Low Hand Temperature Predicts Individual Differences in Cognitive Adaptability During Cold Water Immersion

Timothy L. Dunn
Naval Health Research Center
140 Sylvester Road
San Diego, CA 92106-3521
UNITED STATES OF AMERICA

timothy.l.dunn12.civ@health.mil

Douglas Jones
Naval Health Research Center
140 Sylvester Road
San Diego, CA 92106-3521
UNITED STATES OF AMERICA

douglas.m.jones93.civ@health.mil

Parts of this paper, including data, have been previously reported in: Dunn, T. L., Jones, D., & Heaney, J. H. (2022). Cold entropy: Assessing individual differences in cognitive adaptability during cold stress. *Ecological Psychology, 34*(1-2), 19–47. https://doi.org/10.1080/10407413.2022.2055473.

ABSTRACT

Assessing warfighter cognition presents a challenging case, as processes often unfold in extreme environments. Cold stress, for example, from cold water immersion (CWI) produces well-known physiological responses, though the physiological–cognition interactions that occur during CWI are less understood. Assessing individual differences in cognitive adaptability using dynamical methods can help address current shortcomings. However, many methods aimed at indexing adaptability are impractical in extreme environments. Thus, a gap exists in the ability to index adaptability under heightened experimental noise and to identify informative performance/physiological correlates.

Purpose: *Assess variation in stability and flexibility across individuals during CWI using nonlinear dynamical methods and machine learning, and assess the predictive utility of physiology and baseline variables.*

Methods: *Fifty-seven warfighters (M_{AGE} = 26 years, SD = 5.31; M_{HEIGHT} = 176 cm, SD = 7.79; M_{WEIGHT} = 79.5 kg, SD = 10.52; M_{BMI} = 25.7 kg/m^2, SD = 5.8; and $M_{\%BODYFAT}$ = 17.9%, SD = 6.14) completing CWI training volunteered for the study. Participants were required to immerse their whole body in cold water for 10 minutes outdoors, reclothe, and rewarm passively while completing a simple reaction time task and a match-to-sample task throughout. Average air and water temperatures were held at approximately −5 °C and 1.5 °C, respectively. Sample entropy (SampEn) was computed over both tasks' response time (RT) time series to index stability. Hierarchical agglomerative clustering (HAC) was used to identify unique clusters based on SampEn. Autocorrelation functions (ACFs) were computed over RTs, and errors were considered jointly to assess flexibility across clusters. Lastly, baseline demographics (i.e., age, body mass index [BMI], and % body fat) and physiology (i.e., skin and core temperatures) were used to predict cluster classification.*

Results: *The HAC algorithm produced a two-cluster solution: Cluster 1 (n = 21) was associated with lower stability (SampEn) across both tasks relative to Cluster 2 (n = 36). Cluster 1 additionally showed weaker flexibility in terms of increased ACFs and more omission errors post-CWI across both tasks. A stepwise logistic regression was conducted predicting classification into the clusters. Pre-immersion hand temperature emerged as the only significant predictor, with individuals with lower hand temperatures being more likely to be classified into Cluster 1 (odds ratio = 1.2, 95% confidence interval [1.04, 1.39], p = .01). Using this model, we found that pre-immersion hand temperatures of 14.22 °C and below classified individuals into the low adaptability Cluster 1.*

Conclusion: *Our results demonstrate that nonlinear dynamical methods and machine learning can be used to identify underlying differences in cognitive adaptability during CWI. Importantly, the metric used to index stability (SampEn) models the full information of responses without any data trimming or transformations. Furthermore, low hand temperatures were associated with low stability and weak flexibility in responding (Cluster 1). A focus on warmer skin temperatures, given the susceptibility to rapid cooling of the hands, may improve cognitive adaptability.*

1.0 INTRODUCTION

A complete account of human cognition and behavior requires considering humans embedded in environments in which underlying processes develop and unfold (Clark, 2008; Hutchins, 2010). Operations in extreme environments are a unique case, where distinctive demand characteristics can lead to relatively basic tasks becoming difficult (Paulus et al., 2009). For example, soldiers' rifle assembly and disassembly is worse in extreme cold when initially trained in a relatively comfortable warm environment (Oksa et al., 2006). Furthermore, there exist additional complications with experimentation and measurement in extreme environments, compounding the problem of providing optimal explanation and prediction to decision-makers and those tasked with executing orders. To address these issues, the current study investigated effects of cold stress from cold water immersion (CWI) on adaptive behaviors through a novel approach of examining individual differences in response time (RT) complexity.

1.1 Physiology and Cognitive Behavioral Outcomes of Cold Stress

An extreme environment is defined as an external context that exposes individuals to demanding psychological and/or physical conditions and may have profound effects on cognition and performance (Paulus et al., 2009). Many of these environments are associated with general decrements to performance, though unique perturbations underlie each. For example, cold stress is known to cause specific physiological, psychological, and behavioral responses configured to maintain performance or wholly avoid cold environments. Cold exposure severity is dependent on ambient temperature, exposure medium (air, water), velocity (wind, water movement), duration, exposed surface area (partial or whole body), degree of thermal protection, and psychological state (anxious, fatigued, injured) of the individual being exposed (Castellani et al., 2005; Castellani & Young, 2016; Henriksson et al., 2009). Changes to any of these elements can drastically alter physiology and performance. For example, exposure to air temperature of 10 °C may cause psychological arousal and improve performance, whereas exposure to the same temperature in water impairs performance (Muller et al., 2012; O'Brien et al., 2007).

CWI is often presented in terms of the experiential process, which begins with the cold shock response. The cold shock response during the first minutes of immersion is characterized by hyperventilation, vasoconstriction, tachycardia, and uncomfortable cold sensations (Stocks et al., 2004; Tipton, 1989). Vasoconstriction, the initial response to prevent heat loss, is soon accompanied by shivering to increase heat production in an effort to maintain core temperature. When heat loss is greater than heat production, core temperature falls and, if the cold stimulus is not removed, hypothermia ensues (Danzl & Pozos, 1994). The continuum of physiological processes in response to CWI is well studied, but the same cannot be said for cognitive processes. The information that does exist is often conflicting, incomplete, or does not adequately describe cognitive processes, behavior, or the physiological-cognitive-behavioral interactions that occur during CWI (Martin et al., 2019).

1.2 Assessing Adaptability

The current literature on cognition during cold stress potentially paints an overly general picture of behaviors. Commonly reported averaged group and condition effects lead to granular predictability, such as "worse" performance on reasoning tasks when experiencing cold stress relative to a control (Pilcher et al., 2002). This approach, however, ignores critical aspects of data at the individual level (e.g., Balota & Yap, 2011; Draheim et al., 2019). Or, when no effect is reported as in the case of mathematics tasks during cold stress (Pilcher et al., 2002), there is no ability to account for non-linear patterns potentially present within the intervals of measurement (e.g., in between the pre- and post-measurement points). Moreover, it is possible for two processes to produce nearly identical means and variances but to also have very different underlying dynamics, such as differential temporal dependencies and/or patterns of fluctuations (e.g., Lipsitz & Goldberger, 1992). The stakes are high when humans must operate in extreme environments such as in extreme cold. Therefore, there is a critical need to move beyond basic and convenient methods in favor of more fine-grained modeling of the temporal dynamics of adaptability at the individual level in extreme environments.

Broadly, adaptive behaviors are those that benefit an individual in their environment (Barkow et al., 1992). Put simply, adaptive processes trade off to reduce "bad variability" and increase "functional variability" (i.e., flexibility) to a level of optimization (Hristovski & Balagué, 2020). A wide range of processes fall under adaptive behaviors, with stability and flexibility possibly being the most critical in optimizing performance. For example, an individual must decide whether to pursue their current behavioral policy, thereby maintaining stability, or explore different policies in favor of something more beneficial, demonstrating flexibility (Tardiff, 2019). Stability can be defined as the ability of a system to return to an equilibrium state after a temporary disturbance, where less fluctuation indicates higher stability (Holling, 1973). Within cognitive neuroscience, stability has largely focused on maintenance of working memory under distraction, including goals, tasks sets, and filtering out distracting information (Cools & D'Esposito, 2011). Such processes related to stability are highly adaptable as they support learning and long-term memory formation (Ranganath et al., 2005). Flexibility, in the current context, can be understood as the ability to change processing strategies when faced with novel and unexpected conditions (Cañas et al., 2003). Flexible processes are often linked to attention and executive control, including task switching and inhibiting pre-potent responses based on goals (Friedman & Miyake, 2017). Such processes allow individuals to rapidly reconfigure information to facilitate learning (Cole et al., 2013) and are linked to creativity (Chrysikou, 2018).

Both stable and flexible behaviors can be considered adaptive at many levels, however extreme levels of either can cause decrements. Adaptive gain theory (Aston-Jones & Cohen, 2005) implicates norepinephrine (NE) activity in the adjustments of exploitative and exploratory behaviors, where too-high levels of tonic NE (i.e., over-arousal) can lead to overly flexible and scanning attention in the pursuit of more rewarding alternative behaviors at the cost of a stable exploitation of known rewards (e.g., performing at an optimal level based on task goals). Performance may suffer due to scanning attention to task unrelated stimuli; Cohen et al., 2007), though high flexibility in attention may be beneficial long term if a more rewarding behavior is identified (Aston-Jones & Cohen, 2005). Similarly, overly rigid stability can undermine optimal performance. Decrements are often observed in a reduced ability to task switch (van Schouwenburg et al., 2010).

1.3 Present Investigation

Our aim for the present investigation was to assess variation in adaptability across individuals during CWI using non-linear dynamical methods and clustering, and link these behaviors to variations in cold stress responses (i.e., skin and core temperatures). Measuring performance "in the wild" is difficult given heightened uncertainty and noise most often caused by factors outside of a researcher's control. Thus, many methods aimed at indexing adaptive behaviors are impractical in extreme environments, such as those associated with CWI. A gap exists then in the ability to index stability, flexibility, and their performance/physiological correlates from a pragmatic standpoint. How, then, can we provide better explanation *and* prediction in extreme environments where poor performance holds real-world consequences? Assessing differential patterns in the non-linear dynamics of behaviors may prove fruitful (Van Orden et al., 2003).

We employed entropy analyses to the time series of individuals' RTs in two tasks to index stability. Within dynamical systems, entropy is a singular value indexing complexity over time (Young, 2003) that takes into account the *total information* of time series and has long been used as a model of natural systems (e.g., biology, Brooks et al., 1989; ecology, Harte, 2011; neuroscience, Friston et al., 2006; and distributed cognition, Chemero, 2009). From a theoretical standpoint, entropy provides an ideal metric to make predictions based on patterns of complexity. Work within medicine, biology, and physiology has demonstrated a consistent pattern of entropy decreasing due to perturbations (i.e., the loss of complexity hypothesis; Newell et al., 2006).

The focus of our analyses involves hierarchical clustering of individuals' entropy values for both tasks. It is known that large individual differences in executive abilities, such as stability and flexibility, exist (Friedman & Miyake, 2017). Moreover, assessing individual differences in RTs using alternative methods to basic difference scores has received renewed interest (Draheim et al., 2019). Therefore, computing entropy over RTs as an index of stability and classification metric provides a novel connection across these lines of differential research. We offer no a priori hypotheses about the number or characteristics of the clusters; this approach is data driven. However, subsequent analyses were conducted to assess differences across groups in cognitive flexibility. We specifically considered the autocorrelation function (ACF) of the RT time series and errors as indices of flexible responding. In human behaviors, high autocorrelation has been argued to be indicative of flexibility. For example, long-range autocorrelations are indicative of operations close to a critical point of transition (e.g., a heightened ability to switch across response types; Simola et al., 2017). The alternative account, where stronger autocorrelation is indicative of weaker flexibility, has been proposed as well (i.e., temporal dependencies limit the ability to rapidly reconfigure information and responses; He, 2011).

Lastly, we consider potential differences in body temperatures and baseline demographics as a function of clusters. Specifically, hand, skin, and core temperature were considered given their known influences on performance under cold stress. Extreme cold activates nociceptors and thermal receptors located in the periphery, causing sensations of pain and cold, respectively (Belmonte et al., 2009). These sensations, in turn, drive behavior modification and may disrupt task engagement caused by distraction (Enander, 1987). When hand temperature falls to approximately 15 °C, manual dexterity is impaired and progressively deteriorates as the hands become colder, ultimately leading to a complete loss of function (Heus et al., 1995).

Core body temperature reflects the body's overall heat balance. Several involuntary responses to cold stress, such as vasoconstriction (heat retention) and shivering (heat generation), attempt to maintain core temperature at or near 37 °C. However, if heat loss continues to be greater than heat retention and generation, the outcome is a reduction in core temperature and hypothermia (i.e., core temperature <35 °C), wherein cognitive processes can become significantly altered (Danzl & Pozos, 1994). These measurements, therefore, are critical to evaluate when examining aspects of performance during exposure to cold stress.

2.0 METHOD

2.1 Participants

Only fit-for-full-duty warfighters enrolled in the Cold Weather Medicine course at the Marine Corps Mountain Warfare Training Center (MCMWTC) were permitted to participate. We attempted to enroll as many participants as possible because our main aim was to identify unique groups through clustering. Sixty-four (64) active duty military personnel volunteered for the study. Data were collected across two winters with separate training groups; the first data collection occurred in 2019 (Run 1) and the second in 2020 (Run 2).

2.1.1 Exclusion Criteria

Laptops and wireless acquisition software used for collecting and collating the physiological data failed at several points during the exercise given the field environment. Some laptop cables became disconnected during participant movement and some participants moved out of range for successful wireless signal data transmission. Both situations were immediately corrected by research staff but resulted in minor data loss. Furthermore, behavioral data for some participants were never collected due to disengagement from the tasks. For example, some participants did not respond to the research team when asked to complete the task during rewarming likely because of weather conditions and/or discomfort. Listwise deletion of participants with any missing data would substantially reduce the sample to N = 34 (11% of all cells were missing physiological data due to equipment failure). To overcome this issue and meet the goal of examining potential relations between sample entropy (SampEn) and body temperatures, we chose to impute missing physiological data using the *missForest* package (Stekhoven & Bühlmann, 2012). This package uses random forests techniques trained on the observed values to predict missing values, including complex interactions and non-linear relations. *missForest* also produces an out-of-bag imputation error estimate without needing a test set or elaborate cross-validation (normalized root mean square error = 0.11 for the current data). We note that imputation was not used for missing RT data for either task because these data were used to calculate the main metric of interest (i.e., SampEn). Thus, seven individuals were removed from all analyses for missing RT data, resulting in a final N of 57 (M_{AGE} = 26 years, SD = 5.31; M_{HEIGHT} = 176 cm, SD = 7.79; M_{WEIGHT} = 79.5 kg, SD 10.52; M_{BMI} = 25.7 kg/m^2, SD = 5.8; and $M_{\mathrm{\%BODYFAT}}$ = 17.9%, SD = 6.14).

2.2 Design

The study employed a within-subject design, with all included participants completing the CWI and associated measures.

2.3 Apparatus

The behavioral tasks were designed and administered using E-Prime 3.0 software (Psychology Software Tools, Inc., Pittsburgh, PA, USA) and presented on a 25.7-cm Acer One 10 portable electronic tablet (Acer America Corporation, San Jose, CA, USA). Motor responses were recorded using a USB-tethered response box (Psychology Software Tools). Heart rate was monitored with telemetry chest straps (Polar Electro Oy, Kempele, Finland). Skin temperature was collected using wireless transmitting dermal adhesive patches

(VitalSense®, Phillips Respironics, Bend, OR, USA) placed on the skin over the right pectoralis muscle, medial deltoid, and mid-anterior thigh. Mean skin temperature was calculated using the Burton three-site formula (Ramanathan, 1964). Hand skin temperature was measured on the dorsal side of the non-dominant hand. Core temperature was measured wirelessly with an ingestible VitalSense temperature capsule. All temperature measurements were wirelessly transmitted and stored to portable VitalSense monitors worn by each subject. Lastly, the endpoint of each data collection time point was used for all temperature measurements. Utilizing the endpoint of immersion reflects the full impact that immersion had on temperatures.

2.4 Stimuli

The simple reaction time (SRT) task consisted of a 3.4 × 3.4 cm black star presented at the center of the screen for 40 trials per block. Each stimulus remained on screen for 1000 ms at which time the trial terminated if there was no response. Instructions stated to respond as quickly as possible when the stimulus was presented by pressing any button on the response box. The inter-trial interval ranged randomly from 600 ms to 3000 ms. No performance feedback was given. For the SRT, we operationally defined fast errors (or false alarms) as any RT <180 ms and slow errors (or misses) as any trial where a response was not given within the allotted window (within 1000 ms of stimulus presentation).

The match-to-sample (MTS) task consisted of an encoding and recall phase for every trial for 25 trials per block during the Run 1 data collection and 50 trials per block during Run 2 data collection. For Run 2 data collection, some individuals were fitted with electroencephalography (EEG) equipment during the CWI, which required an increase in the total number of MTS trials relative to Run 1. The EEG results will be discussed in a separate manuscript. A 4 × 4 tri-colored grid was presented for 2000 ms, followed by a 600–3000 ms inter-stimulus interval. The decision phase consisted of two grids presented to the left and right of the screen. Instructions stated to decide which of the two images was presented at the encoding phase. Participants were given a window of 3000 ms to make a response. The leftmost and rightmost buttons on the response box were mapped to a left match and right match, respectively. The remaining middle three buttons were mapped to a non-match response. Non-matching trials occurred 20% of the time within a block and were interleaved in random order. In the MTS, fast errors and slow errors were defined the same (within 3000 ms of stimulus presentation). In addition, memory errors were considered in the MTS (i.e., when an individual chose the wrong array or failed to make a “no match” response).

2.5 Procedure

All methods and procedures were approved by the Institutional Review Board at the Naval Health Research Center, San Diego, California (protocol no. NHRC.2019.0007). The hypothermia lab is a CWI military training exercise that is part of the Cold Weather Medicine course at MCMWTC (elevation 6889 ft/2100 m). The lab is a course requirement (i.e., all students must complete the CWI), however participating in the data collection portion was completely voluntary. The lab occurs 1 week into the Cold Weather Medicine course and consists of 10 minutes of CWI in an outdoor pond located on base, followed by 60 minutes of field rewarming under direct medical supervision. Data were collected during four sessions spanning two winters. For the exercises reported here, the average air and water temperatures were held at approximately 23 °F/ −5 °C and 34.5 °F/ 1.5 °C, respectively.

Potential participants were briefed on the study details, after which they provided voluntary informed consent. No uniformed research staff members were involved in participant recruitment or the consent process. Participants were asked to avoid alcohol, tobacco, and caffeine 24 hours prior to the exercise. A practice session was held the day before the exercise to familiarize participants with the SRT and MTS tasks. Demographic variables and height, weight, and body fat percentage were collected during this session. Body fat percentage was calculated by obtaining circumference measurements from multiple sites and inputting values into the U.S. Navy’s body fat equation. The order of the SRT and MTS tasks was randomized during each administration period. Data were collected at six time points throughout the exercise. On the morning of the exercise (0500), students who had agreed to participate in the study portion of the exercise completed indoor baseline SRT and MTS testing ("Baseline"), followed by a briefing to all students on the procedures for the hypothermia lab (0600). The whole class then moved to a staging area near the pond ("Pre-Immersion"). Beginning at 0700, groups of students (2–6 at a time) entered the pond for 10 minutes, with groups continuously occupying the pond. After completing 10 minutes of CWI, students remained in their wet clothing for 10 minutes ("Post-Immersion"), followed by reclothing to dry clothing. The final segment consisted of students passively rewarming in their sleeping bags for 60 minutes. During this portion, data were collected at the beginning ("Rewarm: Begin"), 15 minutes ("Rewarm: 15 min"), and at the end of rewarming ("Rewarm: 60 min"). Response times were treated as a continuous time-series across the sessions for all analyses.

2.6 Analyses

All analyses were conducted in the statistical programming language R (R Core Team, 2018). All visuals were created using *ggplot2* (Wickham, 2016) and *factoextra* (Kassambara & Mundt, 2017). For inference, we focus on Bayesian hypothesis testing and estimation. Bayes factors (BFs) were computed using the *BayesFactor* package (Morey & Rouder, 2018. A default prior of r = 2 /2 (i.e., "medium") was utilized for analyzing differences in means, and a default fixed r scale = ½ was used for analyses of variance (ANOVAs). For ANOVA results, we used a model comparison procedure where if any factor(s) demonstrated a BF >3 against the error-only random effect (i.e., value ~ participant), the two were compared against each other (e.g., a main effect model vs. interaction model).One hundred thousand Markov chain Monte Carlo simulations were computed for BF stability. Evidence categories for BFs follow those of Lee and Wagenmakers (2013): 1–3 anecdotal/weak, 3–10 moderate, 10–30 strong, 30–100 very strong, >100 extreme. All tests are non-directional and are presented in terms of the alternative hypothesis (some difference) or null (no difference) when a given BF is >1. For estimation, the *BEST* package (Kruschke, 2013) was used. Ninety-five percent highest density intervals (95% HDIs), the simulated mode effect size (i.e., the maximum a posteriori estimate; MAP), and the percentage of the simulated effect size distribution that falls within the region of practical interest (ROPE) are presented. We defined the ROPE as an interval of small effect sizes ranging from −.2 to .2 (Cohen, 1988).

SampEn (*m, r, N*) was computed using the *TSEntropies* package (Tomcala, 2018). SampEn is the negative natural logarithm of the conditional probability that two sequences that are similar for *m* points remain similar at the next point, where *r* is the tolerance for accepting matches, and *N* is the length of the time series (Richman & Moorman, 2000). Therefore, a lower value of SampEn indicates less complexity in a time series. SampEn offers several methodological benefits, such as being robust to observational noise (Yentes et al., 2013) and remaining stable under relatively small discrete data lengths (~200 trials) under most combinations of the *m* and *r* parameters (Yentes et al., 2018). RT data for both tasks were treated as a continuous time series and error trials were not removed. Given known variability in SampEn values as a function of parameter selection (i.e., *r* and *m* values) and stationarity of the times series, sensitivity analyses were conducted. Put briefly, the default values showed stability across parameters. Furthermore, differencing times series to deal with stationarity potentially removes important cross-trial dependencies (e.g., autocorrelation), which are of specific interest in RT data (e.g., Gratton et al., 1992). Thus, raw time-series data were not differenced and SampEn was computed using the default parameters for each participant and task (r = .2*SD and m = 2).

Data-driven hierarchical agglomerative clustering (HAC) analyses was conducted over SRT and MTS SampEn values to identify groups and potential differences in cognitive flexibility. The *cluster* (Mäechler et al., 2019) and *NbClust* packages (Charrad et al., 2014) were used to identify the optimal number of clusters (*k*) using 30 common indices for determining *k*. Ward's method was chosen as the linkage method for clustering. This method attempts to build clusters with dense centers and maximize differences (Murtagh & Legendre, 2014). Two dimensions of performance were considered to assess flexibility. The first was ACF, which provides a measure of how much future values are dependent on past values (Kang et al., 2017) and was computed over time series using the *tsfeatures* package at $k = 1$ (Hyndman et al., 2019). The second was errors in both tasks. For errors, generalized linear mixed-effect models (GLMMs) were fit using the *lme4* package (Bates et al., 2015). Both linear and quadratic orthogonal polynomials were considered over the Trial variable. Trial was scaled and centered for all models. Models considered the Cluster × Trial interaction, and Trial was crossed in random slopes and intercepts, including the polynomial when applicable. Re-leveling of the Cluster factor was used to interpret simple effects (i.e., the effect of trial within each cluster). Unless explicitly noted, all simple effects reported used Cluster 1 as the reference level. Log-likelihood tests were used for model comparison model fits. Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1978) are presented alongside log-likelihood tests (lower values indicate better model fits).

The final portion of the Results focuses on relations between demographics, body composition, body temperatures, and cluster membership. An exploratory stepwise logistic regression was conducted to predict cluster membership based on AIC using forward and backward selection. Regressors included age, height, weight, body fat percentage, body mass index, and skin and core temperatures. Prior work suggests that differences in body fat percentage alter thermal responses to CWI and thus may influence performance (Stephens et al., 2014).

3.0 RESULTS

3.1 Hierarchical Clustering

The HAC method produced a $k = 2$ optimal solution with 21 individuals being assigned to Cluster 1 and 36 individuals being assigned to Cluster 2. Figure 1 displays that Cluster 1 is associated with lower SampEn values for both the SRT and MTS tasks relative to Cluster 2. From this, we can infer that Cluster 1 RTs became more destabilized relative to Cluster 2. In the following analyses, we turn to assessing differences in flexibility of each task's time series across these groups.

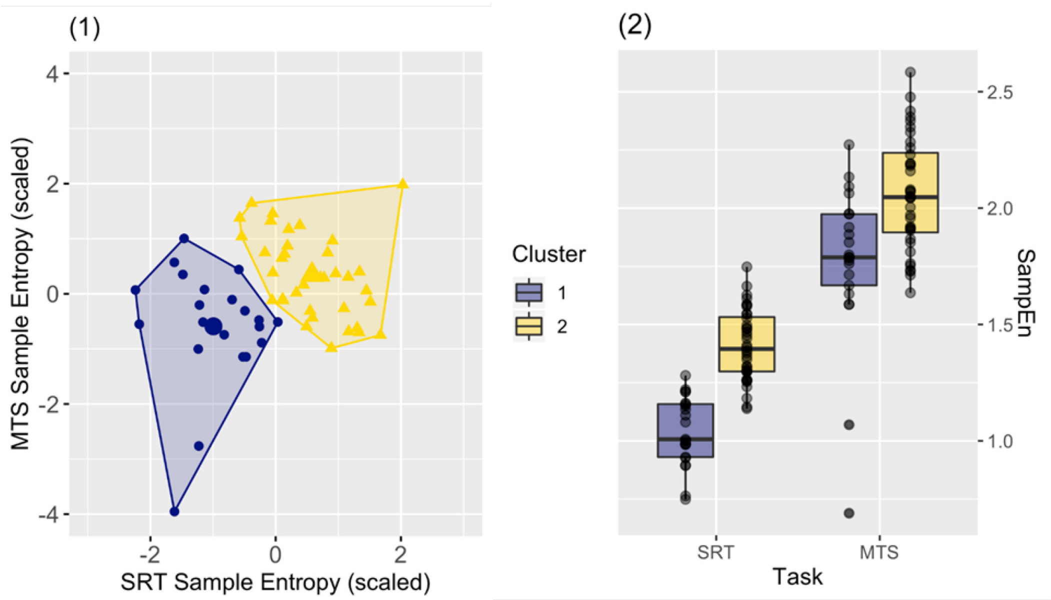

Figure 1: Results (Panel 1) and Box Plots (Panel 2) of $k = 2$ HAC Solution Using Participant Simple Reaction Time (SRT) and Match-to-Sample (MTS) SampEn Values. Points within the plot are individual participants. The larger circle (Cluster 1) and larger triangle (Cluster 2) represent the centroid of each cluster in the left panel.

3.2 ACF

BF ANOVA results favored the joint main effect model (Cluster + Task) over the interaction model, $BF_{ALT} = 3.61$, as well as the individual main effect models, $BF_{ALT} > 1000$ for both comparisons. The joint main effect model best fit the data. ACF values were higher for Cluster 1 for the SRT (Cluster 1, $M = .38$, $SD = .16$; Cluster 2, $M = .24$, $SD = .12$; $BF_{ALT} = 39.48$, MAP = .95, 95% HDI [.31, 1.55], % in $ROPE_{[-.2, .2]} = 1\%$) and for the MTS (Cluster 1, $M = .22$, $SD = .16$; Cluster 2, $M = .08$, $SD = .09$; $BF_{ALT} = 155.64$, MAP = 1.05, 95% HDI [.37, 1.7], % in $ROPE_{[-.2, .2]} = .64\%$). Overall, ACF was higher for the SRT ($M = .29$, $SD = .15$) relative to the MTS ($M = .14$, $SD = .14$), $BF_{ALT} > 100{,}000$, MAP = .86, 95% HDI [.55, 1.19], % in $ROPE_{[-.2, .2]} = .002\%$ (see Figure 2).

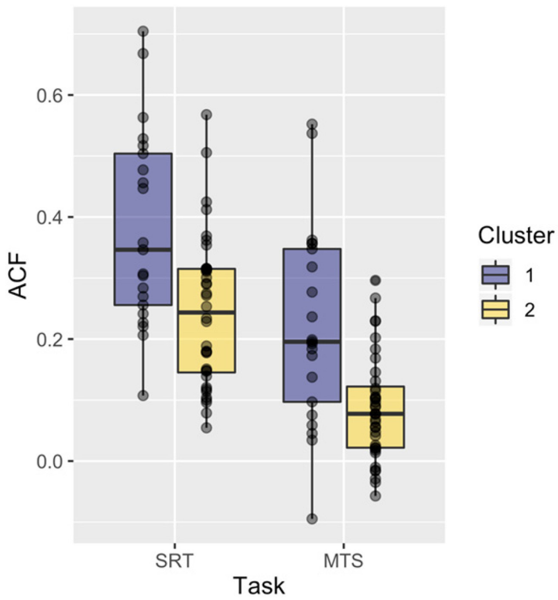

Figure 2: Box Plots for Autocorrelation Function (ACF) as a Function of Cluster and Task. SRT = simple reaction time task; MTS = match-to-sample task.

3.3 Errors

First, we note that there were no differences in fast (i.e., false starts) errors across clusters in the SRT, and no fast or memory errors across clusters in the MTS. Thus, here we focus on slow errors (i.e., response omissions).

3.3.1 Simple Reaction Time Task Errors

The linear Trial × Cluster interaction did not demonstrate a significant interaction effect, $b = .18$, $SE = .15$, $Z = 1.25$, $p > .1$, or trial within either cluster, Zs < 1.81, ps $> .07$ for both. There was a significant simple effect of Cluster, $b = -.96$, $SE = .46$, $Z = -2.12$, $p = .03$, where overall Cluster 2 had a lower probability of making a slow error. The polynomial GLMM demonstrated a significant interaction effect over the quadratic term, $b = 101.77$, $SE = 20.64$, $Z = 4.93$, $p < .001$, revealing the difference in quadratic patterns of errors across the clusters. Cluster 2 demonstrated a significant negative quadratic trend of trial, $b = -43.8$, $SE = 19.47$, $Z = -2.25$, $p = .02$. However, as evidenced by the significant interaction term, Cluster 1 demonstrated a more negative quadratic trend (i.e., more peaked), $b = -145.58$, $SE = 17.78$, $Z = -8.19$, $p < .001$. The polynomial model (AIC = 2062.4, BIC = 2152.7) relative to the linear model (AIC = 2181.4, BIC = 2234.1) was a better fit to the SRT error data, $\chi^2(5) = 128.98$, $p < .001$. Thus, though both clusters demonstrated significant negative quadratic trend over trials (i.e., an inverted-U), Cluster 1's curve was more peaked (see Figure 3, Panel 1).

3.3.2 Match-to-Sample Task Errors

The linear Trial × Cluster interaction did not demonstrate a significant interaction effect, $b = -.03$, $SE = .11$, $Z = -.25$, $p > .1$. There was a significant simple effect of Cluster, $b = -1.02$, $SE = .42$, $Z = -2.46$, $p = .01$, where overall Cluster 2 had a lower probability of making a slow error. Furthermore, the simple effect of trial was signficant in Cluster 1, $b = .17$, $SE = .07$, $Z = 2.38$, $p = .02$, but not in Cluster 2, $b = .14$, $SE = .12$, $Z = 1.25$, $p > .1$. The polynomial GLMM demonstrated a signficant interaction effect over the quadratic term, $b = 32.2$, $SE = 12.04$, $Z = 2.67$, $p = .008$. Cluster 2 demonstrated a signficant negative quadratic trend of trial, $b = -28.7$, $SE = 9.38$, $Z = -3.06$, $p = .002$, where the probability of slow errors increased over time and began flattening and dropping later in the CWI. Cluster 1 demonstrated a more negative quadratic trend, $b = -60.9$, $SE = 8.71$, $Z = -6.99$, $p < .001$, with errors peaking near the middle of the CWI and dropping off sharply. The polynomial model (AIC = 4868.6, BIC = 4957.6) relative to the linear model (AIC = 5055.3, BIC = 5107.2) was a better fit to the data, $\chi^2(5) = 196.65$, $p < .001$. Thus, somewhat similar to slow errors in the SRT, both clusters demonstrated an inverted-U slow error curve, though Cluster 1's curve was more peaked at the center of the CWI (see Figure 3, Panel 2).

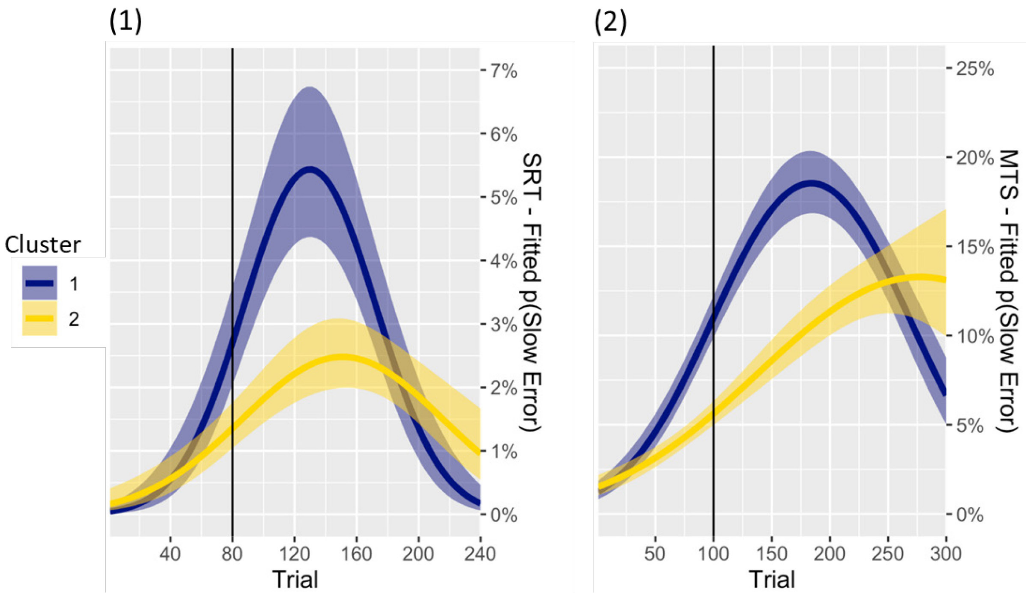

Figure 3: Fitted probability of SRT (panel 1) and MTS (panel 2) errors as a function of trial and cluster. 95% confidence intervals are displayed as the shaded bands. The black vertical line denotes the point within the SRT and MTS trial sequences where participants entered the pond. SRT = simple reaction time task; MTS = match-to-sample task.

3.4 Prediction of Cluster Membership

Results demonstrated Pre-Immersion hand temperature as the best predictor of cluster classification, with individuals with higher Pre-Immersion hand temperature being more likely to belong to Cluster 2, OR = 1.2, 95% confidence interval [1.04, 1.39], $p = .01$, Nagelkerke pseudo-$R^2 = .16$. To reiterate, Cluster 2 was associated with better stability, flexibility, and fewer slow errors in each task. Figure 4 displays the fitted probabilities of Cluster 2 classification for the regression model. The x-intercept indicates an input Pre-Immersion hand temperature value of 14.22 ºC, which corresponds to the 50% probability point in the function. From this model, we can predict that an individual with a Pre-Immersion hand temperature under 14.22 ºC would be more likely to fall into the negatively affected Cluster 1, whereas a temperature above that point would be more likely to fall into Cluster 2.

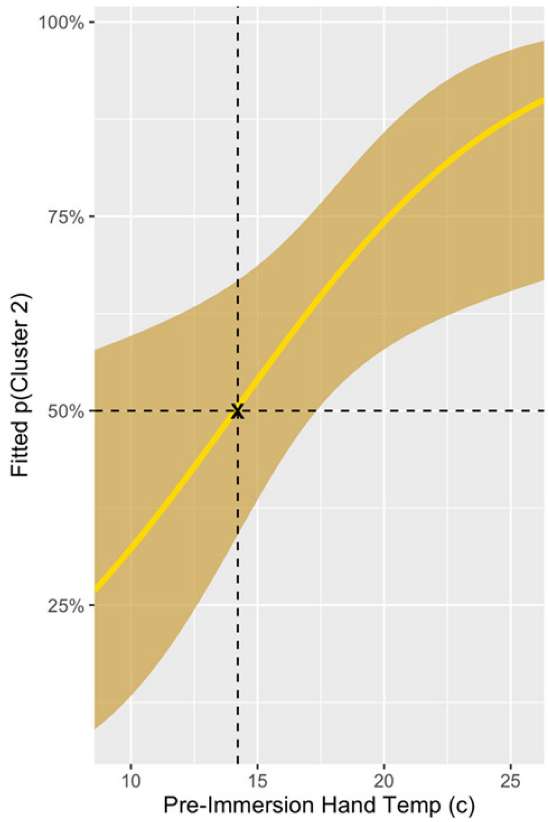

Figure 4: Logistic regression results as a function of pre-immersion hand temperature. The 95% confidence intervals are displayed as the shaded bands. The *x*-intercept indicates an input Pre-Immersion hand temperature value of 14.22 °C, and the *y*-intercept indicates the 50% probability point in the function.

4.0 DISCUSSION

Stability allows for processes such as attention to be focused on goal-oriented information (e.g., overriding distractions). Several physiological responses are initiated upon immersion, involving intense catecholamine release, including NE, which can increase substantially within minutes of CWI (Frank et al., 1997). This neurotransmitter is critical for optimal performance, with increased tonic NE activity being directly implicated in loss of stability, over-arousal, and poor performance due to scanning and labile attention (Aston-Jones & Cohen, 2005). In the current set of results, the putative underlying increase in NE from CWI likely pushed individuals in Cluster 1 to less stability across both tasks relative to Cluster 2. The ability to maintain stability in and of itself comes at a cognitive cost if the value of maintenance wanes (see Westbrook & Braver, 2016). Interestingly, recent work has demonstrated that even the perceived costs of engaging in tasks associated with stability demands (distractor resistance) is larger than those associated with flexibility demands(e.g., updating memory representations; Papadopetraki et al., 2019). Therefore, heightened neurophysiological responses to CWI compounded with the cognitive costs associated with maintenance likely result in an ideal situation for performance decline.

We were additionally able to identify individual differences in ACF and omission errors across the two tasks. Interestingly memory errors did not differ across the clusters in the MTS, suggesting that individuals in both clusters successfully encoded stimuli at a similar rate. We interpret these joint findings as reflecting weaker flexibility, specifically an inability to focus attention to responding within the alloted time. That is, Cluster 1's increased omission errors are likely from uncoupling specific task goals, not from a full disengagement from the tasks, which would be reflected in increased omission errors *and* memory errors. This may be attributed to CWI affecting attentional control mechanisms related to vigilance (i.e., decoupling from maintaining operative and goal-related information in the face of distraction not explicitly introduced through a task; von Bastian et al., 2020) through increased catecholamine (e.g., NE) release triggered by the rapid and intense cooling of the skin (Jacobs et al., 1994). It is important to note that we are not making a directionality conclusion with regard to stability and flexibility (e.g., lower stability led to weak flexibility). Both lower stability and weak flexibility covaried within the parameters of our experiment, leading to overall lower adaptability in Cluster 1 individuals.

Individuals' hand temperature *prior* to the CWI predicted cluster classification. We offer a plausible explanation for this but note that additional investigations are required to determine the underlying mechanisms contributing to this finding. Though we did not query individuals about their cold exposure experiences prior to the study, pre-immersion hand temperature variation can be due to cold habituation (e.g., Enander et al., 1980; Leblanc, 1988). Specifically, localized cold habituation from repeated cold exposure can cause physiological and perceptual adjustments, such as warmer skin temperature, improved thermal comfort, and attenuated shivering (Makinen, 2010). Indeed, these adjustments may be influenced by changes in vasomotor responses resulting from NE attenuation following repeated cold stress (LeBlanc et al., 1975). In the current study, participants were exposed to cold intermittently and organically through participation in other course activities leading up to study participation. It is possible that the degree or frequency of exposure could have differed between participants. For instance, some participants may have removed cold weather gloves more frequently leading to some level of habituation. Although we have no direct evidence to support localized cold habituation for individuals in Cluster 2, it does provide a practical explanation for differences in hand temperature prior to CWI. Often, environments that are too extreme (e.g., CWIs) do not favor opportunities to detect small physiological differences, such as a few degrees of hand temperature between individuals (Young et al., 1986). However, in the presence of mild cold air exposure in the current investigation (23 °F/−5 °C), detection of small differences in hand temperature was possible. Through warmer skin temperature and improved comfort, distraction may be minimized and increased attention given to the immediate task, thereby improving performance (Jones et al., 2017).

5.0 CONCLUSION

A comprehensive view of cognition and behavior requires understanding the human as embedded in complex environments. The current study provided evidence of utility for methods built to take such complexity into account. SampEn computed over RT time series can be used as prediction metric revealing underlying differences in cognitive flexibility. Given the finding that low hand temperatures were associated with lower stability and weak flexibility, future work should aim to disentangle unique contributions from extremity temperatures on attention spanning multiple forms of distraction. By broadening the scope of our methods, more focused predictions and accounts can be created for the extreme contexts that require human adaptability.

6.0 REFERENCES

[1] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723.

[2] Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. Annual Review of Neuroscience, 28, 403–450.

[3] Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. Current Directions in Psychological Science, 20(3), 160–166.

[4] Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). The adapted mind: Evolutionary psychology and the generation of culture. Oxford University Press.

[5] Bates D., Mäechler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48.Belmonte, C., Brock, J. A., & Viana, F. (2009). Converting cold into pain. Experimental Brain Research, 196(1), 13–30.

[6] Brooks, D. R., Collier, J., Maurer, B. A., Smith, J. D., & Wiley, E. O. (1989). Entropy and information in evolving biological systems. Biology and Philosophy, 4(4), 407–432.

[7] Cañas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. Ergonomics, 46(5), 482–501.

[8] Castellani, J. W., O'Brien, C., Sawka, M. N., & Gamble, W. B. (2005). Prevention and management of cold-weather injuries (TB-MED-508). Headquarters, Department of the Army. https://www.usariem.army.mil/assets/docs/publications/articles/2005/tbmed508.pdf

[9] Castellani, J. W., & Young, A. J. (2016). Human physiological responses to cold exposure: Acute responses and acclimatization to prolonged exposure. Autonomic Neuroscience, 196, 63–74.

[10] Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. Journal of Statistical Software, 61(6), 1–36.

[11] Chemero, A. (2009). Radical embodied cognitive science. MIT Press.

[12] Chrysikou, E. G. (2018). The costs and benefits of cognitive control for creativity. In R. E. Jung & O. Vartanian (Eds.), The Cambridge handbook of the neuroscience of creativity (pp. 299–317). Cambridge University Press.

[13] Clark, A. (2008). Supersizing the mind: embodiment, action, and cognitive extension. Oxford University Press.

[14] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed). Erlbaum.

[15] Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1481), 933–942.

[16] Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. Nature Neuroscience, 16(9), 1348–1355.

[17] Cools, R., & D'Esposito, M. (2011). Inverted-U–shaped dopamine actions on human working memory and cognitive control. Biological Psychiatry, 69(12), e113–e125.

[18] Danzl, D. F., & Pozos, R. S. (1994). Accidental hypothermia. New England Journal of Medicine, 331(26), 1756–1760.

[19] Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. Psychological Bulletin, 145(5), 508–535.

[20] Enander, A. (1987). Effects of moderate cold on performance of psychomotor and cognitive tasks. Ergonomics, 30(10), 1431–1445.

[21] Enander, A., Sköldström, B., & Holmér, I. (1980). Reactions to hand cooling in workers occupationally exposed to cold. Scandinavian Journal of Work, Environment & Health, 6(1), 58–65.

[22] Frank, S. M., Higgins, M. S., Fleisher, L. A., Sitzmann, J. V., Raff, H., & Breslow, M. J. (1997). Adrenergic, respiratory, and cardiovascular effects of core cooling in humans. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 272(2), R557–R562.

[23] Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. Cortex, 86, 186–204.

[24] Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. Journal of Physiology-Paris, 100(1-3), 70–87.

[25] Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. Journal of Experimental Psychology: General, 121(4), 480–506.

[26] Harte, J. (2011). Maximum entropy and ecology: A theory of abundance, distribution, and energetics. Oxford.

[27] He, B. J. (2011). Scale-free properties of the functional magnetic resonance imaging signal during rest and task. Journal of Neuroscience, 31(39), 13786–13795.

[28] Henriksson, O., Lundgren, J. P., Kuklane, K., Holmér, I., & Bjornstig, U. (2009). Protection against cold in prehospital care—thermal insulation properties of blankets and rescue bags in different wind conditions. Prehospital and Disaster Medicine, 24(5), 408–415.

[29] Heus, R., Daanen, H. A., & Havenith, G. (1995). Physiological criteria for functioning of hands in the cold: A review. Applied Ergonomics, 26(1), 5–13.

[30] Holling, C. S. (1973). Resilience and stability of ecological systems. Annual Review of Ecology and Systematics, 4(1), 1–23.

[31] Hristovski, R., & Balagué, N. (2020). Theory of cooperative-competitive intelligence: Principles, research directions, and applications. Frontiers in Psychology, 11, Article 2220.

[32] Hutchins, E. (2010). Cognitive ecology. Topics in Cognitive Science, 2(4), 705–715.

[33] Hyndman, R., Kang, Y., Montero-Manso, P., Talagala, T., Wang, E., Yang, Y., & O'Hara-Wild, M. (2019). tsfeatures: Time series feature extraction. R package version 1.0.1. https://CRAN.R-project.org/package=tsfeatures

[34] Jacobs, I., Martineau, L., & Vallerand, A. L. (1994). Thermoregulatory thermogenesis in humans during cold stress. Exercise and Sport Sciences Reviews, 22(1), 221–250.

[35] Jones, D. M., Bailey, S. P., Roelands, B., Buono, M. J., & Meeusen, R. (2017). Cold acclimation and cognitive performance: A review. Autonomic Neuroscience, 208, 36–42.

[36] Kang, Y., Hyndman, R. J., & Smith-Miles, K. (2017). Visualising forecasting algorithm performance using time series instance spaces. International Journal of Forecasting, 33(2), 345–358.

[37] Kassambara, A., & Mundt, F. (2017). factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.5. https://CRAN.R-project.org/package=factoextra

[38] Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. Journal of Experimental Psychology: General, 142(2), 573–603.

[39] Lee, M. D., & Wagenmakers, E. J. (2013). Bayesian cognitive modeling: A practical course. Cambridge University Press.

[40] Leblanc, J. (1988). Factors affecting cold acclimation and thermogenesis in man. Medicine & Science in Sports & Exercise, 20(5 Suppl), S193–S196.

[41] LeBlanc, J., Dulac, S., Côté, J., & Girard, B. (1975). Autonomic nervous system and adaptation to cold in man. Journal of Applied Physiology, 39(2), 181–186.

[42] Lipsitz, L. A., & Goldberger, A. L. (1992). Loss of complexity and aging: Potential applications of fractals and chaos theory to senescence. Journal of the American Medical Association, 267(13), 1806–1809.

[43] Mäechler, M. (2021). Cluster. R package version 2.1.2. https://svn.r-project.org/R-packages/trunk/cluster/

[44] Makinen, T. M. (2010). Different types of cold adaptation in humans. Frontiers in Bioscience-Scholar, 2(3), 1047–1067.

[45] Martin, K., McLeod, E., Périard, J., Rattray, B., Keegan, R., & Pyne, D. B. (2019). The impact of environmental stress on cognitive performance: A systematic review. Human Factors, 61(8), 1205–1246.

[46] Morey, R. D., & Rouder, J. N. (2018). BayesFactor: Computation of Bayes factors for common designs. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor

[47] Muller, M. D., Gunstad, J., Alosco, M. L., Miller, L. A., Updegraff, J., Spitznagel, M. B., & L. Glickman, E. (2012). Acute cold exposure and cognitive function: Evidence for sustained impairment. Ergonomics, 55(7), 792–798.

[48] Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? Journal of Classification, 31(3), 274–295.

[49] Newell, K. M., Vaillancourt, D. E., & Sosnoff, J. J. (2006). Aging, complexity, and motor performance. In J. E. Birren & K. W. Schaire (Eds.), Handbook of the psychology of aging, 163–182. Elsevier.

[50] O'Brien, C., Tharion, W. J., Sils, I. V., & Castellani, J. W. (2007). Cognitive, psychomotor, and physical performance in cold air after cooling by exercise in cold water. Aviation, Space, and Environmental Medicine, 78(6), 568–573.

[51] Oksa, J., Rintamäki, H., & Mäkinen, T. (2006). The effect of training of military skills on performance in cold environment. Military Medicine, 171(8), 757–761.

[52] Papadopetraki, D., Froböse, M. I., Westbrook, A., Zandbelt, B. B., & Cools, R. (2019). Quantifying the cost of cognitive stability and flexibility. bioRxiv. https://doi.org/10.1101/743120

[53] Paulus, M. P., Potterat, E. G., Taylor, M. K., Van Orden, K. F., Bauman, J., Momen, N., Padilla, G. A., & Swain, J. L. (2009). A neuroscience approach to optimizing brain resources for human performance in extreme environments. Neuroscience & Biobehavioral Reviews, 33(7), 1080–1088.

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

[55] R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

[56] Ramanathan, N. L. (1964). A new weighting system for mean surface temperature of the human body. Journal of Applied Physiology, 19(3), 531–533.

[57] Ranganath, C., Cohen, M. X., & Brozinsky, C. J. (2005). Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. Journal of Cognitive Neuroscience, 17(7), 994–1010.

[58] Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology-Heart and Circulatory Physiology, 278(6), H2039–H2049.

[59] Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6(2), 461–464.

[60] Simola, J., Zhigalov, A., Morales-Muñoz, I., Palva, J. M., & Palva, S. (2017). Critical dynamics of endogenous fluctuations predict cognitive flexibility in the Go/NoGo task. Scientific Reports, 7(1), 1–10.

[61] Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics, 28(1), 112–118.

[62] Stephens, J. M., Argus, C., & Driller, M. W. (2014). The relationship between body composition and thermal responses to hot and cold water immersion. Journal of Human Performance in Extreme Environments, 11(2), Article 1.

[63] Stocks, J. M., Taylor, N. A., Tipton, M. J., & Greenleaf, J. E. (2004). Human physiological responses to cold exposure. Aviation, Space, and Environmental Medicine, 75(5), 444–457.

[64] Tardiff, N. (2019). Neurobiological foundations of stability and flexibility (Publication No. 3614) [Doctoral dissertation, University of Pennsylvania]. https://repository.upenn.edu/edissertations/3614

[65] Tipton, M. J. (1989). The initial responses to cold-water immersion in man. Clinical Science, 77(6), 581–588.

[66] Tomcala, J. (2018). TSEntropies: Time series entropies. R package version 9. https://cran.rproject.org/web/packages/TSEntropies/index.html

[67] Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. Journal of Experimental Psychology: General, 132(3), 331–350.

[68] van Schouwenburg, M. R., den Ouden, H. E., & Cools, R. (2010). The human basal ganglia modulate frontal-posterior connectivity during attention shifting. Journal of Neuroscience, 30(29), 9910–9918.

[69] von Bastian, C. C., Blais, C., Brewer, G., Gyurkovics, M., Hedge, C., Kałamała, P., Meier, M., Oberauer, K. Rey-Mermet, A., Rouder, J. N., Souza, A. S., Bartsch, L. M., Conway, A. R. A., Draheim, C., Engle, R. W., Friedman, N. P., Frischkorn, G. T., Gustavson, D. E., et al. Wiemers, E. (2020). Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations. PsyArXiv. https://doi.org/10.31234/osf.io/x3b9k

[70] Westbrook, A., & Braver, T. S. (2016). Dopamine does double duty in motivating cognitive effort. Neuron, 89(4), 695–710.

[71] Wickham., H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag.

[72] Yentes, J. M., Denton, W., McCamley, J., Raffalt, P. C., & Schmid, K. K. (2018). Effect of parameter selection on entropy calculation for long walking trials. Gait & Posture, 60, 128–134.

[73] Yentes, J. M., Hunt, N., Schmid, K. K., Kaipust, J. P., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. Annals of Biomedical Engineering, 41(2), 349–365.

[74] Young, A. J., Muza, S. R., Sawka, M. N., Gonzalez, R. R., & Pandolf, K. B. (1986). Human thermoregulatory responses to cold air are altered by repeated cold water immersion. Journal of Applied Physiology, 60(5), 1542–1548.

[75] Young, L. S. (2003). Entropy in dynamical systems. Entropy, 313, 114–127.