

Development and Evaluation of a Model of Human Comfort and Cognitive Ability for Moderate Differences in Thermal Environment

Shane T. Mueller^[0000-0001-7418-4592], Yin-Yin (Sarah) Tan, and Isaac Flint

Michigan Technological University, Houghton MI 49931, USA
shanem@mtu.edu

Abstract. Past research has established systematic effects of thermal stress on human comfort and cognitive performance. However, this research has primarily focused on extremes of temperature, ignoring moderate temperature ranges typically found in work environments and vehicles. Furthermore, models predicting the psychological impact of thermal environment have typically focused solely on perceived comfort, or when accounting for cognitive performance (e.g., Anno et al., 1996; Mueller et al., 2011) focused on in relatively extreme thermal conditions. As a consequence, there is limited empirical data, and no viable predictive models, for understanding the impact of moderate thermal stress on human comfort and performance. We report on an experimental study with 24 college-age participants that assessed cognitive performance (across a number of cognitive dimensions including manual dexterity, speed, dual-task performance, task switching, executive function, and attention), subjective measures of comfort and workload (including the CALM comfort scale, the affect grid, NASA-TLX, perceived effort, and other measures related to the thermal environment), and physiology (including heart rate, skin temperature, and breathing rate). Participants were tested during three 90-minute sessions in a controlled thermal chamber in which the temperature was either cool (15° C/59° F) room temperature (22.5° C/72.5° F), or warm (30° C /86° F), during which they completed repeated rounds of comfort ratings, cognitive task performance, and rest. Results showed strong responses in physiological and comfort measures over time to differences in thermal environment. However, the thermal environment had differential effects on cognitive measures, with some producing little impairment, and others showing increases or decreases over time that were moderated by thermal environment. The results were examined within a latent variable model that suggests that comfort alone is not an adequate proxy for performance, even for moderate thermal stressors, and more complex predictive models are needed.

Keywords: Performance prediction, mental stressors, Environmental conditions.

1 Background

1.1 Purpose

The goal of this paper is to discuss the design, methodology, and results of a study to investigate how variations in the thermal environment impact aspects of physiological state, comfort, subjective assessments of workload, and cognitive measures of performance. The goal of this project was to develop a statistical latent factor model (based on the Task-Taxon-Task model) that could account for the variety of effects that the thermal environment has on human outcomes, and collect data that demonstrated the utility of a latent factor model in accounting for human response to thermal environment. The practical significance of this study is to understand the human response to moderate variations in their environment—those typical of indoor spaces—to help support future intelligent climate control systems.

Traditional physiological models of thermal stressors on human outcome variables have focused on modeling a single outcome variable (e.g., comfort) based on variations in thermal environment. However, the comfort and performance outcomes are known to have differential and complex relationships to thermal environment. Consequently, traditional models are unable to account for these impacts *prima facie*. However, past data establishing this has focused on extreme thermal environments (e.g., extended physical labor under heat stress conditions in military settings), making the past models ill-suited for the expected smaller ranges of sedentary passengers within the typical thermal environment of vehicle designed space. Consequently, we conducted a new study which used multiple dependent measures under different conditions of thermal stress and time, to develop a latent-variables model to account for the differential effects on various outcome measures.

2 Method

2.1 Participants

Twenty-four participants were recruited voluntarily from the local university and community. The criteria for recruiting participants include adults (1) aged 20-40, (2) at least 4 years of driving experience, (3) normal or corrected-to-normal vision and no color-blindness, (4) no history of heat strain injury, seizure related issues, and (5) no claustrophobia or discomfort in small rooms. Participants fell within a range of gender-dependent height/weight requirements to be included in this project. Participants were required to wear long-sleeve, button-down shirt and long pants, and were instructed to not consume alcohol or drugs 24 hours before any study session. Each participant was paid \$12 for each hour of participation. All methods were approved by the MTU IRB.

2.2 Experimental Design

This study implemented a within-subject experimental design, with all participants completing the same set of tests in a counterbalanced order under different environmental conditions. The study took four sessions, each lasting about 2 hours. The first session was intended as practice and familiarization in order to counteract any learning effects and obtain baseline measures. The remaining three experimental sessions involved manipulated environmental conditions controlled with an environmental chamber: normal condition (22.5°C/72.5°F), cool condition (15°C/59°F), and warm condition (30°C/86°F). The order of these sessions was counterbalanced across participants. Each session was conducted either on different days or by at least two hours apart if different days could not be scheduled. Within each session, participants were assessed in four study cycles each lasting approximately 25 minutes. Each study cycle involved a sequence of cognitive behavioral tasks and self-assessments.

2.3 Procedures

Participants took part in a battery of cognitive tests, including the Anti-Saccade Test, Pursuit Rotor Test, Dual-task Test, Hick's Law Task, Switcher Task, and the Minnesota Dexterity Task (see Table 1 and Figure 1 for detailed descriptions). All computerized tests were implemented using a free, open-source software, Psychology Experiment Building Language (PEBL, [23,25]). The primary independent variables were environmental state (cycle and ambient temperature). The major dependent variables were performance measures, physiological states (i.e., heart rates, breath rates, and skin temperatures), and subjective user comfort and workload ratings.

Each of the four cycles within each session involved the same sequence of tasks. These started with comfort ratings (which lasted about three minutes), then five cognitive tasks (five computerized tasks, each taking 2-5 minutes to complete), followed by the NASA-TLX workload and comfort ratings (3 minutes to complete). The entire sequence took under 25 minutes, and when complete, the participant rested until the next round started (25 minutes after the previous). Tasks were selected based on several criteria: (1) they were likely to be sensitive to decrements in attentional and executive control, which have previously been found to be most highly sensitive to the distraction and discomfort caused by thermal stress; (2) they were relevant for in-vehicle environments, with analogs to many of the control processes and decisions drivers must make, as well as those facing non-drivers who may wish to engage in a focused work task while ignoring external stimuli; (3) they could produce reliable measures in a relatively minimal timeframe (permitting 25-minute cycles), and they could be repeated without inducing substantial learning effects. We used a set of self-report measures related to the thermal environment, perceived workload, and affective state. Physiological monitors recorded real-time heart rate, breathing rate, skin temperature, and several other non-invasive measures (e.g., posture and movement) throughout the sessions, with the Equivital SEM-02 system.

Table 1. Description and Rationale of the measures used in the present experiment.

Task Type	Task	Purpose
Behavioral and Performance	Pursuit Rotor [1,3]	Assess continuous motor control. Coupled with Minnesota Dexterity to provide dual-task manipulation.
	Minnesota Dexterity [24,26]	Assess dual-task aimed movement under dual-task conditions. Manual control.
	Anti-saccade [2,12,14,19,21]	Executive control; suppressing attention-capturing visual stimuli. Maintaining focused attention under distraction.
	Dual-task [8,29,30]	Executive control and task management; responding to multiple simultaneous stimuli.
	Hick's Law [16,28]	Rapid decision making; choose quickly among options with increasing set size.
	Switcher [4,18,21]	Executive control; switch decision making rapidly between rules. Task-switching for managing attention.
Physiological measure	Skin Temperature	Contact measure under clothing using Equivital SEM-02
	Breathing Rate	Belt-derived breathing rate using Equivital SEM-02
	ECG Heart Rate	Contact-derived measure using Equivital SEM-02
Subjective Self-report	Borg Scale/RSME [7,31]	Rating of perceived mental effort
	Comfort [9,13]	Rating of perceived comfort (CALM)
	Hot/Cold [17,27]	Assess perception of temperature.
	Wet/Dry Skin [17,27]	Assess perception of skin moisture (typically related to sweat and humidity)
	Preferred temperature [17,27]	Assess whether they would prefer a warmer or cooler room temperature
	Affect Grid [6]	Two dimensions: Pleasantness and Arousal
	TLX [11,15]	Rating of six dimensions of workload

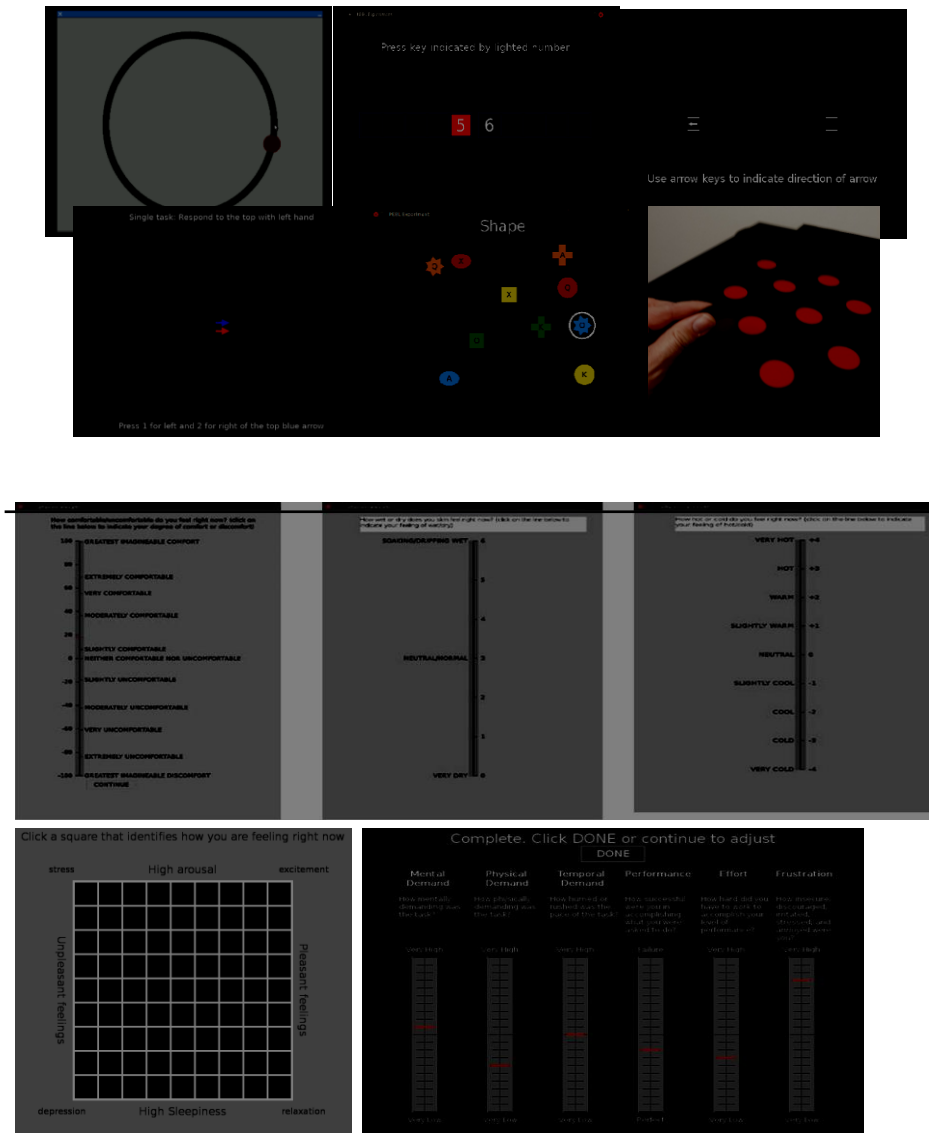


Fig. 1. Measures used in the study, including behavioral tasks (from upper left) Pursuit rotor, Hick's law task, Anti-saccade task, Dual-task test, Switcher task, and Minnesota Dexterity Task. Subjective assessment tasks used in the study, including CALM comfort scale, Wet/Dry, Hot/Cold, Affect grid and NASA-TLX.

Pursuit Rotor Task and Minnesota Dexterity task. The Pursuit Rotor Task employed a 4 (cycles: time in the environment) x 3 (temperatures: cold, normal, & warm) x 2 (level of difficulty: easy & difficult) within-subjects, repeated-measures design. The Minne-

sota Manual Dexterity Test was used along with Pursuit Rotor Task to examine participants' motor skills regarding the capability for eye-hand-finger movement and to create a scenario of dual task performance. All the participants completed the task with five trials in the same sequence. Each trial was set to be 30 seconds with either slow (1 rotation/ 10 seconds = 0.1 RPS) or fast (1 rotation/5 seconds = 0.2 RPS) mouse control rate. In the first trial, participants were instructed to just watch the task (slow) while flipping the disks as quickly as possible. Participants completed the pursuit rotor task alone on the 2nd (slow) and 4th (fast) trials; however, in the 3rd (slow) and 5th (fast) trials, they completed both tasks, executing the task and also flipping the checkers simultaneously.

Anti-Saccade Task. Similar to Pursuit Rotor Task, the Anti-Saccade Task employed a $4 \times 3 \times 2$ factorial within-subject design. The two levels of difficulty include easy level where participants only judged the direction of one single arrow, and difficult level where they had to use both hands to respond to the direction of two arrows.

Dual-task test. Identical to previous tasks, the Dual Task employed a $4 \times 3 \times 2$ factorial within-subject design. In the easy condition, participants were required to respond to the direction of one arrow, while in the difficult condition, they were asked to answer the directions of two arrows.

PEBL Hick's Law Task. PEBL Hick's Law Task was used to measure stimulus-response compatibility, motor control, and interference suppression. The task employed a $4 \times 3 \times 2$ factorial design as previous tasks. There are two levels of difficulty. Easy level has one (i.e., 5) or two (i.e., 5 & 6) numbers for participants to press on the keyboard corresponding to two fingers on their two hands, while difficult level had four (i.e., 4, 5, 6, & 7) or eight numbers (i.e., 2, 3, 4, 5, 6, 7, 8, & 9).

PEBL Switcher Task. PEBL Switcher Task involved two levels of difficulty. In the easy condition, participants were asked to follow one "match rule", such as color, shape, or letter, and searched for the symbol that matches the current symbol. Participants were required to follow more than one "match rule" in the difficult condition.

3 Results

3.1 Physiological Measures

We first examined the physiological measures across time in the three environmental conditions (see Figure 2).

Heart Rate. The warm thermal environment led to a higher overall heart rate. The cool and normal environments began (on average) with a lower heart rate, and the heart rate

reduced by about 10 bpm over the 90-minute session. In contrast, the warm environment began higher and remained elevated. On average, elevated temperature resulted in elevated heart rate that increased over time in comparison to neutral, while cool temperature produced no differential impact on heart rate in comparison to neutral. An ANOVA showed that there was a significant main effect of environment on heart rate ($F(2,42)=9.3, p<.001$; means for CNH = 96.9, 97.5, 102.7), and a significant temperature by time interaction ($F(6,43)=5.01, p=.01$).

Breathing Rate. There was a general elevation in breathing rate over the course of each session. There were substantial differences across individuals, some of which likely arose because of sensor and algorithm errors. Overall, there were not large systematic differences in breathing rate across conditions, which is expected in our study that involved low-impact cognitive testing. When cool and warm environment breathing rates were compared to neutral, no differential effects in comparison to neutral were observed. An ANOVA showed no significant main effect of environmental condition ($F(2,42)=.47, p=.6$; means for CNH = 12.5, 13.0, 11.5), and also no significant temperature by time interaction ($F(6,43)=.46, p=.6$).

Skin temperature. Measured skin temperature rose over the session for each of the conditions—even for the cool condition. This is likely a consequence of the skin temperature being measured underneath clothing in a sedentary environment, where the heat of the body generally increased the measure over time. Notably, in the warm condition, skin temperature was substantially higher even during the first cycle. When skin temperature of the two extreme conditions was compared to neutral, we found that the temperature elevation in the warm condition was relatively constant over time, and that the cool temperature condition produced a relatively constant decrease in skin temperature over time. An ANOVA showed a significant main effect of environmental condition ($F(2,42)=13.8, p<.001$; means for CNH = 32.8, 33.6, 34.7), and a marginally significant temperature by time interaction ($F(6,43)=2.66, p=.08$).

3.2 Subjective Measures

Subjective comfort and workload measures were assessed at the beginning of each session, between each cycle, and after the final cycle. Mean values for each condition are shown in Figure 3. The results indicate that environmental temperature condition has both immediate and time-varying effects on these ratings. A factorial ANOVA showed ambient temperature has significant main effect on CALM ($F(2,44)=19.9, p<.001$), Hot/cold ($F(2,44)=326, p<.001$), Skin wetness ($F(2,44)=18.9, p<.001$), Preferred temperature ($F(2,44)=129, p<.001$), Pleasantness ($F(2,44)=4.0, p=.02$), and Arousal ($F(2,44)=5.6, p=.007$), but not perceived effort ($F(2,46)=2.28, p=.11$). Moreover, most of the interactions were also significant, indicating generally that at least one difference got larger over time, including a significant interaction between cycle and condition for Perceived effort ($F(2,43)=3.4, p=.04$), CALM ($F(2,46)=24.3, p<.001$), Hot/cold ($F(2,46)=23, p<.001$), Preferred temperature

($F(2,46)=24$, $p<.001$), and pleasantness ($F(2,46)=9.3$, $p<.001$). There was no interaction for skin wetness ($F(2,46)=.86$, $p=.43$) or arousal ($F(2,46)=1.5$, $p=.228$). There were no significant main effects or interactions with the NASA-TLX scores.

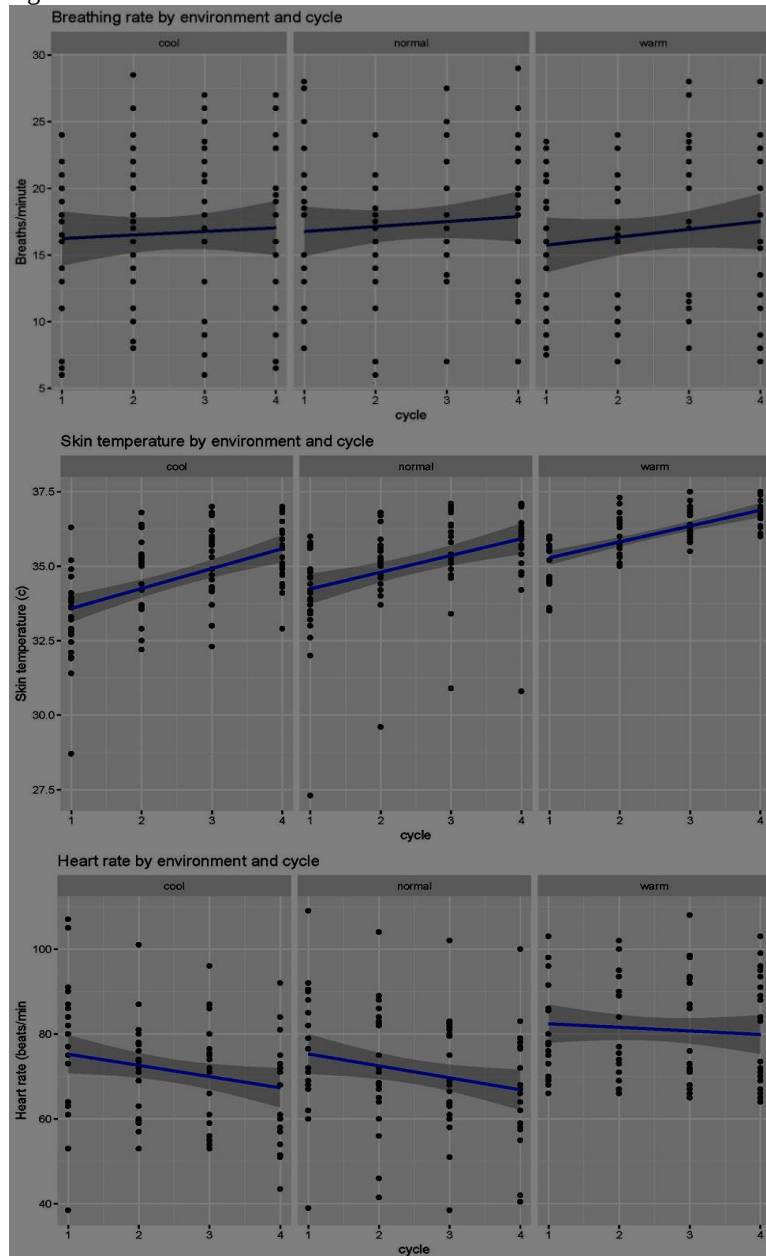


Fig. 2. Physiological responses for each environmental condition, across consecutive 25-minute cycles.

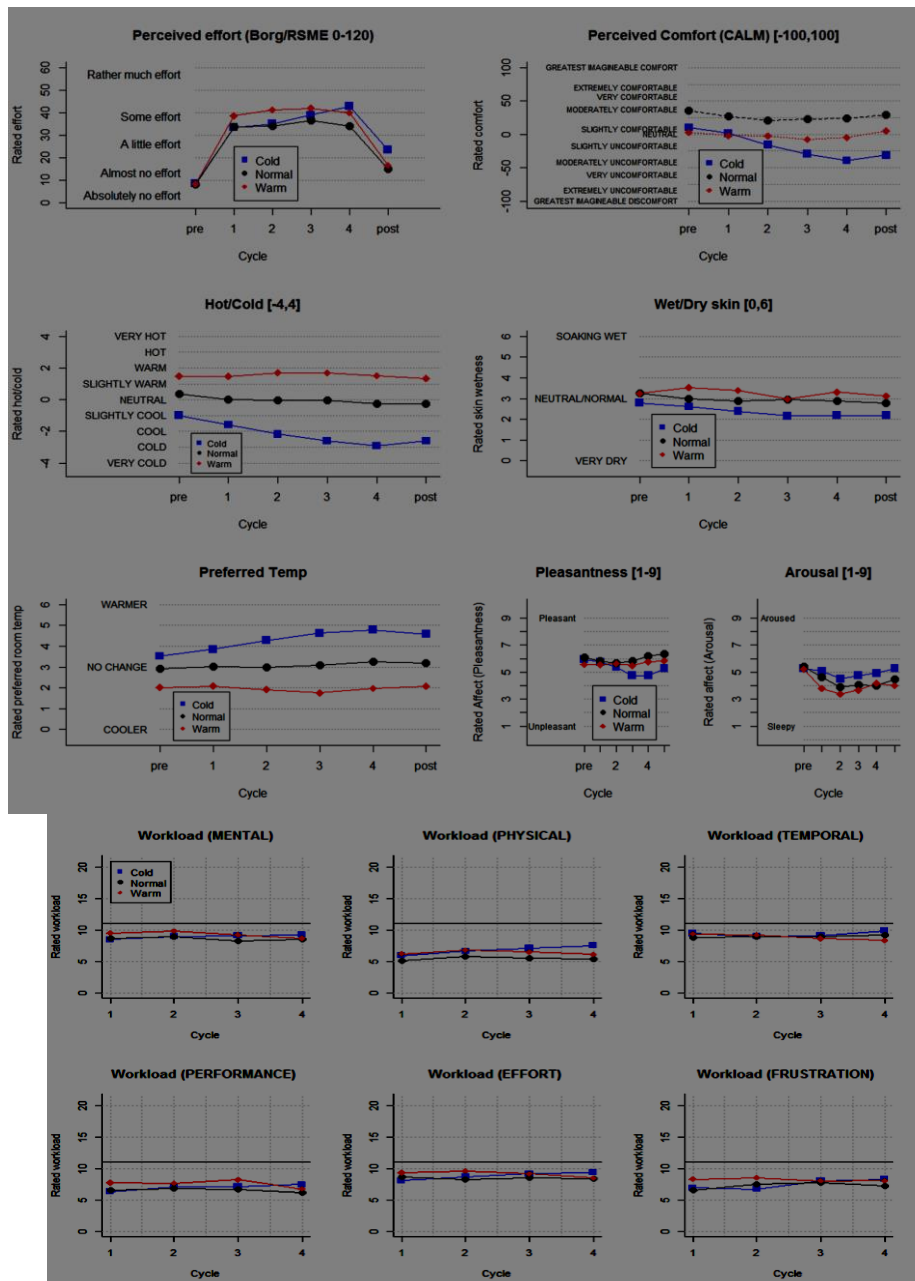


Fig. 3. Subjective comfort, workload, and affect ratings throughout four study cycles.

3.3 Behavioral/Performance Measures

Many of the tasks produced both response time and accuracy, and involved several levels of difficulty. Oftentimes, effects depended on difficulty, and we found some evidence that there may be interactions with gender (see [10]). For this presentation, we will examine the most general effects, via a standardized score. We first computed z-scores or BIS scores (which involve combining speed and accuracy with z-scores, and have been shown to be an unbiased measure of total performance [20]). Each participants scores were standardized across all levels of difficulty. Mean values are shown in Figure 4. Type-II ANOVA results from a mixed-effects lmer regression analysis in which environment was a categorical predictor and cycle is a numeric predictor, allowing a random intercept for each participant, and incorporating task difficulty as a main effect.

Pursuit Rotor. The top left panel of Figure 4 shows the results of participants' performance on the Pursuit Rotor Task. A type-II ANOVA showed significant effects of cycle ($\chi^2(1)=3.8, p=.05$), a significant effect of environment ($\chi^2(2)=16, p<.001$), and no cycle x environment interaction ($\chi^2(2)=.94, p=.63$).

Minnesota Dexterity Task. Simultaneous performance on the Minnesota dexterity task (top right panel) showed significant effects of cycle ($\chi^2(1)=47, p<.001$), environment ($\chi^2(2)=13.6, p<.001$), and a significant interaction $\chi^2(2)=13, p=.001$). Here, warm temperatures produced better performance, and this difference increased with time.

PEBL Anti-Saccade Task. The left panel of the second row of Figure 4 shows the BIS score for the anti-saccade task. A Type-II ANOVA revealed a significant effect of temperature ($\chi^2(2)=5.8, p=.05$) but there were no significant effects of cycle ($\chi^2(1)=.02, p=.87$), and no interactions between ambient temperature and cycle ($\chi^2(2)=1.2, p=.55$).

PEBL Dual-Task Test. The right panel of the second row of Figure 4 shows the BIS scores for the Dual-task test (averaged over all conditions). A Type-II ANOVA revealed that ambient temperature had a significant effect on BIS ($\chi^2(2)=15., p<.001$) a no significant effect of cycle ($\chi^2(1)=17.4, p<.001$), and a marginally-significant interactions between ambient temperature and cycle ($\chi^2(2)=4.67, p=.09$).

PEBL Hick's Law Task. The bottom left panel of Figure 4 shows performance for the PEBL Hick's Law Task. A Type-II ANOVA revealed significant effects of both cycle ($\chi^2(1)=11.7, p<.001$), but no significant effect of environment ($\chi^2(2)=3.3, p=.19$), or cycle x environment interaction ($\chi^2(2)=1.5, p=.47$).

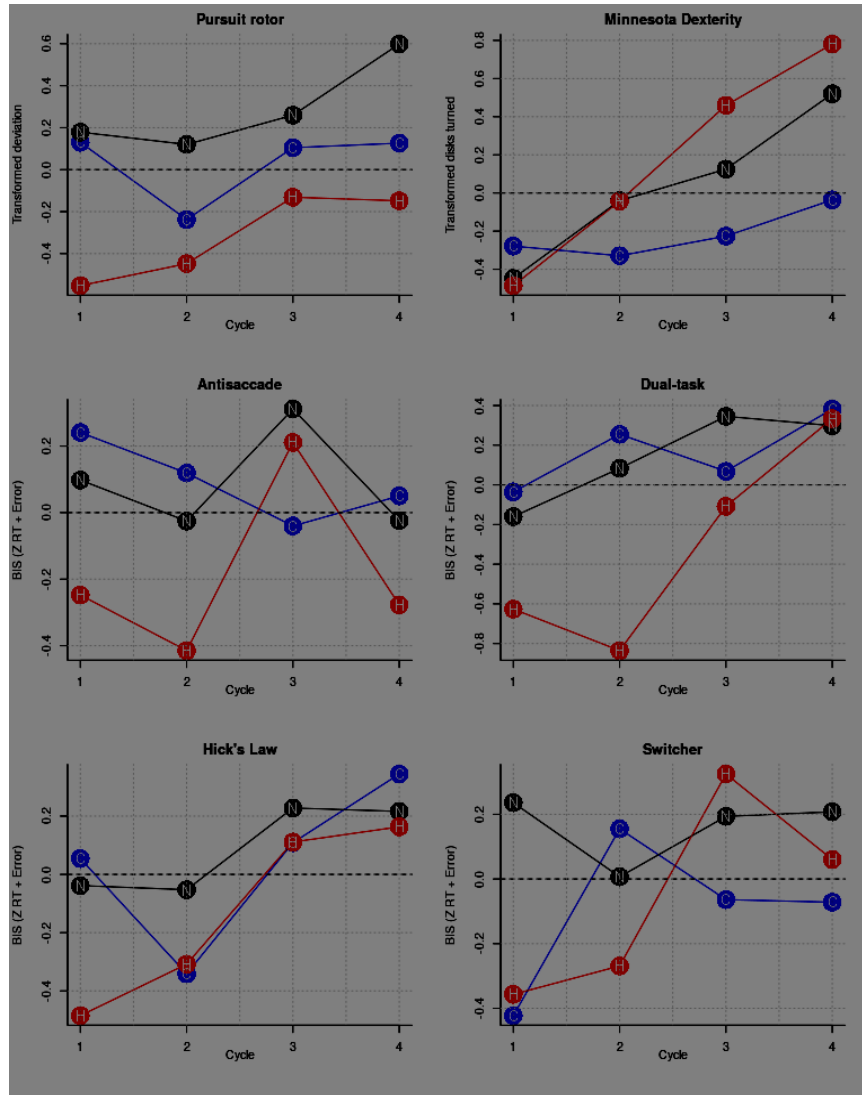


Fig. 4. Participants performance on the behavioral tasks. Performance is displayed as transformed z-scores, or on relevant tests, speed and accuracy were combined into a composite (BIS) measure. In each case, higher values indicate better performance.

PEBL Switcher Task. The bottom right panel of Figure 4 shows performance in the PEBL switcher task. An ANOVA showed marginally-significant effects of environment ($\chi^2(2)=4.45, p=.1$), a significant effect of cycle ($\chi^2(1)=3.6, p=.05$), and no cycle x environment interaction ($\chi^2(2)=2.1, p=.34$).

Summary. Overall, each test showed the impact thermal environmental. For most measures, the neutral condition was the best, and cooler and warmer conditions produced relatively worse performance. But these effects were not uniform across tasks, suggesting a complex relationship between thermal environment and performance.

3.4 Latent Variable Model (T3)

The Task-Taxon-Task (T3) model [5,24,26] was originally developed to model the impact of large environmental and chemical stressors on human performance. Typically, subject-matter experts would provide ratings of the importance of different skill taxa on different tasks, and use this as the basis for fitting performance impact functions for shared variance related to specific taxa. We adopted an alternative data-driven approach to fitting the T3 model. First, we computed performance using the normalized BIS scores, for easy and difficult conditions across each of the six tasks, and for each thermal environment and session. As a result of this coding, a value of 0 indicates average performance, a negative value indicates worse performance, and a positive value indicates better performance.

We used the *flexmix* library within R to fit a mixture of regression models, using as predictors cycle (1 continuous predictor), environment (3 levels), difficulty (2 levels), and a cycle x environment interaction; and using task as a grouping variable. Thus, results were clustered by task, and we used a BIC criterion to identify the optimal number of clusters or taxa. This produced five distinct clusters, but three were only distinguished by the magnitude of the linear effect of difficulty. This effect is interesting, but is mostly dependent on aspects of the task design, and so we selected a corresponding 3-cluster/taxon solution whose membership is shown in Figure 5.

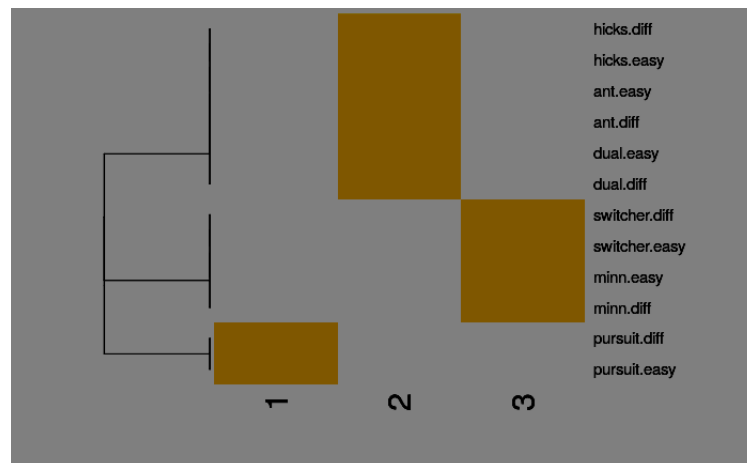


Fig. 5. Three-taxon solution based on behavioral measures. Here, easy and difficult measures group together, with three distinct groups.

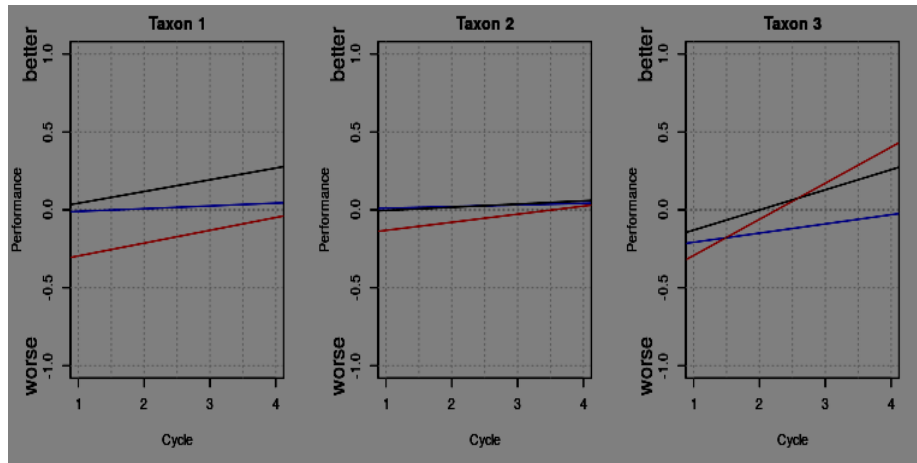


Fig. 6. Results of Taxon 1-3. For each taxa, we see distinct patterns of how performance changes over time, under different temperature conditions.

Each individual taxon is a group of measures that behave similarly in response to time and temperature. The predictions of each of these three models is shown in Figure 6. The patterns of data captured here suggest that overall, performance tended to increase with time under all temperatures. This means that participants were not getting fatigued, but rather were becoming better at the task. This is not simply a learning process, because this happened despite the fact that participants had an entire practice session prior to the testing in which they needed to meet a performance criterion, and the temperature conditions were counterbalanced.

For the first taxon, which was comprised of the two difficulty levels from the Pursuit Rotor task, normal temperature produced the best performance, which improved over time. Warm temperature produced the worst performance, which improved a similar amount. The cold temperature started at about the same level as neutral, but did not improve. This suggests that warm temperature had an immediate negative impact which lessened over time, and cool temperature had no initial impact but its negative effect (in comparison to neutral) increased over time. For the second taxon (several executive function tasks), there was little impact of cold temperatures, and no real improvement over time for the neutral or cold conditions. However, warm temperature again led to an immediate decrement—but this decrement diminished over time. This is a similar pattern as seen in Taxon 1, except with a lack of improvement for neutral. Finally, for taxon 3 (pursuit rotor and switcher), each condition starts out at about the same level, but warm improves the most, then neutral, and cold the least.

This modeling shows that in general, warm temperatures had the largest performance increase within the session. This suggests that our moderately-warm temperatures has an immediate effect of distraction and discomfort, but as people acclimate to the temperature, their cognitive performance returns to what is seen in neutral temperature. The cool condition is slower to have an impact; its initial impact is small, but in 2 of

the taxons, the difference between cold and neutral increases over time. This shows non-symmetric effects of temperature on cognitive function, and non-uniform effects over different cognitive tasks.

4 Discussion

In this study, our goal was to conduct a proof-of-concept study that established the non-unitary effects of thermal environment on human responses, in a highly-controlled laboratory environment. Our results showed impacts on physiology, subjective measures, and cognitive function. We examined several physiological outcomes. We found that heart rate was slightly elevated and did not change substantially in the warm-temperature condition, but neutral and cool temperature lead to lower and decreasing heart rate over time. In contrast, thermal environment had no substantial impact on breathing rate (although breathing rate increased slightly across the session), but thermal environment had a more substantial effect on skin temperature (with cool environment producing lower temperature and warm environment producing a higher temperature than neutral), all of which increased uniformly over time. So, even for relatively direct measures of physiological state, we observed three distinct patterns of results. These patterns (and others) were reproduced in comfort and human performance measures. Notably, self-report workload measures appear to have not been impacted by thermal environment or duration, but different measures of comfort and affect were impacted in different ways—often very strongly. In contrast behavioral measures are not as strong, but we still found impacts of thermal environment on most of the behavioral measures.

This suggests several lessons:

- The impact of thermal environment on performance cannot be uniformly predicted by physiological measures, because thermal environment has different impacts on different types of performance.
- Although subjective measures of comfort and thermal preference are very reliable, they also cannot account for the varied impacts of thermal environment on different measures of performance
- Behavioral measures are less reliable and subject to more individual variability than the subjective measures. Nevertheless, we were able to measure the negative impact of thermal environment on a number of performance tasks.
- Larger studies with more participants, across a larger variety of tasks, will be necessary to more fully map out the cognitive, performance, and comfort impact of thermal environment.

Although the details of the results are informative from a basic theoretical perspective, the larger implications are that thermal environment has non-unitary effects on human comfort and performance. We anticipate that as additional dependent measures and thermal environment manipulations are explored, this fact will remain true. As control of the thermal environment becomes more intelligent, this fact will impact the

basic approach to human-systems integration, control theory, and trust in intelligent automation.

For example, traditional in-vehicle environment control systems required the driver or rider to be instrumental in adapting the environment to their desired state. Until recently, even thermostatic controls (a technology more than a century old) were only available in luxury vehicles. Thermostatic controls work on simple principles that adjust the environment to match a minimum deviation from a set point, which is a well-understood problem, both from an engineering perspective and human factors perspective. Although there are many technical issues in dealing with engineering a proper environment and dealing with variations in anthropometry and individual differences, these problems are addressable because the optimization-control problem is fairly straightforward.

The next generation of systems will aim to work more intelligently. Rather than simply minimizing error to a set-point, they might aim to adapt the environment optimally for many other factors, including driver safety, comfort, work throughput, perceived workload, alertness, stress, and the like. This will require user models that incorporate the differential effects of thermal environment on these different outcomes, including individual differences (possibly stemming from gender, geography, or other factors). Furthermore, this presents several obvious challenges from a human-system integration (HSI) perspective.

For example, an appropriate model may be created that can predict the comfort or work throughput of a typical driver or rider, and these models may even be adaptable to individual differences based on preference or anthropometry. However, the system will likely not be able to infer the goals of the rider directly—for example, whether they want to work, or want to stay awake, or want to take a nap, may be supported by separate environmental-control modes in a vehicle. This means that there are issues in designing a system that a rider can control, give feedback to, change settings, and interrogate the internal state of.

But even getting to that point is a challenge for research and development, and may require substantially different control approaches. Existing naive models based on PID controller schemes may still be running at the core of the system, but these controllers will need to incorporate some specific models of the human within their control scheme. Just as image and audio compression libraries incorporate psychophysical models of human vision and hearing to best optimize the outcome, intelligent control models will need to incorporate models of human response to temperature in order to optimize human comfort and productivity. Furthermore, as we have demonstrated in this study, models of the physiology alone will be insufficient, because there will be no single physiological state that is optimal for all particular goals and outcomes.

Acknowledgements. This research was supported by a gift from the Ford Motor Company to the Michigan Tech Fund. We thank Megan Volaski and Lamia Alam for assisting in data collection, and John Elson for advice and guidance on the conduct of the research, and for comments on this and earlier drafts of this manuscript.

References

1. Adams, J. A. (1952). Warm-up decrement in performance on the pursuit-rotor. *American Journal of Psychology*, 65 (3), 404-414.
2. Ainsworth, B., & Garner, M. (2013). Attention control in mood and anxiety disorders: evidence from the antisaccade task. *Human Psychopharmacology: Clinical and Experimental*, 28(3), 274-280.
3. Ammons, R. B. (1955). Rotary pursuit apparatus: I. Survey of variables. *Psychological Bulletin*, 52(1), 69-76.
4. Anderson, K., Deane, K., Lindley, D., Loucks, B., & Veatch, E. (2012). The effects of time of day and practice on cognitive abilities: The PEBL Tower of London, Trail-making, and Switcher tasks. *PEBL Technical Report Series [On-line]*, #2011-01. <https://sites.google.com/site/pebltechnicalreports/>
5. Anno, G. H., Dore, M. A. & Roth, T. J. (1996). Taxonomic model for performance degradation in combat tasks. Pacific-Sierra Corporation Technical Report, Contract #DNA-001-90-C-0139.
6. Barrett, L. F., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current directions in Psychological Science*, 8(1), 10-14.
7. Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, 14 (5), 377-381.
8. Brown, R. G., & Marsden, C. D. (1991). Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. *Brain*, 114(1), 215-231.
9. Cardello, A. V., Winterhalter, C., & Schutz, H. G. (2003). Predicting the handle and comfort of military clothing fabrics from sensory and instrumental data: development and application of new psychophysical methods. *Textile Research Journal*, 73(3), 221-237.
10. Chang, T. Y., & Kajackaite, A. (2019). Battle for the thermostat: Gender and the effect of temperature on cognitive performance. *PloS One*, 14(5), e0216362.
11. Colligan, L., Potts, H. W., Finn, C. T., & Sinkin, R. A. (2015). Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *International Journal of Medical Informatics*, 84(7), 469-476.
12. Deuter, C. E., Schilling, T. M., Kuehl, L. K., Blumenthal, T. D., & Schachinger, H. (2013). Startle effects on saccadic responses to emotional target stimuli. *Psychophysiology*, 50(10), 1056-1063.
13. Gagge, A. P., Stolwijk, J. A. J., & Hardy, J. D. (1967). Comfort and thermal sensations and associated physiological responses at various ambient temperatures. *Environmental research*, 1(1), 1-20.
14. Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, 18(10), 1279-1296.
15. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*, 52, 139-183. North-Holland.
16. Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4 (1), 11-26.
17. Hultzapple, K., Hirsch, S., Venafro, J., Frumkin, S., Brady, J., Winterhalter, C., & Proodian, S. (2012). Chemical Protection Garment Redesign for Military Use by the Laboratory for Engineered Human Protection Years 2005–2011. In *Performance of Protective Clothing and Equipment: 9 th Volume, Emerging Issues and Technologies*. ASTM International.
18. Jersild, A. T. (1927). Mental set and shift. *Archives of psychology*, 14, 89, 81.

19. Levy, D. L., Mendell, N. R., & Holzman, P. S. (2004). The antisaccade task and neuropsychological tests of prefrontal cortical integrity in schizophrenia: Empirical findings and interpretative considerations. *World Psychiatry*, 3(1), 32-40.
20. Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs (?). *Behavior Research Methods*, 51(1), 40-60.
21. Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100.
22. Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, 7(3), 134-140.
23. Mueller, S. T. (2018). PEBL: The Psychology Experiment Building Language (Version 2.1) [Computer experiment programming language]. Retrieved from <http://pebl.sourceforge.net>.
24. Mueller, S. T. (2010). Adapting the Task-Taxon-Task methodology to model the impacts of chemical protective gear. *The 19th annual conference on Behavior Representation in Modeling and Simulation*, Charleston, SC.
25. Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods* (222), 250–259.
26. Mueller, S. T., Simpkins, B., Anno, G., Fallon, C. K; Price, O.; McClellan, G. E. (2011). Adapting the task-taxon-task methodology to model the impact of chemical protective gear. *Computational and Mathematical Organization Theory*, 17 (3), 251–271.
27. Pierce, J., Hirsch, S., Kane, S., Venafro, J., & Winterhalter, C. (2012). Evaluation of Thermal Comfort of Fabrics Using a Controlled-Environment Chamber. In *Performance of Protective Clothing and Equipment: 9th Volume, Emerging Issues and Technologies*. ASTM International.
28. Proctor, R. W., & Schneider, D. W. (2017). Hick’s law for choice reaction time: A review. *The Quarterly Journal of Experimental Psychology*, 71 (6), 1281-1299.
29. Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science*, 12(2), 101-108.
30. Wickens, C. D. (1991). Processing resources and attention. *Multiple-task performance*, London: Taylor & Francis..
31. Zijlstra, F. R. H. (1993). Efficiency in work behaviour: A design approach for modern tools. Doctoral Thesis. Retrieved from <https://elibrary.ru/item.asp?id=6863266>