The analysis comparing the timing of true periods of spectacle wear to examiner estimate periods of spectacle wear (Table 2.6) was then repeated for examiner 1 grading week 1 data at the coarser sampling intervals as shown in Table 2.7. These findings demonstrate that although percent error decreased for subject 9 with 30 minute temperature sampling, the accuracy of the timing of spectacle wear intervals decreased dramatically with coarser sampling.

Examiner 1 Percent Error for Week 1 Wear at Different Sampling Intervals								
Subject #	**1**	**2**	**3**	**9**	**11**	**19**	**23**	**41**
5 Min Sampling	2%	4%	-1%	37%	9%	-8%	14%	2%
20 Min Sampling	13%	0%	-1%	-20%	-5%	30%	10%	2%
30 Min Sampling	11%	6%	-2%	-12%	-1%	18%	10%	11%
60 Min Sampling	-12%	-11%	1%	30%	20%	-5%	-6%	-3%

Table 2.6. No substantial change in percent error was observed overall with less frequent temperature sampling.

Subject #	5 Min Sampling ON Times	20 Min Sampling ON Times	30 Min Sampling ON Times	60 Min Sampling ON Times
1	99%	98%	95%	91%
2	96%	91%	93%	89%
3	97%	96%	95%	94%
9	76%	36%	43%	29%
11	99%	92%	89%	74%
19	84%	96%	94%	80%
23	99%	98%	96%	94%
41	97%	97%	95%	88%
Average	93%	88%	87%	80%
Subject #	**5 Min Sampling OFF Times**	**20 Min Sampling OFF Times**	**30 Min Sampling OFF Times**	**60 Min Sampling OFF Times**
1	98%	91%	90%	87%
2	97%	96%	94%	90%
3	99%	98%	98%	97%
9	93%	95%	94%	95%
11	92%	97%	92%	95%
19	96%	81%	87%	86%
23	68%	79%	77%	81%
41	98%	97%	90%	91%
Average	92%	92%	90%	90%

Table 2.7. Percentage of agreement between actual timing of examiner 1 marked ON and OFF periods compared to subject reported ON and OFF periods for four different sampling intervals of temperatures.

Discussion

The results of this study indicate that the *Smartbutton Data Logger Temperature Recorder* is successful as a tool to monitor spectacle wear time utilizing the following recommendations. Spectacle wear time approximations were most accurate for the method of examiners marking temperature versus time plots for each subject's data, as seen by the smaller overall range of percent error (Figure 2.3). Although this study initially used a 5 minute temperature sampling rate, utilizing a coarser temperature sampling rate of 30 minutes is likely to be successful and will extend subject follow-up time. Thirty minute sampling did not elevate percent error beyond the threshold of 20% error (Table 2.6), nor did it negatively impact the accuracy of identifying true wear intervals (Table 2.7), except for subject 9 whose worse agreement with 30 minute sampling was likely due to wear intervals briefer than the sampling interval. Also related to findings from subject 9, we recommend that examiners collecting data in warm climates similar to our study location (22 to 37^0 C) advise subjects not to leave spectacles in their car, as this could lead to a higher chance of temperatures recorded that mimic true spectacle wear. Lastly, regarding our two tested mount strategies, we have adopted a preference for the silicone mount type, largely due to the increased risk of losing both the sensor and data if adhesive mounting fails.

Although the data lead us to favor the examiner marking method for wear time approximations, temperature filtering methods were generally robust. The same environmental conditions that negatively impacted the examiner marking method, such as leaving spectacles in a warm car, also impacted the filtering strategies, although with a greater magnitude of error. If the subject who yielded these conditions is dismissed from consideration (Table 2.4, Subject 9), then the filtering methods performed similarly to the examiner marking methods for the large majority of weeks for the remaining subjects. Utilizing the filtering methods may be appealing in that software to plot and mark temperature versus time would not be required, and the speed to analyze findings would be improved. However, some consideration of what range of temperatures to use for classification of spectacle wear is required and could vary by individual or climate. For example, our data found warmer mean ON temperatures for female subjects (Table 2.1) which we predict is related to hairstyles covering the sensor during spectacle wear. The individual temperature filtering method eliminates complications from these inter-subject ON temperature variations; however, to utilize this method, examiners would need to pre-collect temperature data by requiring subjects to wear their spectacles with the sensor for a known time interval in order to calculate that individual's mean ON temperature. In this study, the group mean filtering method had similar median percent error than the individual mean filtering method (Tables 2.2 & 2.4), suggesting that the effort to obtain individual filtering ranges may not yield a meaningful improvement in accuracy. The group mean temperatures used for temperature filtering in this study (28.6^0 to 35.5^0 C) can likely be utilized by investigators deploying the device in similar environmental climates and avoid the need to pre-collect temperature data to define a filtering range; however, further studies are needed to evaluate the data logger's performance in colder climates to determine if the outdoor temperature affects recorded ON temperatures and the recommended filtering ranges.

It is important to consider the level of accuracy when utilizing an objective technique, such as a temperature sensor to monitor spectacle wear. It is expected that most investigators desiring to monitor spectacle wear will prescribe full-time wear, and thus we are most interested in whether or not the sensor can be used to determine if a subject wore their spectacles for the majority of the day. In this study, we evaluated the success of the data logger to monitor spectacle wear by percent error and viewed less than 20% error (at least 80% accuracy) as a reasonable threshold for success. Assuming full time wear is 16 hours per day, a device with 80% accuracy would indicate that the subject wore their spectacles between 12.8 and 19.2 hours. Although this range is large, either end indicates that the subject wore their spectacles for the large majority of the day, indicating good compliance with spectacle wear. We can also consider the subject that did not wear spectacles at all for a given day. If the device over-estimated and revealed the subject wore their glasses approximately 3 hours that day (a number derived from the level of over-estimation observed with a 20% over-estimation error on a full-time wearer), we would know this is overall poor compliance. Ideally, this device would perform without over-estimation, a complication that can be minimized by avoiding environmental spectacle storage situations in which the spectacles reach body temperature.

Although our ultimate goal is to utilize this device to determine whether subjects wear their spectacles full-time, this particular study was designed to evaluate the robustness of the data logger to monitor spectacle wear under a large variety of wear conditions. Subjects were specifically encouraged to adopt a part-time spectacle wear schedule (either alternating with secondary spectacles, contact lenses, or no spectacles), and in particular encouraged not to be predictable in their spectacle wear throughout the week. Subjects 9 & 19 were excellent adopters of this instruction and provided great information for future device deployment. Specifically, both subjects consistently had short intervals of spectacle wear and higher magnitudes of percent error for all methods approximating spectacle wear time. The highest percent error by examiner marking occurred for subject 9 during week 3 (160% error) with examiner 2 approximating 47.2 hours for that full week versus the 18.2 hours of subject reported wear. Averaging this total wear time over 7 days, examiner 2 approximated 6 hours of average wear per day versus the subject's reported 3 hours of wear. Even in this worst case example of examiner over-estimation, neither the examiner's approximation of wear, nor the subject's actual wear reached a level representing full-time spectacle wear, and thus the data logger still performed acceptably in categorizing the wear behavior of this subject. The behavior exhibited by subject 9 is characteristic of someone who only wears spectacles for specific tasks, in this case driving, and is not consistent with the behavior of someone who is completely non-compliant with prescribed full-time wear, who would instead likely leave their spectacles at home or in a backpack, both of which conditions were easily distinguished from true wear time.

The results of this study show similarities to that of Januschowski et al. and Maconachie et al., in that all three sensors provided an accurate, safe, and relatively comfortable means of approximating spectacle wear time.[10, 11] The studies by Jauschowski et al. and Maconachie et al., served as a proof of concept by deploying the sensor to 3 and 4 subjects working within their study groups. Our study enrolled 50 subjects from both the college and the community in an attempt to collect data with a large variety of wear patterns and environmental activities, and to evaluate subject tolerance to two mounting strategies in a large sample. Although successful in general, Januschowski et al. did report the TheraMon microsensor had difficulty discerning spectacle wear with environmental temperatures between 33-37^{0} C. Our data logger did not have the same difficulty with environmental temperatures in that range, with the exception of the subject who routinely left her spectacles in a parked car, a behavior that could be discouraged of future study participants wearing the data logger. Other considerations regarding these sensors are the cosmetic acceptability. The glasses dose monitoring system by Maconachie et al. is not cosmetically favorable due to the bulky nature of the device which could lead to poor compliance. While the TheraMon microsensor is smaller and likely more cosmetically acceptable than the SmartButton Data Logger, the TheraMon microsensor is limited to availability in Europe, and thus the SmartButton may be more readily adopted by investigators outside of Europe. Now that there is evidence of three different thermosensors that can be successfully used to monitor spectacle wear, it appears that the methodology presented in this study can be applied to a variety of thermosensors, as long as the device has an adequate sampling density and is accurate.

One limitation in our study is the fact that there is potential for an inherent error in matching subject logs and temperature plots due to the fixed sampling rate of the thermosensor. The thermosensor only provided data every 5 minutes, but subjects were asked to report wear times to the minute and thus when merging the subject logs with the temperature data, there were many instances where subject reported times did not precisely match the time stamps of the temperature data. Data were reconciled to each other with the rule that a reported ON transition that fell outside of a 5 minute interval would be listed at the next closest time. For example, if the thermosensor recorded data points at 6:30 and 6:35, but the subject reported an ON transition at 6:33, the spectacles would be marked OFF for 6:30 and ON for 6:35, leading to an error of 2 minutes. Given the relatively small number of spectacle ON/OFF transitions over the course of a week of subject wear and the minimum required wear time of 10 hours, these small errors should not significantly impact the overall findings of the study.

Future goals to most effectively utilize the SmartButton Data Logger for spectacle monitoring include efforts to modify the silicone mount type to improve cosmesis while still maintaining good comfort and durability. Secondly, while the method of examiner marking to approximate spectacle wear is not overly burdensome (on average 2-3 minutes to mark one week of subject wear), future efforts include modification of the *Temperature Log Viewer* software to automate plot marking based on both temperature ranges and patterns in the data (e.g. sharp increases or decreases signaling on and off spectacle transitions), eliminating the need for examiner manual marking, or perhaps reducing it to a reconciliatory effort. Lastly, our study only enrolled adult study participants. Evaluating the tolerance of children to wear the SmartButton data logger will be important for its adoption in studies such as those to monitor spectacle wear for the treatment of amblyopia.

Conclusions

The SmartButton Data Logger is a promising device to monitor spectacle compliance in patients. The device was well-tolerated by adult subjects for a four week period, and provided approximations of spectacle wear time with median percent error of less than 10% and overall error falling within 20% for the large majority of subjects.

Appendix

Introduction

The following are the results of preliminary studies performed on the *ACR Systems Smartbutton Data Logger Temperature Sensor*. These studies were used to determine the capabilities of this device. The studies were performed to determine: 1) the variability in sampled temperatures between multiple data loggers in the same environment, 2) time course of the device to reach peak and minimum temperatures when moved from extreme hot to extreme cold, 3) the difference in recorded temperature data when the device is worn on spectacles versus in a pocket or on a stationary surface in the same environment (both hot and cold), and 4) whether or not one side of the device was more sensitive to temperature changes than the other. For reference, the dashed horizontal line in any figure represents human body temperature (37 degrees Celsius), which was predicted to be the temperature that may designate spectacle wear. Lastly, a sample of a subject's spectacle wear log and a temperature versus time plot from the *Temperature Viewer* software are shown.

Variability between Devices

To determine the variability in recorded temperature between multiple devices, 25 thermosensors were set to sample the temperature every 5 minutes for a 17 hour period while left on a desk in the laboratory overnight. As can be seen in figure A1, there was limited variability between devices, with all temperatures within 1 degree Celsius for the entire 17 hour period.

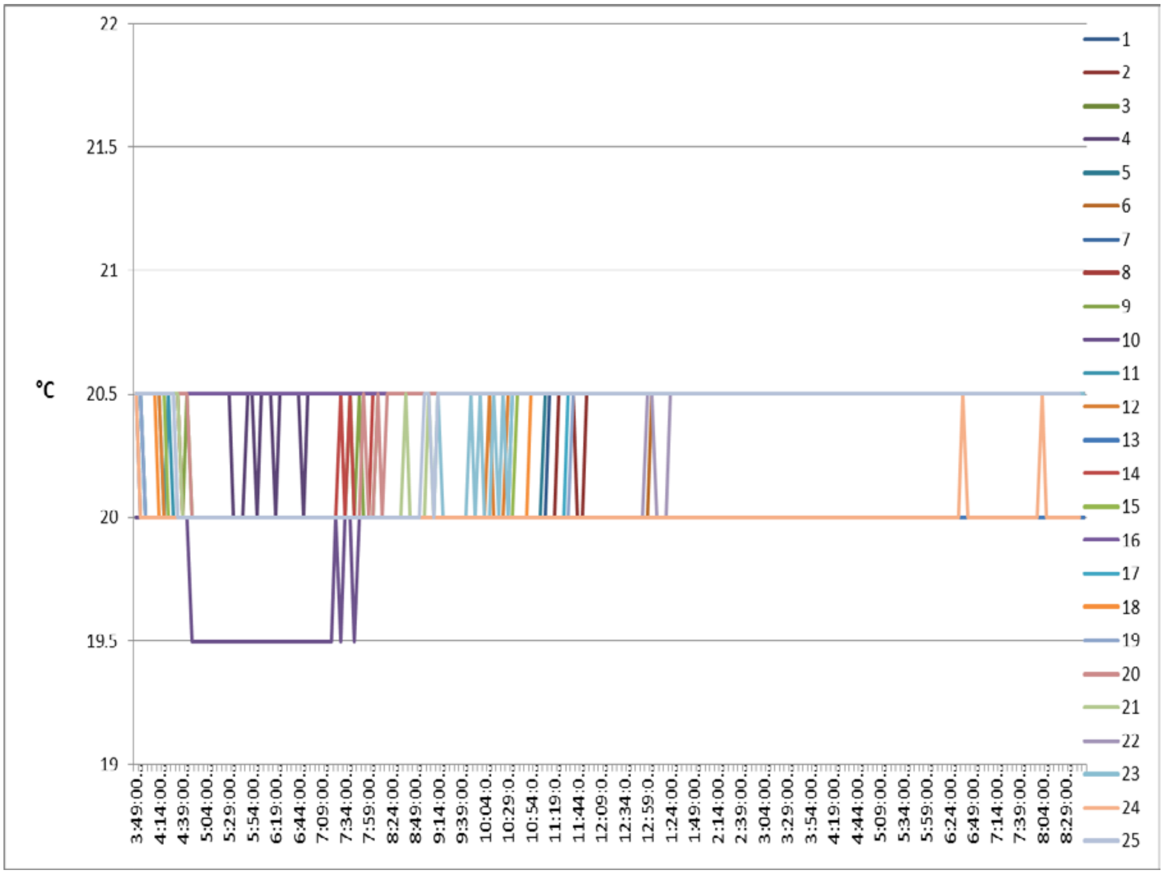


Figure A1. Temperature recorded over 17 hours by 25 individual thermosensors in the same environment. Each colored line represents a different thermosensor sampling temperature every 5 minutes.

Peak Temperatures

In an effort to determine how quickly the thermosensors could reach peak minimum and maximum temperatures, the following experiment was conducted. Six thermosensors, two 'cold' controls (Sensor 1 & 2) left in a refrigerator freezer to simulate peak minimum temperature, two 'hot' controls (Sensor 3 & 4) left outdoors in direct sunlight to simulate peak maximum temperature, and two test sensors (Sensor 5 & 6) were moved in alternating fashion from the same refrigerator freezer to outdoors in direct sunlight every 20 minutes. All devices were set to sample temperature every 5 minutes. As can be seen in figure A2, thermosensors 5 & 6 reached peak minimum and maximum temperatures within 15 minutes when moved from direct sunlight into the freezer. Although 15 minutes may seem like an extended period of time, this represents only three samples of temperature at a 5 minute sampling rate to reach steady state. It should also be noted that such drastic changes in temperature are unlikely to be experienced by human subjects during daily activities and the sensor itself may be insulated by proximity to the body when worn on the spectacles. It should also be noted that the transition from one extreme to the other can be pinpointed at the first 5 minute sample point based on the sharp increase or decrease in the trace.

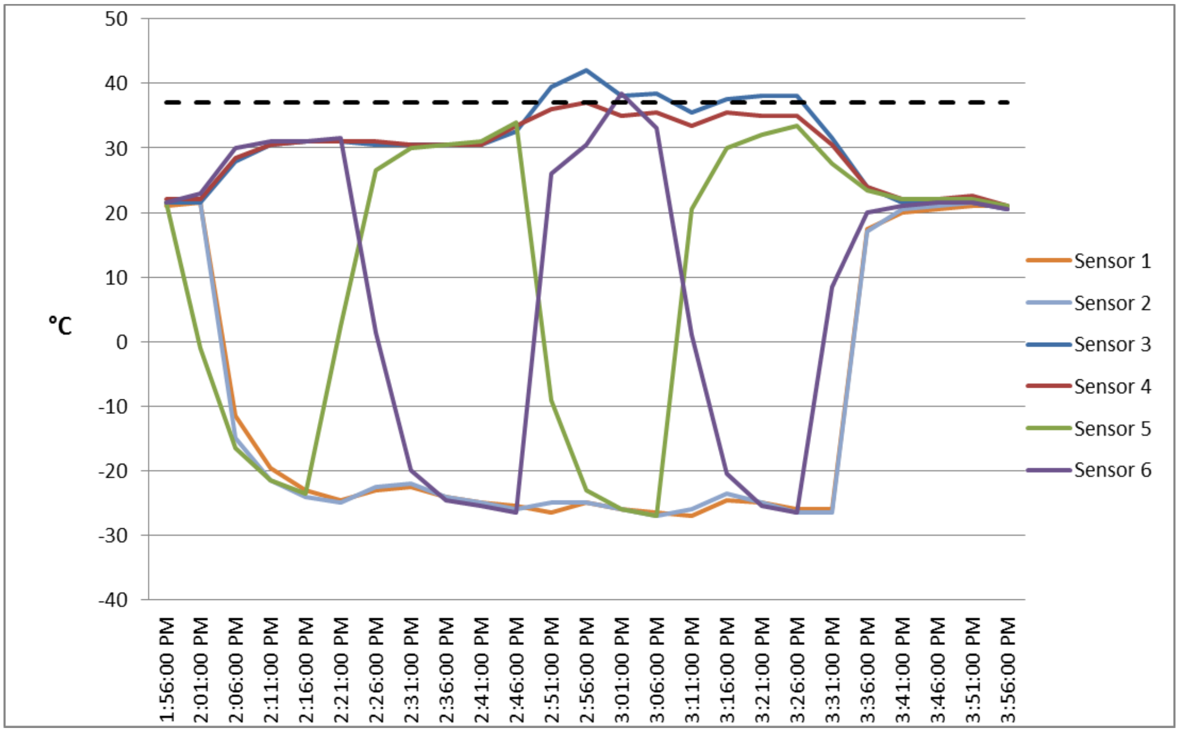


Figure A2. Time course to reach extreme minimum and maximum temperatures. Sensors 1 & 2 were 'cold' controls kept at minimum temperature, sensors 3 & 4 were 'hot' controls kept at maximum temperature, and sensors 5 & 6 were moved from hot to cold every 20 minutes. The black dashed line at 37 degrees Celsius represents average human body temperature.

Spectacles versus Surface

The follow experiment sought to determine if there was a difference in recorded temperature between a thermosensor mounted on spectacles using the silicone mount versus a sensor laying on a flat surface in the same environment. Both sensors were set to sample temperature every 5 minutes. The environmental conditions varied with the time reported on the x-axis, starting inside an air conditioned building (2:01-2:07), to outdoors in direct sunlight (2:07-2:37), to outdoors in a shaded area (2:37-3:13), to back indoors (3:13-4:01). Outdoor temperatures ranged from 34 to 38 degrees Celsius, based on internet weather reports for that day and time. When comparing temperatures logged by the sensors in the different environments, the initial 10 minutes post environmental transition were not included in the analysis. The mean temperatures were significantly different for both indoor ($p<0.001$, spectacles 30.5 ± 0.6 ℃, surface 22.5 ± 1.5 ℃, paired t-test) and outdoor data ($p<0.05$, spectacles 33.4 ± 3.5 ℃, surface 30.2 ± 8.3 ℃, paired t-test). As can be seen in figure A3, although the patterns of both traces appear the same, the thermosensor mounted on spectacles had a smaller range and seemed to be regulated closer to body temperature both indoors and outdoors.

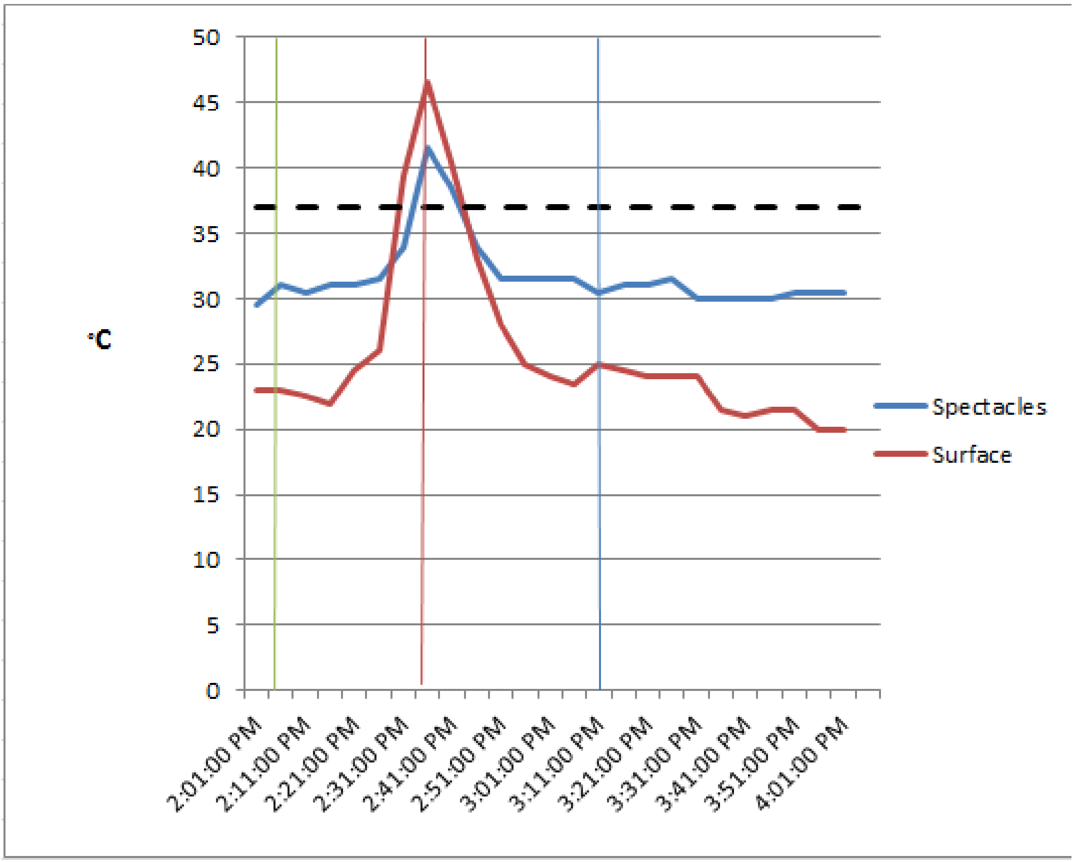


Figure A3. Temperature recorded by a device mounted on spectacles and another device laying on a flat surface in both indoor and outdoor environments. Both devices were moved simultaneously from indoors to outdoors in direct sunlight (green vertical line) to outdoors in shade (red vertical line) to back indoors (blue vertical line).

Spectacle versus Surface and Pocket

The following experiment sought to determine if there was a difference in recorded temperature of a thermosensor mounted on spectacles using the silicone mount versus a thermosensor stored in a pants pocket of the same individual wearing the spectacles versus a thermosensor laying on a flat surface. The study was conducted under cold conditions of a walk-in refrigerator set to 5 degrees Celsius with all 3 thermosensors set to sample temperature every minute. All thermosensors entered the walk-in refrigerator at 1:10 (time depicted on x-axis), the spectacle mounted thermosensor was initially uncovered to start the experiment and then covered by a hooded sweatshirt at 1:25, and all thermosensors exited the walk-in refrigerator at 1:40. Visual inspection of the data made it clear that there is a large difference in recorded temperature between the thermosensor laying on a flat surface and the other two thermosensors. The remaining three conditions (sensor in pocket, sensor on spectacles, sensor on spectacles covered with a hood) were compared by one-way ANOVA. For the sensor in the pocket, the temperatures recorded five minutes after entry into the refrigerator until departure at 1:40 were included in the analysis (24 samples, mean = 17.2 ± 0.3 ℃). For the pre-hooded condition, temperatures recorded five minutes after entry into the refrigerator until the hood was placed on were included in the analysis (10 samples, mean = 21.8 ± 0.3 ℃) and for the hooded condition, temperatures recorded five minutes after the hood was placed over the sensor until departure from the refrigerator at 1:40 were included in analysis (10 samples, mean = 24.9 ± 0.2 ℃). There was a significant difference in recorded temperature between all conditions (one-way ANOVA $p<0.0001$ with post-hoc Tukey) As can be seen in figure A4, the spectacle mounted thermosensor recorded temperatures consistently higher than both the pants pocket sensor and the sensor on the stationary surface. Although the pocket provided insulation for the sensor, the temperatures were significantly lower suggesting that it is possible to discriminate between pocket and spectacle wear utilizing temperature recordings. However, in a cold environment, the exposed spectacle sensor versus the hooded sensor also differed significantly indicating that subjects wearing head coverings or with long hair may have higher ON spectacle temperatures than those whose sensor is not covered during spectacle wear. It should be noted that in this cold environment, none of the sensors reached levels of human body temperature (37 °C).

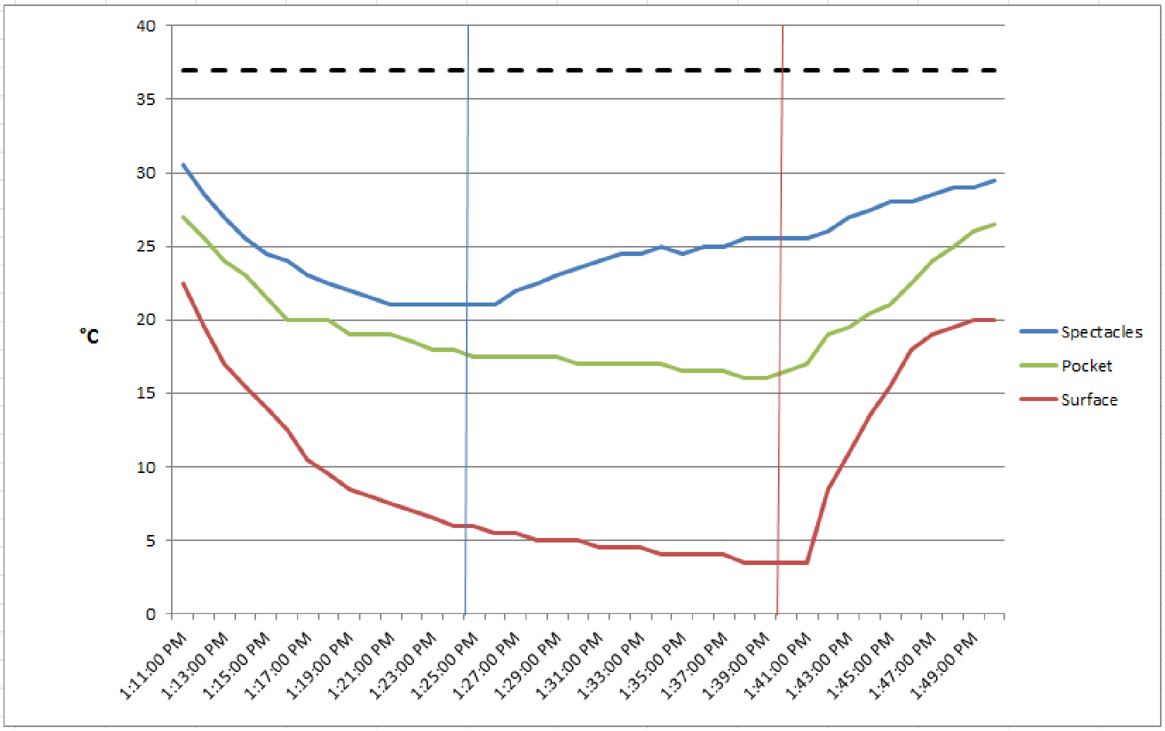


Figure A4. Temperature recorded by 3 devices in cold conditions of a walk-in refrigerator set to 5 degrees Celsius. One sensor was mounted on spectacles worn by an individual, another was place in the pants pocket of the same individual, and the last was laying on a flat surface. Blue vertical line represents when the hooded sweatshirt was worn, red vertical line represents when all thermosensors exited the walk-in refrigerator.

Device Orientation

The following experiment was performed to evaluate potential differences in recorded temperatures between each side of the thermosensor device. One side of the device is labeled with the *ACR Systems* logo and will be referred to as the logo side, and the other side of the device is labeled with a serial number and will be referred to as the number side. Two sensors were held approximately 2 inches from an incandescent lamp in the palm of a hand with either the logo side up or the number side up for the first 15 minutes (1-15 on x-axis). After the first 15 minutes, the devices were both placed on a flat surface the same distance from the lamp for an additional 15 minutes (16-30), without changing the side facing the lamp. Both devices were set to sample temperature every minute. As can be seen in figure A5, there was a significant difference in recorded temperature for the sensors in the palm of hand ($p=0.02$, logo up 38.2 ± 8.0 °C, number up 35.4 ± 4.2 °C, paired t-test) and for the sensors laying on a flat surface ($p<0.0001$, logo up 52.6 ± 7.3 °C, number up 41.4 ± 5.1 °C, paired t-test). This data suggests that there is less variability and temperature is maintained closer to body temperature when the numbers are out and the logo is in contact with the body. In consideration of these results, when deploying this device to subjects the numbers will be facing out with the logo side closest to the body.

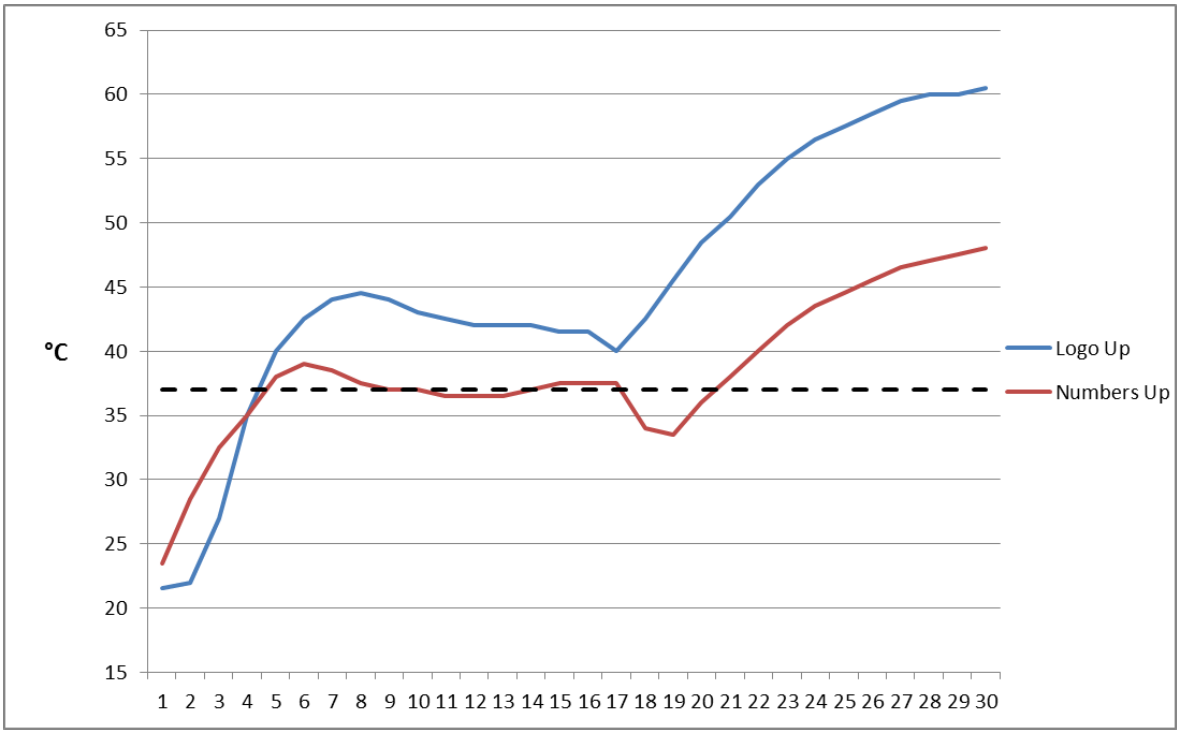


Figure A5. Temperature of two devices oriented in opposite directions placed 2 inches from an incandescent lamp for a 30 minute period. The devices were held in the palm of a hand for minutes 1-15 and laying on a flat surface for minutes 16-30.

Example of Subject Log

Date	Time	Specs ON / OFF	Hair Covered Sensor YES / NO	Additional Details (environment/activity)
6-5-14	7:35 AM – 11:45 AM	ON	YES	Woke up, drove to work, sat in office
6-5-14	11:45 AM – 12:00	ON	YES	Walked outside in 100 degree heat
6-5-14	12:00 – 12:30	ON	YES	Lunch inside
6-5-14	12:30 – 12:45	ON	YES	Walked outside in 100 degree heat
6-5-14	12:45 – 6:30 PM	ON	YES	Work, drove home
6-5-14	6:30 – 7:30 PM	ON	NO	Exercised
6-5-14	7:30 – 8:00 PM	OFF		Showered – glasses on bathroom counter
6-5-14	8:00 – 9:00 PM	OFF		Glasses tucked in shirt pocket
6-5-14	9:00 PM	OFF		Glasses on nightstand

Figure A6. Example of a subject log with reported spectacle wear times and additional environmental details.

Temperature Viewer Sample

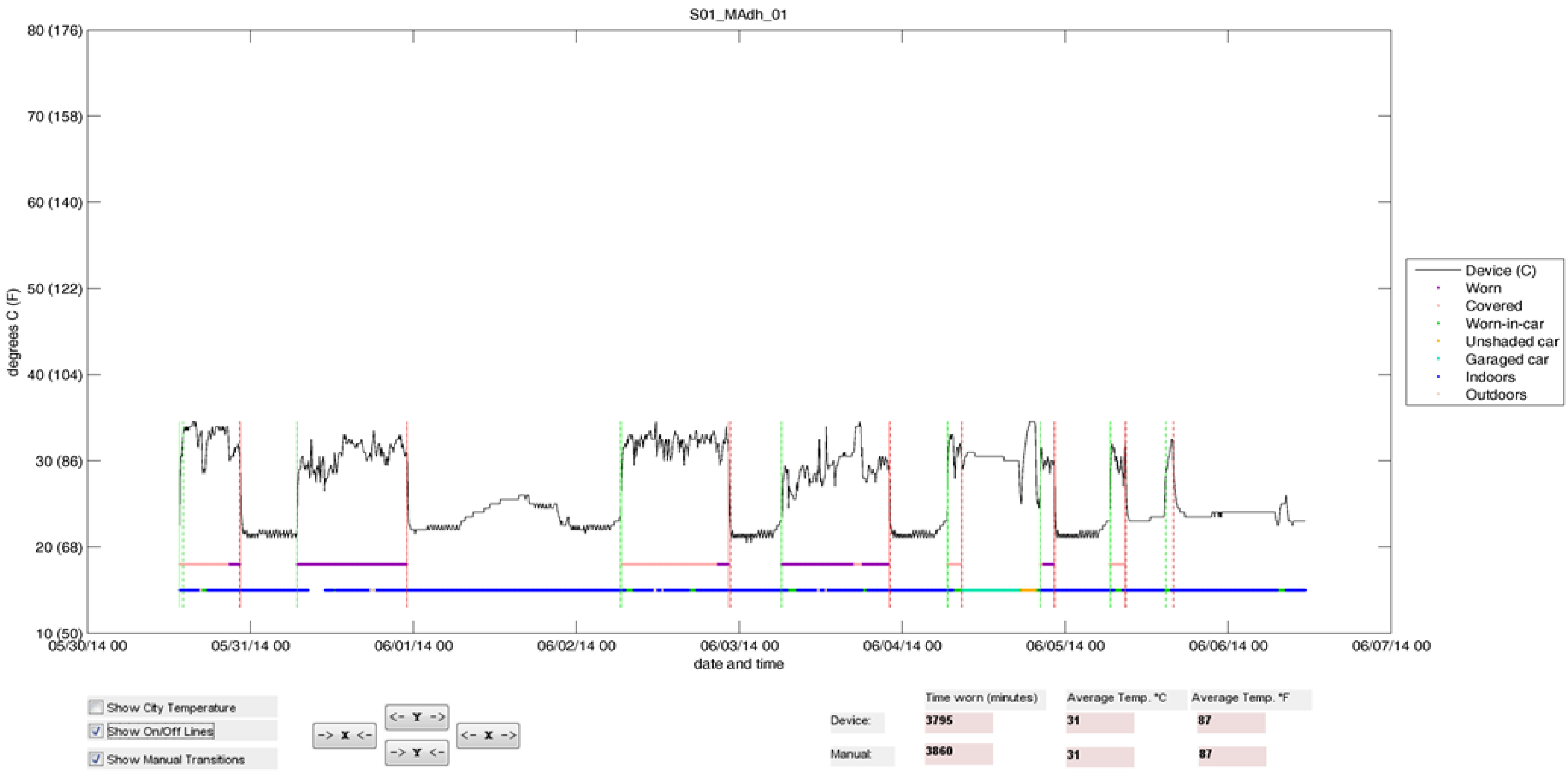


Figure A7. A sample of an examiner marked plot using *Temperature Log Viewer 1.2*. See Chapter 2 *Methods* for a full description of the temperature versus time plot.

Bibliography:

1. Cotter SA, Foster NC, Holmes JM, Melia BM, Wallace DK, Repka MX, Tamkins SM, Kraker RT, Beck RW, Hoover DL, Crouch Iii ER, Miller AM, Morse CL, Suh DW. Optical Treatment of Strabismic and Combined Strabismic–Anisometropic Amblyopia. Ophthalmology 2012;119:150-8.

2. Hwang DJ, Kim YJ, Lee JY. Effect and sustainability of part-time occlusion therapy for patients with anisometropic amblyopia aged ≥8 years. British Journal of Ophthalmology 2010;94:1160-4.

3. Jones-Jordan LA CM, Coffey B, Jackson JM, Manny RE, Rah MJ, Walline JJ. A Comparison of Spectacle and Contact Lens Wearing Times in the ACHIEVE Study. Clinical and Experimental Optometry 2010;93:157-63.

4. Vincent JEN, Satja +; Parry, Amy ++; Mladenovich, Derek [S]; Thein, Nyunt Naing [//]; Amendola, Paul R. [P]. Reported Wearing Compliance of Ready-Made Spectacles at 6 and 12 months. Optometry & Vision Science 2010;87:958-65.

5. Messer DL, Mitchell; Twelker, Daniel. Spectacle Wear in Children Given Spectacles Through a School-Based Program. Optometry & Vision Science 2012;89:19-26.

6. Courage ML AR, Reyno S, Kwa PG. Visual Acuity in Infants and Children with Down Syndrome. Dev Med Child Neurol 1994;36:586-93.

7. John FM BN, Woodhouse JM, Candy TR. Spatial Vision Deficits in Infants and Children with Down Syndrome. IOVS 2004;45:1566-72.

8. Woodhouse JM PV, Saunders KJ, Parker M, Fraser WI, Lobo S, Sastry P. Visual Acuity and Accommodation in Infants and Young Children with Down's Syndrome. Intellect Disabil Res 1996;40:49-55.

9. AB B. Development of a Questionnaire to Measure Spectacle Dependence: The Ohio State University; 2009.

10. Januschowski K, Bechtold TE, Schott TC, Huelber-Januschowski MS, Blumenstock G, Bartz-Schmidt KU, Besch D, Schramm C. Measuring wearing times of glasses and ocular patches using a thermosensor device from orthodontics. Acta Ophthalmol 2013;91:e635-40.

11. Maconachie GE, Farooq S, Bush G, Kempton J, Proudlock FA, Gottlob I. Association between adherence to glasses wearing during amblyopia treatment and improvement in visual acuity. JAMA Ophthalmology 2016;134:1347-53.

12. Horwood AM. Compliance with first time spectacle wear in children under eight years of age. Eye 1998;12:173-8.

13. Khandekar R, Mohammed AJ, Al Raisi A. Compliance of spectacle wear and its determinants among schoolchildren of Dhakhiliya region of Oman: A descriptive study. SQU Journal for Scientific Research - Medical Sciences 2002;4:39-43.

14. Fronius M, Chopovska Y, Nolden J, Loudon SE, Lüchtenberg M, Zubcov A, Pepler L. Occlusion treatment for amblyopia: assessing the performance of the electronic occlusion dose monitor. Strabismus 2006;14:65-70.