A Meta-Analysis of Performance Response Under Thermal Stressors

P. A. Hancock, Jennifer M. Ross, and **James L. Szalma,** University of Central Florida, Orlando, Florida

Objective: Quantify the effect of thermal stressors on human performance. **Background:** Most reviews of the effect of environmental stressors on human performance are qualitative. A quantitative review provides a stronger aid in advancing theory and practice. **Method:** Meta-analytic methods were applied to the available literature on thermal stressors and performance. A total of 291 references were collected. Forty-nine publications met the selection criteria, providing 528 effect sizes for analysis. **Results:** Analyses confirmed a substantial negative effect on performance associated with thermal stressors. The overall effect size for heat was comparable to that for cold. Cognitive performance was least affected by thermal stressors, whereas both psychomotor and perceptual task performance were degraded to a greater degree. Other variables were identified that moderated thermal effects. **Conclusion:** Results confirmed the importance of task type, exposure duration, and stressor intensity as key variables impacting how thermal conditions affect performance. Results were consistent with the theory that stress forces the individual to allocate attentional resources to appraise and cope with the threat, which reduces the capacity to process task-relevant information. This represents a maladaptive extension of the narrowing strategy, which acts to maintain stable levels of response when stress is first encountered. **Application:** These quantitative estimates can be used to design thermal tolerance limits for different task types. Although results indicate the necessity for further research on a variety of potentially influential factors such as acclimatization, the current summary provides effect size estimates that should be useful in respect to protecting individuals exposed to adverse thermal conditions.

INTRODUCTION

How the thermal environment affects human response capacity has been the subject of both theoretical speculation and experimental evaluation since before psychology or even physiology became recognized sciences. Implicit evaluations of thermal effects are evident in the selection of sites for habitation and are influential in the earliest formal military conflicts (Goldman, 2001). More scientifically stringent observations were first generated as a result of practical problems faced by manufacturers whose processes involved the exposure of workers to temperature conditions in which it was uncertain that they could survive. Originally, it was believed that humans could not tolerate levels of thermal exposure that exceeded the boiling point of water, but an early and adventurous empirical demonstration proved the fallacy of this assumption (Blagden, 1775a, 1775b).

The influence of the thermal environment on behavior also played an evident role in patterns of global colonization. For example, Henry Ellis, the then English Governor of Georgia, in what was to later become the United States, made a number of adverse comments about the summer heat in that locale that served, to a degree and for a time, to discourage further European immigration to the area (H. Ellis, 1758). These and other comparable observations provided the partial foundation of the subsequent theory of environmental determinism. The strict interpretation of this theory had it that environmental conditions, among which temperature was conceived as a crucial controlling factor,

Address correspondence to Peter A. Hancock, 3100 Technology Pkwy., Suite 337, Orlando, FL 32826-0544; phancock@pegasus.cc.ucf.edu. *HUMAN FACTORS,* Vol. 49, No. 5, October 2007, pp. 851–877. DOI: 10.1518/001872007X230226.

provided the key to understanding essentially all of human culture and much of behavior (e.g., Huntington, 1919; Von Humboldt, 1816).

The impact of this growing knowledge of thermal effects on behavior, however, was not solely confined to cultural, colonial, or even industrial issues. At the turn of the 20th century, the study of temperature regulation was central to all of physiology, especially following upon Claude Bernard's crucial conception of the "fixity of the interior environment," made in his classic *Cahier Rouge.* In this work, his observations were primarily concerned with the stability of internal process and especially the regulation of body temperature itself (see Hoff, Guillemin, & Guillemin, 1967).

This notion of a regulated internal environment was elaborated and articulated most clearly in Cannon's (1932) influential exposition on the concept of homeostasis. The codification of the notion of controlling feedback systems that underlie bodily system regulation (Hancock, 1980, 1981b) had a fundamental and continuing effect on many fields of study, exerting its strongest influence on the foundations of physiology, psychology, and the broader areas of cognition and neuroscience (see also Wiener, 1954).

At no time were the ongoing concerns for the practical issue of temperature effects on performance far behind contemporary theoretical developments. For example, in the earliest years of the 20th century, the garment industry in Lancashire in England was crucially dependent upon heat and humidity levels that prevented exposed cotton fibers from drying and breaking during cloth manufacture. How these obligatory heat stressor conditions affected workers was a central issue to the British Industrial Fatigue Research Board (e.g., Vernon & Bedford, 1930), which was itself the precursor to many subsequent world organizations in public health and safety. Practical concerns also drove South African interests in heat and humidity effects, as the efficiency of gold mining was crucially dependent upon understanding how workers could acclimatize to and subsequently perform in the exceptionally high heat and humidity levels of the deep mines (see Wyndham, 1969).

These collective studies showed the advantages of artificial acclimatization procedures on work productivity, including beneficial effects on physical, psychomotor, and cognitive activities (e.g., Goldman, 2001; Patterson, Taylor, & Amos, 1998; Wyndham et al., 1964). Nominally, pragmatism also drove thermal research to immoral depths when investigators under the Nazi regime immersed prisoners in freezing water to determine absolute tolerance times for fatal hypothermia and so, putatively, to aid the survival of pilots shot down over the North Sea (see Burton & Edholm, 1955). For a plethora of reasons such experiments have been repudiated as science and discredited as evaluative procedures (see Hancock, 2003).

More recently, human exploration of outer space triggered further efforts to understand how performance capacities varied under thermal extremes, especially in the circumstances experienced during atmospheric reentry, which, it was thought at one time, might be too extreme for human astronauts to tolerate (Blockley & Lyman, 1950, 1951). In general, then, these various studies fall into one of two divisions, one concerned with normal individuals in tolerable but adverse thermal circumstances and a second group of interest primarily to military and industrial agencies concerned directly with survival in most extreme environments.

The experimental approach with the longest history in understanding thermal effects on performance capability is that which treats stress as a property of the environment itself. Derived originally from engineering approaches to the question of material stress exposure, these experiments proceeded by varying some characteristic of the thermal environment and measuring subsequent effects on reflections of human response, such as comfort (Fanger, 1967), psychomotor and cognitive performance (Poulton, 1970; Poulton & Kerslake, 1965), or simple survival (Taylor, 1948).

In recent decades, these types of evaluation have diminished in frequency as risk-averse human participant review boards have become ever less willing to permit the exposure of individuals to potentially damaging, and extreme, environmental conditions. Despite this trend, this form of research is still practiced, particularly in special circumstances such as the imperative demands for military test and evaluation (see Harris, Hancock, & Harris, 2005; Johnson & Kobrick, 2001).

These evolutions in experimental strategies are evident in our meta-analytic observations, which show the diminution of these environmental exposures that have been reported in more recent decades. Contemporary research has most evidently been influenced by Lazarus's influential notion of stress as a transactional process (e.g., Lazarus & Folkman, 1984). Such studies are much more focused upon the process of coping and appraisal in which the stress is an emergent property of the interaction between the individual and the ambient condition. Indeed, recent research has established that perception of heat-induced pain is reduced by relatively high arousal derived from fear induction (Rhudy & Meagher, 2003). In such contexts, however, thermal stressors remain an important theoretical and practical concern. There are a number of reasons thermal stressors still occupy this crucial role.

It is a supportable proposition that thermal variation is the modal form of stress faced by all living organisms, including human beings. Insights derived from the action of thermal stressors therefore serve to inform studies on all forms of general stress effect. This assertion is supported by a number of facts.

First, temperature is a property of the environment but also a property of the individual; thus with temperature, unlike most other environmental sources of disturbance (e.g., noise and vibration), there is a direct analog of the environmental stressor already present in the organism itself. This association between an exogenously measurable characteristic of the environment and an internal representation of the same characteristic allows one to derive reasonably clear causal linkages, which unfortunately prove to be much more complex in the case of other sources of environmental stress.

Second, it is known that like many other organisms, humans oscillate in response capacity across time of day (Kleitman, 1939/1963). This diurnal variation is tied to intrinsic circadian rhythms, and the prime physiological indicator of circadian phase is core body temperature (Aschoff, 1984). Thus, when researchers seek to understand temperature effects, they already have a strong a priori foundation for expecting a direct relationship between body temperature and level of performance.

Third, thermal conditions are stressful to humans in both their excess and their insufficiency. Therefore, unlike comparable environmental sources of disturbance such as vibration, the low end of the ratio scale of thermal exposure is not a comfortable but rather a fatal condition. Thus, extreme cold eventually proves just as fatal as extreme heat. This, of course, is attributable to the thermal dependence of the biochemical platform upon which life itself is erected (see Prosser & Nelson, 1981).

A final, pragmatic advantage in studying thermal effects lies simply in the number of experimental studies conducted on thermal influences as compared with other sources of environmental stress. As the current meta-analysis shows, there is a substantive existing literature, and this also helps establish the veridical pattern of effects more effectively.

The current review builds on a recent meta-analysis of the effects of temperature on performance published by Pilcher, Nadler, and Busch (2002). Although their review was extremely thorough, we have had the opportunity to evaluate a somewhat larger portion of the extant literature, and there are three particular issues that extend substantively upon the information offered by Pilcher et al. (2002).

First, our extended coverage has permitted us a much more detailed evaluation of the variability of the information that composes the present meta-analysis by use of hierarchical meta-analysis. Our more detailed division by task type, and of performance measure within each task type, which again extends beyond the Pilcher et al. (2002) assessment, has permitted us to evaluate different orders of performance, a division that is essential to a deeper level of understanding, as was first observed by Grether (1973). Similarly, evaluation of the joint effects of exposure intensity and duration permit a fine-grained analysis of "interactions" between variables that can contribute to future theoretical work.

Second, we have presented our meta-analytic results against the background of various descriptive relationships concerning stress and performance and their constituent causal theories. Thus we have assessed the effectiveness of the distilled information in respect to contemporary approaches to thermal stressors in particular and overall stress effects in general. Third, the current work includes formal outlier analysis to explicitly address influential data points. These analyses clarified the interpretation of the sampling error variance and residual variance associated with different levels of the moderator variables tested. Consequently, the current work adds to and extends upon the excellent groundwork of Pilcher et al. (2002).

As a consequence of these and allied reasons, we examined the effects of thermal stressors on performance response capacity. Because our concern is primarily with performance, we have not provided quantitative analysis of the effects of thermal stressors on responses such as the perception of comfort, or on physiological or neurological functioning directly, but have focused directly on the influences that such changes have on psychomotor, perceptual, and cognitive capacities.

Issues such as comfort and the effect of affective reaction on pain perception (e.g., see Rhudy & Meagher, 2003) are, of course, pragmatically important issues for concerns such as indoor occupancy. However, such effects are addressed elsewhere (see Fanger, 1967; Hancock, 2006). Here, we report on thermal influences on task performance that contain a preponderance of demands on the information processing capacities of the exposed individual, although some reports do include at least an element of required physical effort. The purpose of the current meta-analytic review was to understand these performance effects in depth.

ANALYTICAL METHOD

Literature Accumulation

A literature search was conducted using the PsycINFO, MEDLINE, and Dissertation Abstracts International databases, using the following initial key words: *thermal, temperature, hot, cold,* and *heat.* After a preliminary listing of articles was obtained, references from the obtained journals were examined and article citations were also input into the Science Citation Index. In a concurrent process, subject matter experts were consulted for articles that had not been identified by the formal search procedure. We also benefited from the aforementioned effort by Pilcher et al. (2002).

All of the identified articles were then searched for additional reference information. When these processes no longer yielded new citations, we compiled our final listing of articles. This process resulted in the identification of 291 articles, reports, dissertations, and theses published between 1925 and 2004. Of these, 49 papers were identified, containing 57 primary studies, that met our six selection criteria for inclusion. These specific papers are identified in the reference list by an asterisk appearing in front of the first author's name (American Psychological Association, 2001).

Identified Criteria for Study Inclusion

All studies were inspected to ensure that they fulfilled the following six criteria for inclusion in the meta-analysis:

1. Each study reported an empirical examination of thermal stressors in which the experimental manipulation involved the explicit application of either a heat or cold stressor.
2. A room-temperature control group was employed for comparison purposes.
3. Sufficient information was provided regarding environmental temperature conditions in order to be able to generate a wet bulb globe temperature (WBGT) index value.
4. Temperature exposures had to be whole body air exposures (e.g., not partial body exposures or water immersion; see, e.g., Goodman, Hancock, Runnings, & Brown, 1984).
5. Each study had to report on at least one type of performance measure (e.g., accuracy or response time in marksmanship, memory, tracking, detection tasks). Studies using comfort votes, physiological measures, or subjective measures alone were thus excluded from the present analysis.
6. Sufficient information regarding performance measures had to be provided to determine effect size estimates.

It is important to note that rejecting many primary studies in a meta-analysis is a common occurrence and is necessary to ensure meaningful results when combining effect sizes across studies. In addition, and as is often the case (e.g., see Pilcher et al., 2002), the modal reason for exclusion of a study was the failure to provide sufficient data for the calculation of such effect sizes.

Identification of the Thermal Conditions in Each Study

Studies included in the meta-analysis had to report the respective aspects of the environmental conditions sufficient to directly represent the environment in the WBGT index values or to provide enough information to subsequently derive such WBGT values. WBGT was used because it provides a composite measure of physical values of air temperature and humidity and has become the accepted international index of thermal conditions (Parsons, 1993, 1995; Yaglou & Minard, 1957).

Studies reporting the environmental temperature in the form of the effective temperature (ET) index (Equation 1; Brief & Confer, 1971), dry bulb (DB) and relative humidity (RH; Equation 2), or DB alone (computed as a range value using Equation 2 with minimum and maximum RH levels) enabled acceptable conversion to WBGT values. In cases where DB and wet bulb temperatures alone were provided, these were used to estimate RH using equations provided by the Southern Region Headquarters National Weather Service (n.d.).

$$\text{WBGT} = (\text{ET} - 13.1)/0.823 \quad (1)$$

$$\text{WBGT} = 0.567(\text{DB}) + 0.393(\text{RH}) + 3.94 \quad (2)$$

The Calculation of Effect Size

Effect sizes used for the current study were the standardized mean differences between the experimental and the control conditions, often referred to as Hedges's *g* (Hedges & Olkin, 1985, p. 78; see also Hedges, Shymansky, & Woodworth, 1989). Many researchers may be more familiar with Cohen's *d* (Cohen, 1988), which is conceptually identical to Hedges's *g* but statistically different (see Hunter & Schmidt, 2004). When means and standard deviations were available the effect size was calculated using Equation 3,

$$g = \frac{(\overline{X}_\text{E} - \overline{X}_\text{C})}{s}, \quad (3)$$

in which $\overline{X}_\text{E}$ = mean of the experimental condition, $\overline{X}_\text{C}$ = mean of the control condition, and s = standard deviation (for the control condition in within-participants designs; pooled standard deviation for between-participants designs). When the means and/or standard deviations were not available, effect sizes were obtained from inferential statistics (e.g., *t* tests) or sums of squares/mean squares (e.g., from ANOVA tables) using equations found in Lipsey and Wilson (2001) and Hedges et al. (1989).

In calculating each effect size, the sign was controlled to ensure that a positive score represented improvement in performance in the experimental group relative to the control group, whereas a negative score indicated performance impairment. The mean weighted effect sizes were computed by weighting each effect size by the reciprocal of its variance, using procedures described in Hedges and Olkin (1985) and with variance formulae in Morris and DeShon (2002). Prior to the weighting procedure, each effect size was adjusted for statistical bias using established procedures (Hedges & Olkin, 1985) and was also adjusted to a common standard deviation (i.e., standard deviations of the experimental and control groups rather than the standard deviation of differences) using procedures described by Morris and DeShon (2002).

To correct the *g* scores for statistical bias, which decreases the accuracy of the estimates, particularly in cases where sample sizes are small, an adjustment was performed. This adjustment (see Hedges & Olkin, 1985, pp. 78–81) provides an unbiased *d*.

$$d = g * \left(1 - \frac{3}{4 * N - 9}\right) \quad (4)$$

The need for this correction comes from the statistical bias associated with *d* as an estimate of the population parameter (δ). The expectation for *d* is given by $E(d) = \delta/J(N-2)$, where $J(m) = 1 - [3/(4m-1)]$. Thus, the expected value of *d* is not the population parameter δ but is proportional to it (see Hedges & Olkin, 1985, p. 79).

It is important to note that most studies used in the current analysis reported multiple *g* scores for multiple temperatures, tasks, or durations, resulting in a total of 528 effect sizes (181 for cold and 347 for heat). Many of the studies included multiple effect sizes derived from common participant samples, so these estimates are not independent of one another. Violations of the independence assumption can lead to underestimation of the variance because of sampling error (Cheung & Chan, 2004; Martinussen & Bjornstad, 1999). To avoid such violations we averaged the effect sizes within studies prior to estimating means and variances (see Lipsey & Wilson, 2001).

Note, however, that in the moderator analyses this averaging procedure was conducted within each level of the respective moderator variable. For instance, if a particular study contributed an effect size for each of two task categories, these effect sizes were not averaged for that moderator analysis but were included in their respective groups. As the levels of the moderator variables were never formally (i.e., statistically) compared, this does not represent a violation of the independence assumption. However, for the global analysis and other moderator analyses not including task, these effect sizes were averaged. Thus, the number of studies within each level of a moderator variable is not necessarily the sum of the total number of studies used in the global analysis. Such problems associated with the independence assumption are intrinsic to meta-analyses generally and not simply a characteristic of the current work.

In addition to the weighted mean effect size, two variance estimates were computed: variability attributable to sampling error (s_e^2) and variability of the effect sizes (s_g^2). These values were used to estimate the variability attributable to differences in the population effect sizes, δ (s_δ^2). Thus,

$$s_\delta^2 = s_g^2 - s_e^2. \quad (5)$$

A large s_δ^2 indicates that there is variability among the observed effect sizes that cannot be accounted for by sampling error and that there are likely to be one or more variables that additionally moderate the magnitude of the effect in question (see Hunter & Schmidt, 2004, p. 288). If all of the variance in the effect sizes were accounted for by sampling error, then $s_\delta^2 = 0$. The 95% confidence intervals reported were computed using the standard deviation corresponding to s_e^2.

In some cases the estimate of the variance of the effect sizes can be less than the estimate of sampling error variance. This is conceptually similar to a treatment mean square in an *F* ratio being less than the mean square error. Indeed, Hedges and Olkin (1985) recommended a statistical test using these variances to test hypotheses regarding any moderator variables. We have not adopted this approach as this test can be biased, particularly when the sample size is small (Hunter & Schmidt, 2004), and our purpose is to transcend the limitations of significance tests by application of meta-analytic techniques.

Instead we use the "75% rule" guideline recommended by Hunter and Schmidt (2004). According to this guideline, if 75% of the observed variance (s_g^2) is attributable to sampling error (s_e^2), it is likely that most of the other 25% of the variance is also artifactual. Cases in which this condition is not met can be interpreted as instances in which the residual variance (i.e., s_δ^2, the variance not accounted for by sampling error) is likely to be "real" variance attributable to differences in conditions across studies.

Hunter and Schmidt (2004) noted that the 75% rule is as at least as powerful as more formal significance tests for homogeneity, and with small numbers of effect sizes it is actually more powerful. Perusal of the tables reported here indicates that in general, application of the 75% rule results in the conclusion that there is substantial residual variance (i.e., $s_e^2/s_g^2 < 0.75$). Thus, in general, the results reported in the present tables indicate that there likely are moderator variables unaccounted for in the present analyses. This underscores the heterogeneity across studies of the effect of thermal stressors on human performance.

RESULTS

The 181 effect sizes for cold and 347 effect sizes for heat were obtained from a total of 57 primary studies (derived from the 49 asterisked sources in the reference list). Several moderator analyses were computed based on the temperature range, the tasks that were employed, the type of dependent measure or measures used, and the duration of the exposure to the thermal stressors. These moderator analyses were accomplished separately and, when possible, hierarchically (e.g., by task within each temperature range).

The latter strategy is analogous to an analysis of interactions among variables. A limitation of such analyses is that division of effect sizes into separate categories quickly reduces the absolute number of these effect sizes upon which the summary statistics can be estimated. Thus, there are instances in which some levels of a moderator variable contain relatively few studies. Such estimates, as we note in each specific instance, should be interpreted with the appropriate caution. However, these cases reveal the paucity of studies examining particular combinations of relevant variables and therefore point toward the need for further experimentation to reliably identify these specific combinatorial effects.

Outlier Analysis

The residual variances observed in these respective analyses tend to be quite large relative to the error variances. One possible reason for this is the presence of outliers or influential data points. Prior to the present overall analysis, one study (Beshir, El-Sabagh, & El-Nawawi, 1981) was removed because the effect sizes were extreme outliers. (The effect sizes for their heat stressor conditions were $g = -12.76$ and $g = -15.62$ for the 26°C and 30°C WBGT conditions, respectively; see also Pilcher et al., 2002.)

The potential presence of other, more subtly influential data points was tested formally using techniques developed by Huffcutt and Arthur (1995). They introduced a measure conceptually similar to the difference in fit value (DFFITS; the change in the predicted value from the exclusion of a particular case) that accounted for the influence of sample size associated with each effect size. This measure, the sample-adjusted meta-analytic deviancy statistic (SAMD), is therefore an index of the degree to which an individual effect size influences the overall mean effect size (see Huffcutt & Arthur, 1995, pp. 328–329 for pertinent details). This procedure was employed here.

Global Effects of Heat and Cold

As Table 1 shows, the overall effect size of temperature on performance was $g = -0.34$. Consequently, performance under thermal stressors proved on average to be approximately one third of a standard deviation or about 11% worse than performance at a comparative thermoneutral temperature. This unsurprising finding confirms the expectation that thermal stressors adversely affect human information processing and psychomotor capacities. As the confidence interval of this value excludes zero, we are confident as to the reality of this effect.

Table 1 also divides the present results into the different respective effect sizes for heat and cold. The effect sizes for each form of thermal stressor were comparable, with $g = -0.29$ for heat and $g = -0.26$ for cold. Thus heat and cold stressors each exert similar deleterious influences, and neither condition contains zero within its confidence interval. The results for cold effects are very similar in extent to those reported by Pilcher et al. (2002; see Table 1): –0.26 versus –0.26. However, the effect for heat (–0.15 vs. –0.29) and the global overall effect (–0.19 vs. –0.34) are substantively different from those reported by Pilcher et al. (2002; see Table 1).

These differences most probably accrue from the different extent of the literature surveyed and the manner in which specific analyses were conducted, particularly with reference to the assumption of independence of effects. Further, recent developments in meta-analytic techniques have permitted us to estimate effect sizes and their associated sampling errors, an opportunity not available to Pilcher et al. (2002) working at an earlier date.

Outlier analysis. Analysis of the entire data set revealed eight outliers that represented influential data points among the 57 total studies. These are shown in Table 2.

Intensity Effects

In a subsequent hierarchical analysis, a more detailed examination was conducted that derived a finer discrimination with respect to the range of temperatures. The first division was performed on

Continued on page 860

TABLE 1: Effects of Thermal Stressors on Performance Divided Into Respective Categories of Analysis

Category	k	$\bar{g}$	s_e^2	s_g^2	s_δ^2	s_e^2/s_g^2	95% CI (s_e^2)	n
Global	56	–0.34	0.07	1.01	0.94	0.07	$-0.41 < \delta < -0.27$	2037
Heat	40	–0.29	0.06	0.47	0.41	0.13	$-0.36 < \delta < -0.21$	1810
Cold	21	–0.26	0.09	2.57	2.48	0.04	$-0.39 < \delta < -0.14$	363
Heat ET > 85°F (29.4°C)	31	–0.25	0.08	0.67	0.59	0.12	$-0.35 < \delta < -0.16$	649
Heat ET < 85°F (29.4°C)	14	–0.27	0.04	0.24	0.20	0.17	$-0.37 < \delta < -0.17$	1420
Below 78.3°F (25.7°C) ET	5	–0.97	0.13	0.65	0.52	0.20	$-1.28 < \delta < -0.65$	279
78.3°F < ET < 85°F (25.7°C < ET < 29.4°C)	12	–0.18	0.02	0.11	0.09	0.18	$-0.26 < \delta < -0.09$	1311
85°F < ET < 95.4°F (29.4°C < ET < 35.2°C)	14	–0.40	0.08	0.21	0.13	0.38	$-0.55 < \delta < -0.26$	468
Above 95.4°F (35.2°C) ET	21	–0.39	0.10	1.82	1.72	0.05	$-0.52 < \delta < -0.25$	223
Below 52°F (11.1°C) ET	12	–0.78	0.15	2.30	2.15	0.07	$-0.99 < \delta < -0.35$	116
Above 52°F (11.1°C) ET	9	0.35	0.05	1.27	1.22	0.04	$0.20 < \delta < 0.49$	247
Perception	24	–0.92	0.14	1.82	1.68	0.08	$-1.07 < \delta < -0.77$	389
Cognitive	30	–0.18	0.04	0.51	0.47	0.08	$-0.25 < \delta < -0.11$	1638
Psychomotor	19	–0.46	0.10	2.01	1.91	0.05	$-0.60 < \delta < -0.31$	388
Accuracy	50	–0.28	0.07	0.80	0.73	0.09	$-0.35 < \delta < -0.20$	1824
Speed	28	–0.45	0.09	2.57	2.48	0.04	$-0.56 < \delta < -0.34$	652
Less than 1 hr	13	–0.13	0.04	1.01	0.97	0.04	$-0.24 < \delta < -0.02$	250
1 to 2 hr	19	–0.60	0.14	2.34	2.20	0.06	$-0.77 < \delta < -0.44$	176
2 to 3 hr	17	–0.76	0.11	3.02	2.91	0.04	$-0.91 < \delta < -0.60$	381
More than 3 hr	8	–0.15	0.04	1.28	1.24	0.03	$-0.30 < \delta < -0.01$	900

Note. Here, k represents the number of studies that fall into each respective category of analysis; $\bar{g}$ represents the mean effect size of that category; s_e^2 represents the sampling error variance; s_g^2 represents the variance of the effect sizes; s_δ^2 represents the difference between s_e^2 and s_g^2 and represents the variance in the effect sizes not attributable to sampling error (see Hunter & Schmidt, 2004); n represents the number of participants analyzed at each level; CI = confidence interval; ET = effective temperature.

TABLE 2: Outlier Analyses for Hierarchical Meta-Analysis

Category	Outlier	k	$\bar{g}$	s_e^2	s_g^2	s_δ^2	s_e^2/s_g^2	N	95% CI (s_e^2)
Global									
	Beshir et al. (1981); Blockley & Lyman (1950); Lockhart (1968); Lockhart & Kiess (1971); Mayo (1955); Pepler & Warner (1968); van Orden et al. (1996)	49	–0.47	0.10	0.39	0.29	0.25	1087	$-0.56 < \delta < -0.38$
ET									
>85°F	Blockley & Lyman (1950)	27	–0.36	0.09	0.49	0.40	—	603	$-0.46 < \delta < -0.25$
<85°F	Allen & Fischer (1978); Hygge & Knez (2001); Mayo (1955); Peccolo (1962)	10	–0.17	0.03	0.09	0.06	0.33	457	$-0.27 < \delta < -0.07$
<52°F	Lockhart & Kiess (1971); van Orden et al. (1996)	10	–0.56	0.14	1.14	1.00	0.14	76	$-0.79 < \delta < -0.33$
>52°F	Lockhart (1968); Pepler & Warner (1968); Sharma & Panwar (1987)	6	–0.10	0.06	0.14	0.08	0.43	126	$-0.29 < \delta < 0.10$
Duration									
<1 hr	Lockhart (1968); van Orden et al. (1996)	11	–0.28	0.09	0.83	0.74	0.11	214	$-0.45 < \delta < -0.10$
1–2 hr	Blockley & Lyman (1950); Lockhart & Kiess (1971)	17	–0.71	0.14	0.81	0.67	0.17	148	$-0.89 < \delta < -0.53$
2–3 hr	C. R. Bell & Provins (1963); Beshir et al. (1981); Griffiths & Boyce (1971); Lockhart & Kiess (1971); Wyon (1969)	13	–0.41	0.09	0.45	0.36	0.20	238	$-0.57 < \delta < -0.25$
>3 hr	Lockhart & Kiess (1971); Mayo (1955); Pepler & Warner (1968)	5	–0.57	0.10	0.30	0.20	0.33	82	$-0.85 < \delta < -0.29$
	Task: ET > 85°F								
Perceptual	Courtright (1976); Hocking, Silberstein, Lau, Stough, & Roberts (2001)	22	–0.29	0.08	0.24	0.16	0.33	372	$-0.43 < \delta < -0.15$
Cognitive	P. A. Bell (1978); Blockley & Lyman (1950)	10	0.03	0.05	0.28	0.23	0.18	168	$-0.10 < \delta < 0.16$
Motor	Epstein, Keren, Moisseiev, Gasko, & Yachin (1980); Teichner & Wehrkamp (1954)	10	–0.33	0.09	0.29	0.20	0.31	227	$-0.52 < \delta < -0.14$
	Task: ET < 85°F								
Cognitive	Hygge & Knez (2001); Mayo (1955); Peccolo (1962)	10	–0.60	0.08	0.40	0.32	0.20	534	$-0.79 < \delta < -0.42$
	ET > 85°F								
Accuracy	Blockley & Lyman (1950); Epstein et al. (1980)	25	–0.38	0.09	0.43	0.34	0.21	447	$-0.50 < \delta < -0.26$
Speed	Courtright (1976); Epstein et al. (1980); Hocking et al. (2001); Pepler & Warner (1968)	12	–0.08	0.06	0.33	0.27	0.18	269	$-0.21 < \delta < .06$
	ET < 85°F								
Accuracy	Hygge & Knez (2001); Peccolo (1962); Pepler & Warner (1968)	8	–0.18	0.03	0.11	0.08	0.27	357	$-0.29 < \delta < -0.06$
Speed	Holmberg & Wyon (1969)	6	–0.19	0.02	0.05	0.03	0.40	218	$-0.31 < \delta < -0.06$

		Duration < 2 hr						
ET > 85°F	Blockley & Lyman (1950); Courtright (1976); Pepler (1960); Razmjou & Kjellberg (1992)	15	–0.20	0.10	0.19	0.09	0.53	141	$-0.36 < \delta < -0.04$
ET < 85°F	Hygge & Knez (2001)	6	–0.33	0.03	0.11	0.08	0.27	202	$-0.47 < \delta < -0.19$
ET < 52°F	Lockhart & Kiess (1971); van Orden et al. (1996)	10	–0.54	0.14	1.14	1.00	0.12	76	$-0.77 < \delta < -0.31$
ET > 52°F	Lockhart (1968)	3	–0.13	0.08	0.03	0	1	78	$-0.44 < \delta < 0.18$
		Duration > 2 hr						
ET > 52°F	Pepler & Warner (1968)	3	–0.67	0.11	0.22	0.11	0.50	53	$-1.06 < \delta < -0.29$
		Heat						
Perceptual								
Accuracy	Bursill (1958); Chiles (1958); Razmjou & Kjellberg (1992)	9	–0.30	0.10	0.11	0.01	0.91	87	$-0.50 < \delta < -0.09$
Speed	Courtright (1976); Hocking et al. (2001)	9	–0.20	0.06	0.32	0.26	0.19	231	$-0.36 < \delta < -0.05$
Cognitive								
Accuracy	Allen & Fischer (1978); P. A. Bell (1978); Blockley & Lyman (1950); Hygge & Knez (2001); Peccolo (1962)	14	–0.15	0.03	0.27	0.24	0.11	1072	$-0.24 < \delta < -0.06$
Motor								
Accuracy	P. A. Bell (1978); Beshir et al. (1981)	12	–0.66	0.13	0.77	0.64	0.17	112	$-0.87 < \delta < -0.45$
Speed	Razmjou (1996)	4	0.25	0.04	0.06	0.02	0.67	68	$0.04 < \delta < 0.45$
		Cold						
Cognitive								
Accuracy	Pepler & Warner (1968)	9	–0.15	0.09	0.56	0.47	0.16	205	$-0.34 < \delta < 0.04$
Speed	van Orden et al. (1996)	5	0.55	0.15	0.36	0.21	0.42	61	$0.21 < \delta < 0.89$
Motor								
Accuracy	Lockhart (1968)	4	–0.24	0.09	0.06	0	1	84	$-0.54 < \delta < 0.06$
Speed	Lockhart & Kiess (1971)	3	0.06	0.04	0.24	0.20	0.17	56	$-0.17 < \delta < 0.28$

Note. Each statistic is estimated with its respective outliers removed; 85°F = 29.4°C, 52°F = 11.1°C.

the heat stressor effects and was a straight division of effects above or below the threshold of 85°F (29.4°C) ET. This boundary forms the threshold of Lind's (1963) "prescriptive zone," which was integrated into a more comprehensive description of heat effects by Hancock and Vercruyssen (1988) as the threshold of the "zone of thermal tolerance."

The value of 85°F (29.4°C) ET represents the thermal condition in which the body begins the process of obligatory heat storage. In this circumstance, although the individual is dissipating heat at the maximal rate, he or she experiences a dynamic increase in core body temperature. Analogous to the "death zone" in high-altitude climbing, such conditions cannot be permanently tolerated and, unless alleviated in some fashion, will eventually result in permanent harm. Consequently, this value has been taken as a crucial thermal threshold (Hancock, 1984, 1986b).

As is evident from Table 1, the average effect size of these two divisions is very similar (i.e., –0.25 vs. –0.27). However, what is evident is that there is almost threefold greater variability in those studies that make up the group above 85°F (29.4°C) ET (i.e., $s_g^2 = 0.67$) versus those below (i.e., $s_g^2 = 0.24$). This suggests that there are additional influences involved in the performance outcome that composes the former group.

We believe that exposure time and task type play very important roles in this outcome. However, it may be possible that other environmental factors are also influential here. Although it has yet to be completely established whether 85°F (29.4°C) ET represents the equivalent watershed in performance variation as it does in changing the thermophysiological status of the exposed individual, the threefold difference in effect size variability is suggestive. What has become progressively more evident in research on stress effects in general is that mean performance change alone is insufficient to represent a full picture of what is going on as stress increases and, instead, that increasing variability in performance scores represents a crucial indicator.

The present variation is one between effect sizes and does not strictly characterize individual performer variation per se. However, given the degree to which variation across studies can be taken as an important indicator, we can continue to assert that 85°F (29.4°C) ET, and its WBGT equivalent (87.4°F, 30.8°C), remains a useful theoretical and practical threshold for the parsing of heat stressor effects. In future work it may be able to be more explicit in terms of the actual value at which various forms of performance become unstable.

Outlier analysis. An outlier analysis of this category identified one outlier in the ET > 85°F (29.4°C) category and four outliers in the ET < 85°F (29.4°C) category. When these outliers were removed and the meta-analysis recalculated, a somewhat different pattern of results was observed. Higher temperatures (ET > 85°F or 29.4°C) now resulted in larger effects than temperatures below this threshold. The confidence intervals associated with each temperature category in the reanalysis now overlap only slightly. In contrast, with the outliers still included, the effect sizes for the two categories were almost identical (–0.25 and –0.27; see Table 1). Consequently, conclusions that are drawn depend upon whether these five outlier studies are included in the moderator analysis.

However, one possible reason there were outliers in the ET < 85°F (29.4°C) might be sample size. Three of the four studies in that category (the exception was Allen & Fischer, 1978) had sample sizes much larger than those in used in the other studies in that category. It has been argued that removing outliers prior to analysis renders the mean effect size more "representative" of the population of studies. Taken together, the results indicate that Lind's (1963) prescriptive zone differentiates the effectiveness of thermal stressors in impairing performance but that other variables exist that moderate this division. To confirm, this means that whether the 85°F (29.4°C) ET threshold is a useful threshold of performance differentiation depends on the presence of certain other conditions.

Empirical Division of Temperature Range

Table 1 further parses the impact of heat. However, this latter differentiation is made on the basis of an empirical rather than a theoretical division. This new grouping gives four naturally distinct ranges: below 78.3°F (25.7°C) ET; from 78.3°F to 85.0°F (25.7° to 29.4°C) ET; from 85.0°F to 95.4°F (29.4° to 35.2°C) ET; and 95.4°F (35.2°C) ET and above (see Table 3). These values were derived by taking the median division of the studies below 85°F (29.4°C) ET and the median temperature used in studies above 85°F (29.4°C) ET.

For the lowest intensity the effect size is large,

TABLE 3: Studies Reflecting the Four Heat Stressor Temperature Ranges

Report	Temperature	Duration (in Hours)	Adjusted Effect Size
Temperature Range: Below 78°F (25.6°C)			
Allen & Fischer, 1978, Experiment 1	74°F (23.3°C)	0.67	–1.71
Bateman, 1980	73.7°F (23.2°C)	missing	–1.45
Holmberg & Wyon, 1969, Experiment 1	75.5°F (24.1°C)	0.67	–0.16
Wyon, 1969	73.3°F (22.9°C)	2	0.06
Pepler & Warner, 1968	77.4°F (25.2°C)	3	0.16
Temperature Range: 78°–85°F (25.6°–29.4°C)			
Hygge & Knez, 2001	79.2°F (26.2°C)	2	–1.16
Holmberg & Wyon, 1969, Experiment 2	80.2°F (26.8°C)	0.67	–0.55
P. A. Bell, 1978	81.4°F (27.4°C)	missing	–0.48
Moreland & Barnes, 1970	85°F (29.4°C)	2	–0.16
Holmberg & Wyon, 1969, Experiment 1	80.2°F (26.8°C)	0.67	–0.12
Reddy, 1974	81.2°F (27.4°C)	2	–0.09
Mayo, 1955	83.7°F (28.7°C)	80	–0.04
Mackworth, 1946	84.9°F (29.4°C)	3	0.03
Teichner & Wehrkamp, 1954	83.6°F (28.7°C)	missing	0.08
Pepler & Warner, 1968	83.2°F (28.4°C)	3	0.12
Givoni & Rim, 1962	82.7°F (28.2°C)	2	0.14
Wyon, 1969	79.2°F (26.2°C)	2	0.46
Temperature Range: 85°–95°F (29.4°–35°C)			
P. A. Bell, 1978	92.6°F (33.7°C)	missing	–0.72
Weiner & Hutchinson, 1945	95.1°F (35.1°C)	missing	–0.60
Poulton & Kerslake, 1965	88.6°F (31.4°C)	0.33	–0.27
Bateman, 1980	92.6°F (33.7°C)	2	–0.20
Mackworth, 1946	93.1°F (34°C)	3	–0.20
Faerevik & Reinertsen, 2003	89.8°F (32.1°C)	3	–0.15
Givoni & Rim, 1962	92.9°F (33.9°C)	2	–0.14
Pepler & Warner, 1968	89.3°F (31.9°C)	3	0.11
Reilly & Parker, 1988	88.6°F (31.4°C)	6	0.16
Tikuisis, Keefe, Keillor, Grant, & Johnson, 2002	94.6°F (34.8°C)	2	0.17
Razmjou & Kjellberg 1992	94.9°F (35°C)	1.33	1.17
Temperature Range: Above 95°F (35°C)			
Mackworth, 1946	102°F (38.9°C)	3	–3.94
Epstein et al., 1980	124.9°F (51.6°C)	2.25	–2.83
Hocking et al., 2001	100.6°F (38.1°C)	missing	–1.94
Courtright, 1976	99.5°F (37.5°C)	1.1	–1.77
Pepler, 1960, Experiment 2	108.4°F (42.4°C)	0.67	–1.18
Bursill, 1958, Experiment 2	120.1°F (48.9°C)	1.73	–0.88
Epstein et al., 1980	99°F (37.2°C)	2.25	–0.70
Reardon, Fraser, & Omer, 1998	100.8°F (38.2°C)	1.78	–0.69
Cian, Barraud, Melin, & Raphel, 2001	124.7°F (51.5°C)	2	–0.59
Bateman, 1980	103°F (39.5°C)	2	–0.57
Bursill, 1958, Experiment 1	120.1°F (48.9°C)	1.73	–0.47
P. A. Bell, Loomis, & Cervone, 1982	95.9°F (35.5°C)	2.58	–0.41
Razmjou, 1996	104.9°F (40.5°C)	1.3	–0.05
Givoni & Rim, 1962	104°F (40°C)	2	–0.04
Chiles, 1958, Experiment 1	101°F (38.4°C)	1	–0.01
Colquhoun & Goldman, 1972	113.9°F (45.5°C)	2	0.00
C. R. Bell, 1964, Experiment 1	140°F (60°C)	missing	0.04
C. R. Bell, 1964, Experiment 2	132.8°F (56°C)	missing	0.04
Cian et al., 2000	124.7°F (51.5°C)	2	0.04
Chiles, 1958, Experiment 2	109.1°F (42.8°C)	1	0.16
Teichner & Wehrkamp, 1954	100.7°F (38.2°C)	missing	0.61
C. R. Bell & Provins, 1963, Experiment 2	117°F (47.2°C)	2.75	0.72
Curley & Hawkins, 1983	100°F (37.8°C)	2.58	1.10
Blockley & Lyman, 1950	185.6°F (85.4°C)	0.6, 0.57, 1.23	2.73

Note. Some papers contain multiple studies that were analyzed in this meta-analysis.

$g = -0.97$, whereas that of the most intense temperature exposures is moderate, $g = -0.39$. However, as observed in Table 1, the variability of the effect size in the highest heat stressor group is now nearly 17 times higher than that of the lowest group. The error variance for the highest heat stressor group is comparable to that of two of the other three temperature ranges, including the lowest. Thus, factors other than sampling error drive the differences in variability.

With the exception of the lowest temperature range, it is clear that the effect size variation sequentially increases across the three remaining categories. This gives rise to the proposition that performance is relatively stable over much of the temperature range but exhibits radical variation at the highest extreme. This observation, that performance variation is a crucial indicator of incipient failure, is a central characteristic of the extended-U theory of stress and performance (Hancock & Warm, 1989). Table 1 also illustrates similar trends for the influence of cold. Above approximately 52°F (11.1°C), there is a small to moderate increment in performance, which, in respect of observations on transient effects, is not surprising (Poulton, 1976). However, below this temperature there is a large deleterious effect. Thus, as with heat, there is a clear stress intensity effect, which, although not unexpected, is important to confirm.

Outlier analysis: Cold. At temperatures below 52°F (11.1°C), two outliers were identified (Table 2). Removal of these resulted in a substantial attenuation of the mean effect size and the residual variance, whereas error variance remain relatively unaffected. At temperatures above 52°F (11.1°C), three outliers were identified. Inspection of the data indicated that effect sizes for these studies were much larger than for the others in that same category. Removal of these studies and reanalysis yielded a much smaller effect size and residual, with similar level of error variance.

Applying Hunter and Schmidt's (2004) 75% rule, we conclude that the six remaining studies are homogenous. However, the confidence interval associated with this latter mean does include zero. Thus, the analysis without the influential data points includes a null effect, whereas inclusion of these studies indicates a small to medium *positive* effect on performance. This is not attributable to changes in statistical power because the error variances in the two analyses are similar. Therefore, six of the nine studies in this category indicate a reliable null effect of moderate cold on performance.

Task Type

In Table 1 we have also parsed the temperature effects into three subdivisions based upon the differentiation of information-processing stages (see Lachman, Lachman, & Butterfield, 1979; Pilcher et al., 2002; Wickens & Hollands, 2000). This tripartite differentiation splits effects into perceptual, cognitive, and psychomotor response capacities, which is important because change in response capacity under thermal stressors is not equivalent in each case (see Grether, 1973; Hancock, 1982; Ramsey, 1995; Ramsey & Morrissey, 1978). These do not represent a specific model of human information processing but, rather, are general task performance categories. There are detrimental effects for thermal stressors on all three categories, with the highest impact being on perception, the next highest on psychomotor response, and the smallest on cognitive tasks. The variability of the effect sizes is again greater in the higher impact categories of perception and psychomotor response.

Performance Measures

In performance analysis, a primary division of the dependent variable is often split between the speed of a response and the accuracy of that response. In many circumstances speed is traded for accuracy and vice versa (Fitts, 1954). Although the number of effect sizes for accuracy was almost double that for response time, there was sufficient evidence to provide a stable representation of each.

Table 1 shows that thermal stressors generate substantial performance degradation on both speed and accuracy, with the larger effect being on speed. This outcome argues against a simple speed-accuracy trade-off being responsible for results previously observed, which, if in reality were present, would represent a strategic change rather than an absolute reduction in response capacity. The variability between the effect sizes that compose these separate influences on response time and accuracy is evidently high, although the error variances are comparable. With a threefold difference it is clear that response-time effect sizes are composed of individual effects of much greater variation across studies, indicating the likely influence of moderator variables on response time that are not so evident for accuracy.

Duration of Exposure

One crucial factor that permeates all stress effects is exposure time. To evaluate such effects, we subdivided studies according to the duration of exposure (see Table 1). Duration exerts a systematic effect, and in respect to the first three categories (up to 3 hr), there is a sequential increase in effect size with increasing exposure time. Thus, performance is affected adversely to an ever greater degree as time on task proceeds. Although this effect is relatively stable, the effect size variability that accompanies these temporal effects also increases.

This pattern is consistent with the expectation of a time/intensity effect in research on stress in general and the degree of variation also being a mark of progressive instability (see Hancock & Warm, 1989). However, this pattern is broken in the final category of exposure (beyond 3 hr). As indicated in Table 1, durations above 3 hr prove essentially as benign as exposures below 1 hr, and the error variances are similar. This outcome is, however, an artifact. A perusal of original studies shows that only the lower temperature ranges (i.e., less intense exposures) are tolerable for this extended period of time. How these duration effects are modulated by acclimatization is also surely an important issue. However, as indicated in Table 4, so few studies specifically included an acclimatization regimen that there is at present insufficient information to derive acceptable meta-analytic results on such effects. This is clearly an important issue for future research.

Our findings show an overriding interaction between exposure time and tolerable exposure temperature. This is not immediately evident, as meta-analytic effect sizes are initially the equivalent of main effects. The time by intensity outcome is fundamentally an interaction. This general issue is reexamined later when we discuss multifactor effect sizes. Results from the present meta-analysis serve to confirm duration of exposure effects, and this quantitative summation is consistent with previous, more qualitative observations on heat stressors (Hancock, 1981a, 1982, 1986a; Hancock & Vasmatzidis, 1998; Ramsey, 1995) and similar observations on stress in general (e.g., Hancock & Desmond, 2001; Hockey, Gaillard, & Coles, 1986).

Outlier analysis. The outlier analyses for the duration categories revealed influential data points in each category. Two outliers were observed in both the <1-hr category and the 1- to 2-hr category. Five outliers were observed in the 2- to 3-hr category, and three outliers were observed in the longest duration category (above 3 hr). After removal of these influential data points, reanalysis resulted in much smaller s_g^2 and a different pattern of effect sizes across duration categories (see Table 2). The mean effect sizes in the two lower duration categories and the highest duration category increased, whereas the effect in the 2- to 3-hr category decreased. Examination of the relevant confidence intervals suggests that there is a substantial increase in effect from <1 hr to the 1- to 2-hr category and that beyond this, the effect of duration varies as a function of other moderating variables and sampling error.

Hierarchical Analysis: Task Category and Performance Measure

The findings in regard to task category and performance measures were analyzed hierarchically by heat and cold individually (see Table 5). As is clear, heat stressors detrimentally influence all three performance categories. Most affected by heat stressors are perceptual capabilities, followed by motor responses, with cognitive attributes least affected but still subject to decrement. Thus the heat stressor findings replicate the overall thermal effects.

For cold, this pattern is again confirmed, with the motor and perceptual capacities being affected to the greatest degree. However, the pattern for cold is not exactly equivalent to that for heat. Perceptual abilities are most affected, but the level of that effect is greater (i.e., –1.13 vs. –0.78). Motor response represents the intermediate category, with similar levels of decrement in heat and cold (i.e., –0.42 vs. –0.31). However, a clear difference is evident in the area of cognition (see Table 6). The effect of cold on cognition is an *increment* in performance capacity, as compared with the decrement for heat (i.e., +0.41 vs. –0.23). This may be attributable to the transient warming effects of cold on core body temperature, as we discuss later.

As performance is differentiated by task, results can be divided by heat and cold into their effects on response time and accuracy (see Table 5). Heat is deleterious to both speed and accuracy, although variation in effect size distribution is greater for response time. Error variances are comparable, and

Continued on page 866

TABLE 4: Reports With Heat Acclimatization and Core Body Temperatures

Paper	Heat Acclimatization? If Yes, How Long?
Allen & Fisher, 1978 (Exp. 1)	No
Bateman, 1980	No
C. R. Bell & Provins, 1963	No
C. R. Bell, 1964 (Exp. 1)	No
C. R. Bell, 1964 (Exp. 2)	No
P. A. Bell, 1978	No
P. A. Bell et al., 1982	No
Blockley & Lyman, 1950	No
Bursill, 1958 (Exp. 1)	Yes: Artificially heat acclimatized over a period of a fortnight by exposing them for 5 days in 1 week to a condition of 95°/85°F (35°C/29.4°C) and a subsequent 5 days to 105°/95°F (40.6/35°C) for 3 hr daily, with an air velocity of 120 feet/min (~36.6 m/min).
Bursill, 1958 (Exp. 2)	Yes: Artificially heat acclimatized over a period of a fortnight by exposing them for 5 days in 1 week to a condition of 95°/85°F (35°C/29.4°C) and a subsequent 5 days to 105°/95°F (40.6/35°C) for 3 hr daily, with an air velocity of 120 feet/min (~36.6 m/min).
Chiles, 1958 (Exp. 1)	No
Chiles, 1958 (Exp. 2)	No
Cian et al., 2001	No
Cian et al., 2000	No
Colquhoun & Goldman, 1972	No
Courtright, 1976	No
Curley & Hawkins, 1983	Yes: 10-day acclimatization.
Epstein et al., 1980	No
Faerevik & Reinertsen, 2003	No
Givoni & Rim, 1962	No
Hocking et al., 2001	No
Holmberg & Wyon, 1969 (Exp. 1)	No
Holmberg & Wyon, 1969 (Exp. 2)	No
Hygge & Knez, 2001	No
Mackworth, 1946	No
Mayo, 1955	Yes: 7-month course in electronics, two classes held at different temperatures and compared for end of year grades.
Moreland & Barnes, 1970	No
Peccolo, 1962	Yes: Experimental conditions those standard in the classroom, that children are already used to (i.e., acclimatized to).
Pepler, 1960 (Exp. 2)	No
Pepler & Warner, 1968	No
Poulton & Kerslake, 1965	No
Razmjou, 1996 (Exp. 2)	No
Razmjou & Kjellberg, 1992	No
Reardon et al., 1998	Unclear: "Activities during the first week included uniform and helmet fitting, simulator training, and heat stress acclimatization in an environmental chamber" (p. 570).
Reddy, 1974	No
Reilly & Parker, 1988	No
Teichner & Wehrkamp, 1954	No
Tikuisis et al., 2002	No
Weiner & Hutchinson, 1945	Yes: "In the main series of experiments the 6 subjects, aged 25 to 35, were well acclimatized to working in hot humid environments" (p. 154).
Wyon, 1969	No

TABLE 5: Effects of Thermal Stressors

Category	k	$\bar{g}$	s_e^2	s_g^2	s_δ^2	s_e^2/s_g^2	95% CI (s_e^2)	n
Heat								
Perception	17	–0.78	0.12	2.04	1.92	0.06	$-0.95 < \delta < -0.62$	297
Cognitive	20	–0.23	0.04	0.44	0.40	0.09	$-0.31 < \delta < -0.15$	1477
Psychomotor	14	–0.31	0.09	0.40	0.32	0.23	$-0.47 < \delta < -0.15$	284
Accuracy	36	–0.33	0.06	0.58	0.52	0.10	$-0.41 < \delta < -0.25$	1625
Reaction time	20	–0.26	0.08	1.79	1.71	0.04	$-0.38 < \delta < -0.13$	561
Cold								
Perception	8	–1.13	0.17	0.76	0.59	0.22	$-1.42 < \delta < -0.84$	100
Cognitive	12	0.41	0.05	0.98	0.93	0.05	$0.28 < \delta < 0.54$	249
Psychomotor	8	–0.42	0.08	4.08	4.00	0.02	$-0.62 < \delta < -0.23$	160
Accuracy	18	0.05	0.07	1.22	1.15	0.06	$-0.07 < \delta < 0.18$	307
Reaction time	12	–0.11	0.07	3.79	3.72	0.02	$-0.25 < \delta < 0.04$	215
Perception								
Heat ET > 85°F (29.4°C)	17	–0.78	0.12	2.04	1.92	0.06	$-0.95 < \delta < -0.62$	297
Heat ET < 85°F (29.4°C)	—	—	—	—	—	—	—	—
Cognitive								
Heat ET > 85°F (29.4°C)	12	–0.01	0.05	0.69	0.64	0.07	$-0.13 < \delta < 0.12$	320
Heat ET < 85°F (29.4°C)	12	–0.34	0.04	0.30	0.26	0.13	$-0.45 < \delta < -0.23$	1388
Motor								
Heat ET > 85°F (29.4°C)	12	–0.32	0.09	0.50	0.41	0.18	$-0.49 < \delta < -0.14$	264
Heat ET < 85°F (29.4°C)	4	0.01	0.11	0.05	—	2.20	$-0.20 < \delta < 0.22$	192
Heat ET > 85°F (29.4°C)								
Accuracy	27	–0.37	0.09	0.89	0.80	0.10	$-0.49 < \delta < -0.26$	464
Reaction time	16	–0.33	0.09	2.47	2.38	0.04	$-0.47 < \delta < -0.18$	367
Heat ET < 85°F (29.4°C)								
Accuracy	13	–0.30	0.04	0.28	0.24	0.14	$-0.40 < \delta < -0.19$	1392
Reaction time	7	0.08	0.02	0.16	0.14	0.13	$-0.03 < \delta < 0.18$	298

Note. Results are shown as a function of task category and performance measure and by temperatures above and below the 85°F (29.4°C) effective temperature (ET) threshold of the "prescriptive zone" (Lind, 1963). Performance is differentiated by processing task demand and by changes in response capacity as expressed in measures of speed and accuracy.

TABLE 6: Thermal Stressor Papers With Positive Effect Sizes for Cold in the Cognitive Task Category

Report	Adjusted Effect Size
Enander, 1987	0.00
Griffiths & Boyce, 1971	0.03
Reddy, 1974	0.16
H. D. Ellis, Wilcock, & Zaman, 1985, Experiment 2	0.21
H. D. Ellis, 1982, Experiment 1	0.26
H. D. Ellis, 1982, Experiment 2	0.81
H. D. Ellis et al., 1985, Experiment 1	0.82
Pepler & Warner, 1968	1.33
Sharma & Panwar, 1987	1.84

thus the outcome is sufficiently stable to confirm these respective influences.

The cold influences are much less deleterious in general. Cold stressors reduce response time capability but only by a relatively small degree. The most interesting finding is that cold exerts almost no effect on accuracy. Here, the zero (no effect) is contained within the 95% confidence interval for both response time and accuracy. However, the large variance in *g* for the cold conditions indicates that the small effect sizes may be attributable to the influence of other moderating variables. The general conclusion, that heat is somewhat more damaging than cold, is in line with physiological accounts of thermal stressor effects in general (Hancock, 1986a).

Joint Effects of Intensity and Task Type

Subdivision of the respective stress effects has two conflicting tendencies. First, subdivision allows examination of the influences of multiple factors. As one cannot truly understand performance without both speed and accuracy measures, so one cannot understand stress effects without grasping the combined impact of time and intensity. Unfortunately, each subdivision reduces the number of effects involved and thus the reliability of the outcome. Soon, one is down to a single effect from a single study or, even worse, no effect at all.

This is illustrated in Table 5, where information is parsed by task demand category and heat stressor level above and below the 85°F (29.4°C) ET (see Hancock & Vercruyssen, 1988; Lind, 1963). This results in a very different pattern for heat effects contingent upon the type of task being performed. Unfortunately, perceptual capacity could not be examined because of the problem of the restricted number of effects, as mentioned earlier. Fortunately, this problem does not hold for the other task demand components.

Motor responses to heat prove very systematic. Above 85°F (29.4°C) ET there is a moderately strong decrement, and the error variance confidence interval excludes a zero effect. Conversely, motor response seems to be unaffected below the 85°F (29.4°C) ET threshold, as shown by the confidence interval, in which the range of *g* scores brackets the zero effect. The most interesting pattern, however, is evident for cognitive activity. Although there is essentially no effect upon cognitive tasks over the 85°F (29.4°C) ET threshold, there is a reasonably strong negative impact on cognitive performance below the 85°F (29.4°C) ET threshold. It would appear that studies employing a less intense heat stressor actually lead to a greater cognitive task decrement. This bifurcation of effects across the tasks apparently confirms earlier observation of 85°F (29.4°C) ET as a crucial threshold.

Outlier Analysis

Perceptual tasks. Among studies using perceptual tasks at temperatures above 85°F (29.4°C) ET, two outliers were identified. Removal of these resulted in a substantially smaller mean effect size, as well as smaller sampling error variance and residual variance.

Cognitive tasks. For studies with cognitive tasks and exposures below 85°F (29.4°C) ET, three outliers were identified (see Table 2). After removal of these studies, subsequent reanalysis revealed a much larger mean effect size and the residual variance was smaller. The error variance was comparable to that associated with the analysis with the original data included. For the group of studies in which cognitive tasks were used at exposures greater than 85°F (29.4°C) ET, two outliers were identified (Table 2).

Psychomotor tasks. Table 2 also shows that the evaluation of motor tasks at temperatures above 85°F (29.4°C) ET revealed two outliers. However, removal of these from analysis did not substantively change the mean effect size or the error variance, although the residual variance was cut in half. No outliers were observed at temperatures below 85°F (29.4°C) ET, possibly because of the limited number of studies and the absence of any residual variance.

Performance Measures Within Each Heat Stressor Category

As response time and accuracy are differentially affected by heat stressors, we again subdivided these influences using the heat threshold we have previously identified (Table 6). This shows that accuracy is ubiquitously affected by heat stressors, regardless of whether the heat is above or below the 85°F (29.4°C) ET threshold. The outcome for response time is rather different. Below the threshold value there is essentially no effect for heat on response time, and by implication heat stressor degradation in this region is attributable fundamentally to a change in response accuracy. However, as one enters the range of heat stressors that cannot be compensated for by normal physiological response (Hancock, 1982; Houghton & Yagloglou, 1923; Lind, 1963), response time in relation to imposed task demands increases and thus performance becomes significantly poorer as both aspects of response capacity are affected.

Outlier analysis. Outlier analyses of accuracy and speed effects within each temperature category revealed substantively different conclusions when these influential data points were removed from the analysis. Thus, at temperatures above 85°F (29.4°C) ET, two outliers were observed for the accuracy analysis, and removal of these yielded similar values of mean effect size and error variance for the ET > 85°F (29.4°C) category and less residual variance. For the speed analysis, four outliers were identified. However, the analysis of speed in the same heat category indicated that removal of the outliers resulted in a negligible effect. Even with an associated smaller sampling error variance, the confidence interval still included zero. However, the residual variance was over four times greater than the error variance.

For studies utilizing exposures below 85°F (29.4°C) ET, three outliers were identified for accuracy and one for speed (Holmberg & Wyon, 1969). Removal of these from the analysis resulted in a smaller effect for accuracy, with similar error variance but smaller residual variance. By contrast, removal of one outlier for the speed analysis changed the outcome from a null to a small negative effect. Thus, accuracy is impaired more above 85°F (29.4°C) ET, whereas impairment of speed occurs more at temperatures below 85°F (29.4°C) ET. Although this is not a true speed-accuracy trade-off, it does reflect differences in the effect of heat stressors dependent upon the order of performance examined.

Joint Effects of Intensity and Duration

In Table 7, we parsed the data according to thermal level and exposure duration. Although this division promises to be most informative, it does reduce the number of some effect sizes to a problematic level. We expected to see a time by intensity interaction such that the longer exposures at the higher thermal stressor levels would prove to be the most deleterious. For heat stressors we did not see this pattern. Rather, it is the shorter exposures that result in greater degradation. This is surprising but may reflect Poulton's (1976) invocation of an acclimation factor. Even the variability of the effect-size factor is not an explanation for this effect. This observation has important implications for length of duty in heat stressor conditions, as we will discuss later. This temporal factor also plays a role in activities such as mission planning and shift work.

Cold provides a clearer picture. Above approximately 52°F (11.1°C) ET cold acts to improve performance, but as exposure time increases this improvement is attenuated. Below 52°F (11.1°C)

TABLE 7: Time by Temperature Effects for Heat and Cold Stressors

Category	k	$\bar{g}$	s_e^2	s_g^2	s_δ^2	s_e^2/s_g^2	95% CI (s_e^2)	n
				Heat				
<ET 85°F (29.4°C)								
<120 min	7	−0.47	0.05	0.23	0.18	0.22	$-0.64 < \delta < -0.31$	330
>120 min	3	0.03	0.01	0.01	0	1.00	$-0.08 < \delta < 0.14$	809
>ET 85°F (29.4°C)								
<120 min	19	−0.29	0.09	0.67	0.58	0.13	$-0.43 < \delta < -0.15$	187
>120 min	7	−0.03	0.05	0.57	0.52	0.09	$-1.19 < \delta < 0.13$	143
				Cold				
<ET 52°F (11.1°C)								
<120 min	12	−0.71	0.15	2.10	1.95	0.07	$-0.93 < \delta < -0.50$	116
>120 min	2	0.13, 6.03[a]						28
>ET 52°F (11.1°C)								
<120 min	4	0.66	0.11	1.61	1.50	0.07	$0.33 < \delta < 0.99$	94
>120 min	4	0.42	0.02	1.46	1.44	0.01	$0.27 < \delta < 0.57$	125

[a] Because only two studies exist in this category, each effect size is reported rather than the mean.

ET, durations of less than 1 hr are more deleterious than anything over 52°F (11.1°C) ET (see Table 7). Problematically, estimating the effects of exposures over 2 hr at lower temperatures was limited by the very small number of effect sizes. This latter estimate depends on only two effect sizes, each of which provides a very different outcome. For now, it is best to treat this result as unreliable and adopt a conservative approach by considering this condition at least equally as damaging as less than a 2-hr duration. Given the importance of such conditions, exploring such conditions may well be viewed as a priority area for future research.

Outlier analyses: Heat. Examination of the duration by intensity interaction for heat indicated no outliers among effects greater than 2 hr, although this might be attributable to the small number of studies in these categories. Outliers were, however, observed among studies with less than 2-hr durations. Thus, for the ET > 85°F (29.4°C) category, four outliers were identified. Removal of these studies and reanalysis resulted in a smaller mean effect size but a comparable sampling error variance. However, the total variance was much lower, and the residual variance was reduced to zero upon removal of these latter influential data points. This indicates that the apparent heterogeneity in this category is attributable to 4 of the 19 total studies that compose the group (see Table 2).

For the group of studies in which ET < 85°F (29.4°C) and duration of exposure was less than 2 hr, one outlier was identified. Removal of this study reduced the mean effect size and resulted in a lower residual variance and slightly lower error variance. Considering these mean differences and the high degree of overlap between confidence intervals, the apparently anomalous effect in the original analysis – that temperatures *greater* than 85°F (29.4°C) ET induced a *smaller* effect – is likely attributable to the influence of the outlier studies identified. This is further evidence that the utility of Lind's (1963) prescriptive zone in understanding performance effects is dependent on other moderating conditions, in addition to duration and temperature of exposure.

Outlier analyses: Cold. Analysis of the duration and cold temperature category interaction also results in a different interpretation of the data. Only two studies employed the combination of exposure temperatures less than 52°F (11.1°C) and durations greater than 2 hr. For durations less than 2 hr, two outliers were identified. The sample sizes for these latter studies were both $n = 20$, whereas the other 10 studies in that category all had $n \leq 12$. Removal of these studies and reanalysis yielded a smaller mean effect size associated with substantially smaller residual variance. Sampling error variance was similar, however. The interpretation therefore did not change with removal of the outliers, except that the new magnitude is in the range of a medium rather than a large effect.

A different story emerges at temperatures above 52°F (11.1°C), however. At durations greater than 2 hr, one outlier was identified. Reanalysis following removal of this study resulted in a mean effect of similar magnitude but in the opposite direction. The effect size for this study was $g = 1.33$, whereas those of the other three were $g = -1.03$, -0.75, and 0.21. Further, the small number of studies resulted in an increase in error variance and a substantial decrease in residual variance. Thus, an interpretation of heterogeneity of variance with all four studies changes to one of homogeneity with the study of Pepler and Warner (1968) removed.

At durations less than 2 hr, the removal of the one outlier resulted in a much smaller and negative mean effect size, although the confidence interval contained zero. This is unlikely to be attributable to a change in power, as omission of Lockhart (1968) reduced sampling error variance from $s_e^2 = 0.11$ to $s_e^2 = 0.08$. Residual variance was reduced to zero, indicating that the heterogeneity among the larger effect sizes observed in the original analysis is attributable almost entirely to the Lockhart (1968) study. Inspection of the individual effect sizes reveals the reason for this pattern. For the Lockhart (1968) study the value of $g = 2.60$, whereas for the other three studies the effect sizes were $g = -0.10$, -0.36, and 0.03.

Hierarchical Analysis: Joint Effects of Task and Performance Measure

In Table 8, we have extended the process of decomposition to its greatest practical degree. The division is by thermal stressor into heat and cold, and within this division we have looked at task category and dependent variable. It is perhaps easiest to deal with these categories sequentially, and the first observation is that heat stressors deleteriously influence perception through a reduction in response accuracy and an increase in response time. For the cognitive task category, heat reduces response accuracy but has a minimal impact on response time.

TABLE 8: Breakdown of Heat and Cold Stressor Effects by Task Category and Performance Measure

Category	k	$\bar{g}$	s_e^2	s_g^2	s_δ^2	s_e^2/s_g^2	95% CI (s_e^2)	n
Heat								
Perception								
Accuracy	12	–0.41	0.10	0.27	0.17	0.37	$-0.59 < \delta < -0.22$	123
Reaction time	11	–0.91	0.13	3.13	3.00	0.04	$-1.12 < \delta < -0.70$	248
Cognitive								
Accuracy	19	–0.27	0.04	0.57	0.53	0.07	$-0.36 < \delta < -0.18$	1461
Reaction time	4	0.02	0.03	0.01	—	3.00	$-0.14 < \delta < 0.19$	140
Psychomotor								
Accuracy	13	–0.59	0.13	0.72	0.59	0.18	$-0.78 < \delta < -0.40$	256
Reaction time	5	0.68	0.13	1.32	1.19	0.10	$0.37 < \delta < 1.00$	77
Cold								
Perception								
Accuracy	6	–1.07	0.16	1.20	1.04	0.13	$-1.39 < \delta < -0.74$	77
Reaction time	5	–0.85	0.14	0.86	0.72	0.16	$-1.18 < \delta < -0.52$	46
Cognitive								
Accuracy	10	0.05	0.05	0.56	0.51	0.09	$-0.09 < \delta < 0.19$	225
Reaction time	6	0.64	0.19	0.47	0.28	0.40	$0.29 < \delta < 0.99$	133
Psychomotor								
Accuracy	5	0.58	0.09	1.78	1.69	0.05	$0.32 < \delta < 0.85$	100
Reaction time	4	–1.10	0.44	3.07	2.63	0.14	$-1.75 < \delta < -0.45$	76

In the psychomotor task category, heat stressors exert strategic effects. The meta-analysis shows that heat reduces psychomotor accuracy; however, the evidence also shows *facilitation* in response time. The effect sizes indicate that the facilitation in speed is greater than the decrement in accuracy; however, whether the absolute trade-off of these two causes an overall deficiency or increment is still uncertain, as such trade-offs are contingent upon the character of the specific task at hand.

The results for cold stressors are also systematic but again sparse. The obvious thing is that cold, like heat, has a significant decrement on both the speed and accuracy of perception. For cognitive tasks, however, the effect size associated with accuracy was small, with the confidence interval including zero. In contrast, cold temperatures produced evidence of a moderately large *increment* in speed for cognitive tasks ($g = 0.64$; i.e., better or faster response time). This is also evidence of global speed-accuracy trade-off, and again the absolute effect of this is difficult to determine, being task specific.

In respect of psychomotor activity, there is a further anomalous finding. The average effect size for accuracy shows an evident *increment* in performance (see Poulton, 1976). This is a real effect but one that would be very much interrupted with the onset of shivering. Response time shows a decrement indicating a different form of speed-accuracy trade-off than that observed in heat stressor conditions. Thus, in the presence of heat stressors individuals engaged in psychomotor tasks are faster but less accurate, whereas the reverse is apparently the case in cold environments.

Outlier Analyses: Heat Stressors

Perceptual tasks. Three outliers were observed for the accuracy analysis and two for the speed analysis (see Table 2). With these outliers included, speed was negatively impacted more than accuracy, but the effects were both negative and there was evidence of heterogeneity in the variability across effect sizes. When these outliers were removed, the mean effect sizes are of similar magnitude, and error variance was similar for accuracy and smaller for speed. Although the residual variance was substantially smaller in both cases, there was still evidence of heterogeneity in effect sizes for speed when applying the 75% rule. In contrast, the residual variance for accuracy was only 9% of the total variance, indicating a high degree of consistency in the effect of heat on perceptual performance accuracy (i.e., the effect sizes were homogeneous).

Cognitive tasks. Five outliers were observed for the accuracy category (Table 2). Four of these ranged in effect from $g = -1.16$ to -1.71. The exception was Blockley and Lyman (1950), who obtained an effect size of $g = 2.63$. Removal of these outliers and reanalysis yielded a smaller effect with a similar error variance. The residual variance was cut in half, but the pattern still indicated heterogeneity in the remaining effect. For speed, no outliers were observed. Thus, the outlier analyses for studies using cognitive tasks reveal no substantively different conclusions, with or without the outliers included, with the one exception of the analysis for the magnitude of the mean effect size for accuracy.

Psychomotor tasks. Two outliers were observed in the accuracy condition and one outlier in the speed condition (Table 2). Removal of the effect size obtained from P. A. Bell (1978) did not substantively change the outcome of the analysis. In regard to speed, removal of the outlier reduced the magnitude of the mean effect size, and it also substantially reduced both the error variance and the residual variance. Thus, the remaining four studies are homogeneous with respect to the effect of heat on psychomotor speed.

Cold

Cognitive tasks. There was one outlier/influential data point with respect to cognitive task accuracy (Pepler & Warner, 1968) and one for speed (van Orden, Benoit, & Osga, 1996). Removal of each outlier resulted in a smaller effect for accuracy, with larger error variance and slightly smaller residual variance. For speed, the mean effect size and error variance were slightly smaller without the outlier, and the residual variance was reduced by more than half. However, there was still evidence of heterogeneity, so the substantive interpretation does not change when this outlier is removed.

Psychomotor and perceptual tasks. One influential data point was identified for accuracy. Removal of this study resulted in a smaller effect in the opposite direction, although the 95% confidence interval (CI) included zero, rendering this effect somewhat unreliable ($\bar{g} = -0.24$, 95% CI: $-0.54 < \delta < 0.06$). This change is not attributable to less statistical power without that study, as the error variances of the two analyses were the same ($s_e^2 = 0.09$). Further, the observed variance was substantially smaller ($s_g^2 = 0.06$), indicating that $s_\delta^2 = 0$. Thus, the variability among the four remaining studies can be attributed entirely to sampling error (i.e., the effect sizes are homogenous). No outliers were observed for the studies using perceptual tasks.

One outlier was identified for speed (Lockhart & Kiess, 1971), and removal of this study rendered a change of mean effect from $\bar{g} = -1.10$ to $\bar{g} = 0.06$, with zero included in the 95% CI ($-0.17 < \delta < 0.28$). However, the error variance ($s_e^2 = 0.04$) and the residual variance ($s_\delta^2 = 0.20$) were also substantially reduced. Given the small number of studies, regardless of whether the outlier is included or excluded, such results should be interpreted with caution. However, the effect size associated with the Lockhart and Kiess (1971) study ($g = -4.35$) was substantially larger than that associated with the other three ($g = 0.25$, 0.37, and -0.87). Thus, this study obtained results quite different from those of the other studies in the same category.

DISCUSSION

The current meta-analysis is in general agreement with those of Pilcher et al. (2002) and confirms their proposition that variables such as duration and intensity of exposure moderate the relation between exposure to thermal stressors and human performance. This study also extended the findings of Pilcher et al. (2002) in several respects. First, this analysis established that the magnitude, and in some cases the direction, of the effect of thermal conditions on performance depended on particular combinations of exposure range (i.e., heat/cold), task type, performance measure, and the duration and intensity of exposure.

Second, the moderator variables influenced not only the mean effect size but also the variance associated with it, revealing a substantial degree of heterogeneity. Third, analysis of influential data points revealed that the magnitudes of the variances were, in some cases, influenced by outlier studies. Finally, these analyses have important implications for future empirical and theoretical work. We will examine the theoretical implications of this work and that of Pilcher et al. (2002).

With any meta-analysis, what can be derived from the procedure is a quantitative description of the state of present understanding to the degree that it can be known. This description is an empirical map that can accommodate competing theoretical accounts, at least to the degree that any quantitative data can address a theory for which they were not initially derived to evaluate. Although we are very interested in the various theories that have been developed to account for stress effects in general (e.g., Lazarus & Folkman, 1984; Selye, 1976; Ursin & Eriksen, 2004) and thermal stressors in particular (Boregowda, 1999; Enander, 1989; Enander & Hygge, 1990), we wish to focus first upon the different descriptive functions that have sought to link stress to performance efficiency.

A number of proposed relationships describe the function that relates the level of stress to performance capacity. The most influential and ubiquitous of these is the inverted U, which dominates the landscape, especially in undergraduate and beginning psychology texts. First derived from experiments on discrimination learning in mice (Yerkes & Dodson, 1908), the general function was given particular impetus by Hebb's (1955) influential affirmation. Despite an ongoing litany of criticisms, the inverted U still inhabits this dominant position, and despite its obvious flaws, it continues to be taught as one of the fundamental "laws" of psychology (see Hancock & Ganey, 2003). We do not rehearse these various objections here. Instead we explore what the present meta-analytic findings mean for this and the other descriptive relationships that have been proposed.

Although the inverted U is the dominant descriptive relationship (which we have illustrated only in its general form; see Figure 1a), it is by no means the only one. For example, Näätänen (1973) argued that if the individual could keep his or her attention focused directly on task-relevant cues and not become distracted or overwhelmed by irrelevant stimuli, then there should be no descending arm to the inverted U. In such circumstances performance should continue to increase as stress grows, but with the rate of gain in performance efficiency decreasing progressively with each sequential increase in the level of imposed stress. This results in the function illustrated in Figure 1b, as compared with the classic inverted U.

This proposal has, most recently, been elaborated and developed by Gaillard (2005), who argued for the utility of the construct of "concentration" on distinguishing changes associated with change in stress level. It is also consistent with the findings that emotional states associated with high arousal (particularly fear) reduce the subjective perception of pain, thereby mitigating the effects of such pain-inducing environmental stimuli.

As an elaboration of these earlier postulations, Hancock and Warm (1989) developed the "extended-U" description of stress and performance change, which is illustrated in Figure 1c (see Hancock & Warm, 1989). In this conception, stress in the real world is conceived as being a tolerable form of interference that results in only minor levels of change up to specific threshold limits. At these thresholds (which vary as a function of the type of behavior measured also changes), each of the various compensatory mechanisms that act to mitigate stress effects begins to fail, and their respective capacities are rapidly overwhelmed as adaptive response now proves insufficient for full compensation (see Figure 2).

Thus Hancock and Warm's (1989) description shows a central, plateau region where performance is *relatively* stable, bound by regions of evident failure characterized by sudden and radical failure and, hence, the reference to this description as the extended U. As is evident in Figure 2, the threshold at which comfort fails is much lower in terms of stress level, as compared with psychological functioning, which itself is affected before the onset of physiological failure. This more general

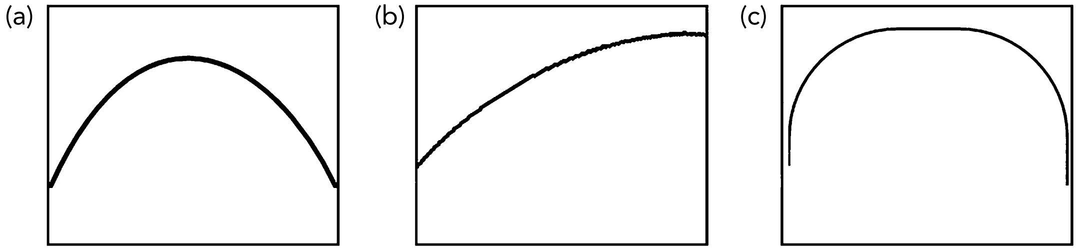

Figure 1. Stress-performance functions associated with three prominent stress theories of human performance. (a) The Yerkes-Dodson (1908) inverted-U description; (b) Näätänen's (1973) sequential increase description; (c) the Hancock and Warm (1989) extended-U description.

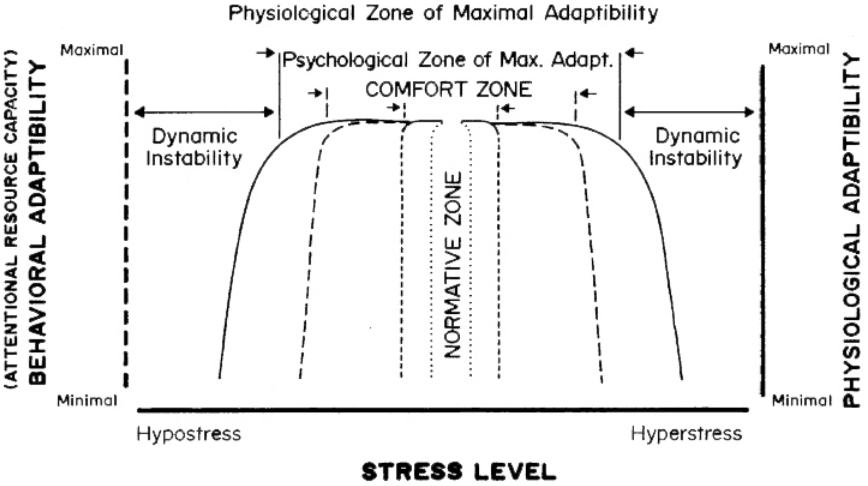

Figure 2. The extended-U model describing the relation of stress to performance and adaptation. From Hancock and Warm (1989).

description was postulated to apply to physiological as well as performance degradation, in which there is a formal relation between the progressive failures in physiological and psychological functioning, respectively (Figure 2; see also Hancock & Warm, 1989).

As further experimental evaluations of thermal effects proceed, it is critical to provide more quantitative assessments of this postulated isomorphism between physiological change and performance degradation. Confirmation of this linkage would serve to reinforce the notion that perceptual and cognitive performance change can itself be used as a stress assessment index. This formal integration would bring together conceptions of stress from the distinct worlds of medical evaluation and ergonomic assessment.

Given these descriptions of individual response at the different physiological and psychological levels of analysis, the observed global decrement in performance with exposure to thermal stressors (i.e., both heat and cold stressors) can now be interpreted as support for both the inverted-U and extended-U hypotheses. This outcome, however, can be seen to somewhat contradict Näätänen's (1973) description of the stress/performance function. However, the post hoc explanation for this latter discrepancy would be that participants failed to focus on task-relevant cues, a proposition that is not immediately testable from any of the studies examined. Post hoc rationalization is always of concern, especially when postulating an unobserved (and sometimes unobservable) variable as the crucial causal factor.

However, performance did prove to be relatively stable over much of the tolerable temperature range and then showed radical decrement at the highest extremes. This pattern is particularly consistent with the propositions intrinsic to the extended-U theory of stress and performance and directly contradicts the inverted-U description, which explicitly defines a single, optimal point of response efficiency. As we have examined explicitly different orders of performance, the usual post hoc adjustment contingent upon an unidentified level of so-called task complexity cannot now save the inverted-U description, as it has done in many previous commentaries (but see Hancock & Ganey, 2003). We also observed that there is, in general, an increased negative impact on performance with increased duration of exposure to thermal stressors, although this generalization must be tempered with respect to the specific observations we have made on the longest time intervals evaluated.

For our generalized findings concerning thermal stressors the Hancock and Warm (1989) model fits with the data very well. However, as analyses get more fine grained, it becomes necessary to make more specific inferences in regard to task, dependent measure, and other factors. For instance, we found a great deal of heterogeneity in our lowest heat temperature range, with two studies with large effects (–1.71: Allen & Fischer, 1978; –1.45:

Peccolo, 1962) driving the overall negative effect in this category, whereas all the other studies demonstrated small effects around the zero level (–0.16: Holmberg & Wyon, 1969; 0.16: Pepler & Warner, 1968; 0.06: Wyon, 1969). This indicates a greater need for specificity in the division of studies at this level to discover the driving moderating forces of these disparate results in what are expected to be relatively homogenous results.

Another example in which there was a great deal of disparity was in regard to task effects. Perceptual and psychomotor effects frequently proved much larger than cognitive effects. This also supports the postulation in the Hancock and Warm (1989) model of dynamic maximal adaptability in regard to its diagnostic ability. Hancock and Warm (1989) regarded the immediate task as the proximal source of threat and disturbance, but not all tasks are equally stressful. Some tasks, such as perceptual and motor demands and their combination in dual-task situations, tend to be more attention demanding and therefore fail earlier and to a greater degree than do more cognitive tasks (see Vasmatzidis, Schlegel, & Hancock, 2002).

Whereas most task types were impacted by heat and cold in a relatively similar manner, cognition actually benefited from cold but was negatively affected by heat. The positive effect of cold on cognition, as compared with the degradation by heat (i.e., +0.41 vs. –0.23), might be attributable to the transient warming effects of cold on core body temperature through the thoracic pooling of blood (see Enander, 1984, 1986; Poulton, 1976). This is supported by the analysis of a median split in cold temperature ranges that demonstrated a beneficial effect for slight cold but a large decrement in extreme cold, in which such transient warming effects would be eventually offset by long-term cooling. Therefore, it was the nature of the experiments, with more studies implementing milder cold stressors, that leads to the recorded beneficial effect. This serves as a reminder that stress is not always a bad thing.

In regards to duration, heat seemed to demonstrate some degree of benefit for increased duration, whereas cold appeared to not demonstrate a benefit with increased duration. However, this difference must be tempered by knowledge of the actual conditions and recognition that one can acclimate to heat but acclimation to cold is presently thought to be untenable. The present analysis was unable to examine the effect of heat acclimatization on performance directly, given an insufficient number of qualifying studies, but there is a strong theoretical rationale for its prophylactic effects. Thus, it is important for both theoretical and practical purposes that such experiments be conducted so that quantification of this benefit can be specified. Specifying the influence of this and other such modifying factors, especially adaptation to the stress and to the task (Hancock, 1986a), awaits future collective efforts.

Our final observation is that heat generated a common decrement in both accuracy of response and in response time, but cold demonstrated a decrement only to response time. However, when heat was divided along Lind's (1963) threshold, the lower temperature range showed no decrement for response time.

CONCLUSIONS

Hancock and Warm (1989) argued that the nature of the task itself is a crucial influence on the individual's response to stress. The current analysis supports this contention by showing that the effect of thermal environments on human performance varies as a function of task type (see Grether, 1973). It is likely, however, that task type interacts with other moderators – particularly exposure duration and intensity – in influencing both the magnitude and direction of the effect of thermal stressors (Hancock & Vasmatzidis, 2003; Pilcher et al., 2002).

Further, even in the hierarchical analyses, the variances within most categories indicated the likely influence of other moderators. In such cases, however, there were insufficient studies to explicate these more complex interactions. Thus, future work should examine how these variables jointly influence performance under conditions of exposure to thermal stressors. If such interactions could be explored and quantified (e.g., Loeb & Jeantheau, 1958), it would permit the construction of vectors representing the collective impact of the variables on adaptation under stress (see Hancock & Warm, 1989).

PRACTICAL APPLICATIONS

One of the issues involved in using stress research for design is the difficulty in quantifying precisely how a particular stressor influences human performance (Poulton, 1965). The current meta-analytic review represents one step toward a solution to this problem by providing quantitative estimates of the effect of a particular stressor on performance of a variety of task types. It is only a partial solution, as it does not address how multiple stressors (e.g., temperature, noise, time pressure, task complexity) combine to influence response outcome (see Hancock & Pierce, 1985). Nor, at the present, does it provide completely reliable estimates of specific combinations of task and environmental characteristics even within one stressor category (e.g., the joint effect of task, temperature range, and dependent measure).

As a result of still-evident limitations on the present empirical database, a further limitation of all meta-analyses is that one can test only for combinations of moderator variables that have been investigated to a sufficient degree. For example, the use of acclimatized individuals exposed to heat stressors is an advisable strategy, and yet much needs to be done in relation to perceptual and cognitive performance to show how such acclimatization influences the performance on tasks of primarily mental demand. The current meta-analytic review underscores the necessity for further empirical research that is purposefully targeted at elucidating these complex interactions.

The mean effect sizes reported here represent the current "best estimates" of thermal effects on performance and can therefore be used to estimate the degree to which particular task-environment combinations will result in performance decrement or performance increment (e.g., Poulton & Kerslake, 1965). For instance, these results indicate that certain combinations of task and environmental characteristics are more likely to induce performance decrement (e.g., performance on perceptual tasks under cold conditions) than are other combinations (e.g., response time in cognitive tasks with exposure to heat stressors). We can therefore recommend their use as guidelines for practitioners who wish to incorporate the effect of thermal stressors on operators into their systems design.

It is not only military personnel who have to work under highly adverse thermal conditions. From the glass blowers of Venice to the crab fishermen of Alaska, many occupations require that workers face a significant thermal hazard in their everyday occupations. Traditional ergonomic strategies of isolation, augmenting support, or restricting exposure times continue to be applicable. However, as researchers and practitioners seek to defend people against the deleterious effects of thermal stressors, it is crucial to capture a quantitative assessment of the composition of such threats. The present meta-analysis provides this knowledge to those who design for, supervise, or operate alongside personnel in adverse thermal environments.

ACKNOWLEDGMENTS

The views expressed here are those of the authors and do not necessarily reflect the official policy or position of the Department of the Army, the Department of Defense, or the U.S. government. This research was specifically supported by a grant from the Army Research Laboratory through Micro Analysis and Design, P. A. Hancock and J. L. Szalma, Principal Investigators (Grant #64018042). We wish to thank Dr. Laurel Allender, John Lockett, Sue Archer, and Celine Richer, for providing administrative and technical direction for the grant.

This work was also facilitated by the Department of Defense Multidisciplinary University Research Initiative (MURI) program, P. A. Hancock, principal investigator, administered by the Army Research Office under Grant DAAD19-01-1-0621. The authors also wish to thank Dr. Sherry Tove, Dr. Elmar Schmeisser, and Dr. Mike Drillings for providing administrative and technical direction for the latter grant. Further, we are very appreciative of the insightful comments of the three anonymous reviewers provided on an earlier version of this paper, for which we are indebted.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis. Not all such references are cited.

*Allen, M. A., & Fischer, G. J. (1978). Ambient temperature effects on paired associate learning. *Ergonomics, 21*, 95–101.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

*Angus, R. G., Pearce, D. G., Buguet, A. G. C., & Olsen, L. (1979). Vigilance performance of men sleeping under arctic conditions. *Aviation, Space, and Environmental Medicine, 50,* 692–696.

Aschoff, J. (1984). Circadian timing. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (pp. 442–468). New York: New York Academy of Sciences.

*Bateman, R. P. (1980). *The effects of increased ambient temperature on dual task performance*. Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

*Bell, C. R. (1964). Climate, body temperature, and vigilance performance. In *Proceedings of the Second International Congress on Ergonomics* (pp. 169–172). London: Taylor & Francis.

*Bell, C. R., & Provins, K. A. (1963). Relation between physiological responses to environmental heat and time judgments. *Journal of Experimental Psychology, 66,* 572–579.

*Bell, P. A. (1978). Effects of noise and heat stress on primary and subsidiary task performance. *Human Factors, 20,* 749–752.

*Bell, P. A., Loomis, R. J., & Cervone, J. C. (1982). Effects of heat, social facilitation, sex differences, and task difficulty on reaction time. *Human Factors*, *24,* 19–24.

*Beshir, M. Y., El-Sabagh, A. S., & El-Nawawi, M. A. (1981). Time on task effect on tracking performance under heat stress. *Ergonomics, 24,* 95–102.

Blagden, C. (1775a). Experiments and observations in an heated room. *Philosophical Transactions, 65,* 111–123.

Blagden, C. (1775b). Further experiments and observations in an heated room. *Philosophical Transactions, 65,* 484–494.

*Blockley, W. V., & Lyman, J. H. (1950). *Studies of human tolerance for extreme heat: III. Mental performance under heat stress as indicated by addition and number checking tests* (Tech. Rep. 6022). Wright Patterson Air Force Base, OH: U.S. Air Force.

Blockley, W. V., & Lyman, J. H. (1951). *Studies of human tolerance for extreme heat: IV. Psychomotor performance of pilots as indicated by a task simulating aircraft instrument flight* (Tech. Rep. 6521). Wright Patterson Air Force Base, OH: Air Force Research Laboratories.

Boregowda, S. C. (1999). Thermodynamic modeling and analysis of human stress responses. *Dissertation Abstracts International, 60*(3-B), 1185.

Brief, R. S., & Confer, R. G. (1971). Comparison of heat stress indices. *American Industrial Hygiene Association Journal, 32,* 11–16.

*Bursill, A. E. (1958). The restriction of peripheral vision during exposure to hot and humid conditions. *Quarterly Journal of Experimental Psychology, 10,* 113–129.

Burton, A. C., & Edholm, O. G. (1955). *Man in a cold environment: Physiological and pathologic effects of exposure to low temperatures.* London: Edward Arnold.

Cannon, W. B. (1932). *The wisdom of the body.* New York: W. W. Norton.

Cheung, S. F., & Chan, D. K. S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology, 89,* 780–791.

*Chiles, W. D. (1958). Effects of elevated temperatures in performance of a complex mental task. *Ergonomics, 2,* 89–96.

*Cian, C., Barraud, P. A., Melin, B., & Raphel, C. (2001). Effects of fluid ingestion on cognitive function after heat stress or exercise-induced dehydration. *International Journal of Psychophysiology, 42,* 243–251.

*Cian, C., Koulmann, N., Barraud, P. A., Raphel, C., Jimenez, C., & Melin, B. (2000). Influences of variations in body hydration on cognitive function: Effect of hyperhydration, heat stress, and exercise-induced dehydration. *Journal of Psychophysiology, 14,* 29–36.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Colquhoun, W. P., & Goldman, R. F. (1972). Vigilance under induced hypothermia. *Ergonomics, 15,* 621–632.

*Courtright, J. F. (1976). *Effects of whole and partial body exposure to dry heat on certain performance measures.* Unpublished master's thesis, Purdue University, West Lafayette, IN.

*Curley, M. D., & Hawkins, R. N. (1983). Cognitive performance during a heat acclimatization regimen. *Aviation, Space, and Environmental Medicine, 54,* 709–713.

Ellis, H. (1758). An account of the heat of the weather in Georgia. *Philosophical Transactions, 50,* 754–756.

*Ellis, H. D. (1982). The effects of cold on the performance of serial choice reaction time and various discrete tasks. *Human Factors, 24,* 589–598.

*Ellis, H. D., Wilcock, S. E., & Zaman, S. A. (1985). Cold and performance: The effects of information load, analgesics, and the rate of cooling. *Aviation, Space, and Environmental Medicine, 56,* 233–237.

Enander, A. E. (1984). Performance and sensory aspects of work in cold environments: A review. *Ergonomics, 27,* 365–378.

Enander, A. E. (1986). *Sensory reactions and performance in moderate cold.* Unpublished doctoral dissertation, University of Uppsala, Uppsala, Sweden.

*Enander, A. E. (1987). Effects of moderate cold on performance of psychomotor and cognitive tasks. *Ergonomics, 30,* 1431–1445.

Enander, A. E. (1989). Effects of thermal stress on human performance. *Scandinavian Journal of Work, Environment, and Health, 15*(Suppl.), 27–33.

Enander, A. E., & Hygge, S. (1990). Thermal stress and human performance. *Scandinavian Journal of Work, Environment, and Health, 16*(Suppl.), 44–50.

*Epstein, Y., Keren, G., Moisseiev, J., Gasko, O., & Yachin, S. (1980). Psychomotor deterioration during exposure to heat. *Aviation, Space, and Environmental Medicine, 51,* 607–610.

*Faerevik, H., & Reinertsen, R. E. (2003). Effects of wearing aircrew protective clothing on physiological and cognitive responses under various ambient conditions. *Ergonomics, 46,* 780–799.

Fanger, P. O. (1967). *Thermal comfort.* Copenhagen, Denmark: Danish Technical Press.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Gaillard, A. W. K. (2005, July). *Concentration – An instrument to augment cognition*. Paper presented at the 1st International Conference on Augmented Cognition, Las Vegas, NV.

*Givoni, B., & Rim, Y. (1962). Effect of the thermal environment and psychological factors upon subjects' responses and performance of mental work. *Ergonomics, 5,* 99–114.

Goldman, R. F. (2001). Introduction to heat-related problems in military operations. In K. B. Pandolf, R. E. Burr, C. B. Wenger, & R. S. Pozos (Eds.), *Medical aspects of harsh environments.* (pp. 3–49). Washington, DC: Borden Institute.

Goodman, D., Hancock, P. A., Runnings, D. W., & Brown, S. L. (1984). Temperature induced changes in neuromuscular function: Central and peripheral mechanisms. *Perceptual and Motor Skills, 59,* 647–656.

Grether, W. F. (1973). Human performance at elevated environmental temperatures. *Aerospace Medicine, 44,* 747–755.

*Griffiths, I. D., & Boyce, P. R. (1971). Performance and thermal comfort. *Ergonomics, 14,* 457–468.

Hancock, P. A. (1980). Simulated and experimental temperature responses in man during exercise in varying environments. *Computers in Biology and Medicine, 10,* 1–9.

Hancock, P. A. (1981a). Heat stress impairment of mental performance: A revision of tolerance limits. *Aviation*, *Space, and Environmental Medicine*, *52,* 177–180.

Hancock, P. A. (1981b). The simulation of human core temperature. *International Journal of Bio-Medical Computing, 12,* 59–66.

Hancock, P. A. (1982). Task categorization and the limits of human performance in extreme heat. *Aviation*, *Space, and Environmental Medicine, 53,* 778–784.

Hancock, P. A. (1984). Effect of environmental temperature on display monitoring performance: An overview with practical implications. *American Industrial Hygiene Association Journal, 45,* 122–126.

Hancock, P. A. (1986a). The effect of skill on performance under an environmental stressor. *Aviation, Space, and Environmental Medicine, 57,* 59–64.

Hancock, P. A. (1986b). Sustained attention under thermal stress. *Psychological Bulletin, 99,* 263–281.

Hancock, P. A. (2003). The ergonomics of torture: The moral dimension of evolving human-machine technology. In *Proceedings of the Human Factors and Ergonomic Society 47th Annual Meeting* (pp. 1009–1011). Santa Monica, CA: Human Factors and Ergonomics Society.

Hancock, P. A. (2006). Thermal comfort. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (2nd ed., pp. 1859–1862), London: Taylor & Francis.

Hancock, P. A., & Desmond, P. A. (Eds.). (2001). *Stress, workload and fatigue.* Mahwah, NJ: Erlbaum.

Hancock, P. A., & Ganey, H. C. N. (2003). From the inverted-U to the extended-U: The evolution of a law of psychology. *Human Performance in Extreme Environments, 7,* 5–14.

Hancock, P. A., & Pierce, J. O. (1985). Combined effects of heat and noise on human performance: A review. *American Industrial Hygiene Association Journal, 46,* 555–566.

Hancock, P. A., & Vasmatzidis, I. (1998). Human occupational and performance limits under stress: The thermal environment as a prototypical example. *Ergonomics, 41,* 1169–1191.

Hancock, P. A., & Vasmatzidis, I. (2003). Effects of heat stress on cognitive performance: The current state of knowledge. *International Journal of Hyperthermia, 19,* 355–372.

Hancock, P. A., & Vercruyssen, M. (1988). Limits of behavioral efficiency for workers in heat stress. *International Journal of Industrial Ergonomics, 3,* 149–158.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31,* 519–537.

Harris, W. C., Hancock, P. A., & Harris, S. C. (2005). Information processing changes following extended stress. *Military Psychology, 17,* 115–128.

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychological Review, 62,* 243–254.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

Hedges, L. V., Shymansky, J. A., & Woodworth, G. (1989). *A practical guide to modern methods of meta-analysis.* Washington, DC: National Science Teachers Association.

Hockey, G. R. J., Gaillard, A. W. K., & Coles, M. G. H. (Eds.). (1986). *Energetics and human information processing.* Dordrecht, Netherlands: Martinus Nijhoff.

*Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C., & Roberts, W. (2001). Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology, 128*(A), 719–734.

Hoff, H. H., Guillemin, L., & Guillemin, R. (1967). *The cahier rouge of Claude Bernard.* Cambridge, MA: Schenkman.

*Holmberg, I., & Wyon, D. (1969). The dependence of performance in school on classroom temperature. *Educational and Psychological Interactions, 31,* 1–20.

Houghton, F. C., & Yagloglou, C. P. (1923). Determining equal comfort lines. *Journal of the American Society Heating and Ventilation Engineers, 29,* 165–176.

Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80,* 327–334.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting for error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

Huntington, E. (1919). *World-power and evolution.* New Haven, CT: Yale University Press.

*Hygge, S., & Knez, I. (2001). Effects of noise, heat and indoor lighting on cognitive performance and self-reported affect. *Journal of Environmental Psychology, 21,* 291–299.

Johnson, R. F., & Kobrick, J. L. (2001). Psychological aspects of military performance in hot environments. In K. B. Pandolf, R. E. Burr, C. B. Wenger, & R. S. Pozos (Eds.), *Medical aspects of harsh environments* (pp. 135–159). Washington, DC: Borden Institute.

Kleitman, N. (1963). *Sleep and wakefulness.* Chicago: Chicago University Press. (Original work published 1939)

Lachman, R., Lachman, J. L., & Butterfield, E. C. (1979). *Cognitive psychology and information processing.* Hillsdale, NJ: Erlbaum.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Lind, A. R. (1963). A physiological criterion for setting thermal environmental limits on everyday work. *Journal of Applied Physiology, 18,* 51–56.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

*Lockhart, J. M. (1968). Extreme body cooling and psychomotor performance. *Ergonomics, 11,* 249–260.

*Lockhart, J. M., & Kiess, H. O. (1971). Auxiliary heating of the hands during cold exposure and manual performance. *Human Factors, 13,* 457–465.

Loeb, M., & Jeantheau, G. (1958). The influence of noxious environmental stimuli on vigilance. *Journal of Applied Psychology, 42,* 47–49.

*Mackworth, N. H. (1946). Effects of heat on wireless telegraphy operators hearing and recording Morse messages. *British Journal of Industrial Medicine, 3,* 143–158.

Martinussen, M., & Bjornstad, J. F. (1999). Meta-analysis calculations based on independent and non-independent cases. *Educational and Psychological Measurement, 59,* 928–950.

*Mayo, G. D. (1955). Effect of temperature upon technical training. *Journal of Applied Psychology, 39,* 244–246.

*Moreland, S., & Barnes, J. A. (1970). *Exploratory study of pilot performance during high ambient temperatures/humidity* (Aberdeen Research and Development Tech. Rep. 5184.12.646). Aberdeen Proving Ground, MD: U.S. Army.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimate in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125.

Näätänen, R. (1973). The inverted-U relationship between activation and performance. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 155–174). New York: Academic Press.

Parsons, K. C. (1993). *Human thermal environments: The effects of hot, moderate and cold environments on human health, comfort and performance.* London: Taylor & Francis.

Parsons, K. C. (1995). International heat stress standards: A review. *Ergonomics, 38,* 6–22.

Patterson, M. J., Taylor, N. A. S., & Amos, D. (1998). *Physical work and cognitive function during acute heat exposure before and after heat acclimization* (DSTO-TR-0683). Melbourne, Australia: Defence Science and Technology Organisation, Aeronautical and Maritime Research Laboratory.

*Peccolo, C. (1962). *The effect of thermal environment on learning: A pilot study.* Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.

*Pepler, R. D. (1960). Warmth, glare and a background of quiet speech: A comparison of their effects on performance. *Ergonomics, 3,* 68–73.

*Pepler, R. D., & Warner, R. E. (1968). Temperature and learning: An experimental study. In *Proceedings of the 74th ASHRAE Conference* (pp. 211–219). Atlanta, GA: American Society of Heating Refrigerating and Air Conditioning Engineers.

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

Poulton, E. C. (1965). On increasing the sensitivity of measures of performance. *Ergonomics, 8,* 69–76.

Poulton, E. C. (1970). *Environment and human efficiency.* Springfield, IL: Charles C. Thomas.

Poulton, E. C. (1976). Arousing environmental stresses can improve performance, whatever people say. *Aviation, Space, and Environmental Medicine, 47,* 1193–1204.

*Poulton, E. C., & Kerslake, D. M. (1965). Initial stimulating effect of warmth upon perceptual efficiency. *Aerospace Medicine, 36,* 29–32.

Prosser, C. L., & Nelson, D. O. (1981). The role of nervous systems in temperature adaptation of poikilotherms. *Annual Review of Physiology, 43,* 281–300.

Ramsey, J. D. (1995). Task performance in heat: A review. *Ergonomics, 38,* 154–165.

Ramsey, J. D., & Morrissey, S. J. (1978). Isodecrement curves for task performance in hot environments. *Applied Ergonomics, 9,* 66–72.

*Razmjou, S. (1996). Mental workload in heat: Toward a framework for analyses of stress states. *Aviation, Space, and Environmental Medicine, 67,* 530–538.

*Razmjou, S., & Kjellberg, A. (1992). Sustained attention and serial responding in heat: Mental effort in the control of performance. *Aviation, Space, and Environmental Medicine, 63,* 594–601.

*Reading, J. E., Kincaid, P. S., Roberts, D. E., Hesslink, R. L., & Pozos, R. S. (1994). *The effect of shivering on rifle shooting performance in U.S. Marines* (Tech. Rep. 94-5). San Diego, CA: U.S. Naval Health Research Center.

*Reardon, M. J., Fraser, E. B., & Omer, J. M. (1998). Flight performance effects of thermal stress and two aviator uniforms in UH-60 helicopter simulators. *Aviation, Space, and Environmental Medicine, 69,* 569–576.

*Reddy, S. P. (1974). *Sedentary job performance within the thermal comfort zone.* Unpublished master's thesis, Texas Tech University, Lubbock, TX.

*Reilly, R. E., & Parker, J. F. (1988). *Effect of heat stress and prolonged activity on perceptual-motor performance* (Tech. Rep. CR-1153). Arlington, VA: National Aeronautics and Space Administration.

Rhudy, J. L., & Meagher, M. W. (2003). Negative affect: Effects on an evaluative measure of human pain. *Pain, 104,* 617–626.

Selye, H. (1976). *The stress of life.* New York: McGraw-Hill.

*Sharma, V. M., & Panwar, M. R. (1987). Variations in mental performance under moderate cold stress. *International Journal of Biometeorology, 31,* 85–91.

Southern Region Headquarters National Weather Service. (n.d.). *Relative humidity and dewpoint temperature from temperature and wet-bulb temperature.* Retrieved April 6, 2007, from http://www.srh.noaa.gov/elp/wxcalc/formulas/rhTdFromWetBulb.html

*Tanaka, M., Tochihara, Y., Yamazaki, S., Ohnaka, T., & Yoshida, K. (1983). Thermal reaction and manual performance during cold exposure while wearing cold-protective clothing. *Ergonomics, 26,* 141–149.

Taylor, C. L. (1948). *Interim report, Committee on Aviation Medicine.* Washington, DC: National Research Council.

*Teichner, W. H., & Wehrkamp, R. F. (1954). Visual-motor performance as a function of short-duration ambient temperature. *Journal of Experimental Psychology, 47,* 447–450.

*Thomas, J. R., Ahlers, S. T., House, J. F., & Schrot, J. (1989). Repeated exposure to moderate cold impairs matching-to-sample performance. *Aviation, Space, and Environmental Medicine, 60,* 1063–1067.

*Tikuisis, P., Keefe, A. A., Keillor, J., Grant, S., & Johnson, R. F. (2002). Investigation of rifle marksmanship on simulated targets during thermal discomfort. *Aviation, Space, and Environmental Medicine, 73,* 1176–1183.

Ursin, H., & Eriksen, H. R. (2004). The cognitive activation theory of stress. *Psychoneuroendocrinology, 29,* 567–592.

*van Orden, K. F., Benoit, S. L., & Osga, G. A. (1996). Effects of cold air stress on the performance of a command and control task. *Human Factors, 38,* 130–141.

Vasmatzidis, I., Schlegel, R. E., & Hancock, P. A. (2002). Dual task performance under heat stress. *Ergonomics, 45,* 218–239.

Vernon, H. M., & Bedford, T. (1930). *A study of heating and ventilation in schools* (Industrial Health Research Board, Medical Research Council Report). London: His Majesty's Stationery Office.

Von Humboldt, A. (1816). *Voyage aux regions equinoxiales* [Voyages to Equatorial Regions]. Paris: Greek, Latin, and German Library.

*Weiner, J. S., & Hutchinson, J. C. D. (1945). Hot humid environment: Its effect on the performance of a motor co-ordination test. *British Journal of Industrial Medicine, 2,* 154–157.

Wiener, N. (1954). *The human use of human beings: Cybernetics and society.* Boston: Houghton Mifflin.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance.* Upper Saddle River, NJ: Prentice Hall.

Wyndham, C. H. (1969). Adaptation to heat and cold. *Environmental Research, 2,* 442–469.

Wyndham, C. H., Strydom, N. B., Munro, A., MacPherson, R. K., Metz, B., Schaff, G., et al. (1964). Heat reactions of Caucasians in temperate, in hot, dry, and in hot, humid climates. *Journal of Applied Physiology, 19,* 607–612.

*Wyon, D. P. (1969). *The effects of moderate heat stress on the mental performance of children* (Tech. Rep. D8/69). Stockholm: National Swedish Institute for Building Research.

Yaglou, C. P., & Minard, D. (1957). Control of heat casualties at military training centers. *American Medical Association Archives of Industrial Health, 16,* 302–316 and 405 (corrections).

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459–482.

P. A. Hancock is the Provost Distinguished Research Professor at the University of Central Florida, Orlando, FL. He received his Ph.D. in motor performance from the University of Illinois at Urbana-Champaign in 1983.

Jennifer M. Ross is a fifth-year graduate student in the Applied Experimental Human Factors Ph.D. Program at the University of Central Florida, where she earned her M.A. in modeling and simulation in 2004.

James L. Szalma is an assistant professor in the Department of Psychology at the University of Central Florida. He received his Ph.D. in applied experimental psychology/human factors from the University of Cincinnati in 1999.

Date received: February 9, 2006
Date accepted: January 10, 2007