

ORIGINAL RESEARCH REPORT

The Human Penguin Project: Climate, Social Integration, and Core Body Temperature

Hans IJzerman*, Siegwart Lindenbergt†, İlker Dalğar§, Sophia S. C. Weissgerber||, Rodrigo C. Vergara¶, Athena H. Cairo**, Marija V. Čolić††, Pinar Dursun‡‡, Natalia Frankowska§§, Rhonda Hadil|||, Calvin J. Hall**, Youngki Hong¶¶, Chuan-Peng Hu***, Jennifer Joy-Gaba**, Dušanka Lazarević††, Ljiljana B. Lazarević††, Michal Parzuchowski§§, Kyle G. Ratner¶¶, David Rothman**, Samantha Sim†††, Cláudia Simão†††, Mengdi Song***, Darko Stojilović††, Johanna K. Blomster§§§, Rodrigo Brito|||, Marie Hennecke¶¶¶, Francisco Jaume-Guazzini††††, Thomas W. Schubert†††.†††, Astrid Schütz§§§§, Beate Seibt§§§ and Janis H. Zickfeld§§§

Social thermoregulation theory posits that modern human relationships are pleiomorphically organized around body temperature regulation. In two studies ($N = 1755$) designed to test the principles from this theory, we used supervised machine learning to identify social and non-social factors that relate to core body temperature. This data-driven analysis found that complex social integration (CSI), defined as the number of high-contact roles one engages in, is a critical predictor of core body temperature. We further used a cross-validation approach to show that colder climates relate to higher levels of CSI, which in turn relates to higher CBT (when climates get colder). These results suggest that despite modern affordances for regulating body temperature, people still rely on social warmth to buffer their bodies against the cold.

Keywords: Social Integration; Social Thermoregulation Theory; Attachment Theory; Embodiment; Machine Learning

One key motivating force for bonding across mammals is their need to regulate body temperature (Ebensperger, 2001). Without adequate temperature regulation, they die. Distributing body heat across conspecifics makes responding to environmental fluctuations in temperature less costly energetically by lowering metabolic rate. Perhaps more surprisingly, the literature on human adults has also shown a link between feelings of trust (or psychological warmth) and temperature regulation (for an overview, see IJzerman & Hogezeil, 2018). It is unclear however to what extent these effects extend to relationships beyond

immediate close ties and whether “social warmth” indeed protects people’s core body temperatures from the cold.

In this report, we investigated whether the quality of one’s social networks relates to higher core body temperatures, even when environmental temperatures are lower. We did so through a high-powered pilot study and an even higher-powered cross-national study. In the pilot study, we first identified *which* variables best predict core body temperature by using a powerful exploratory method we borrowed from artificial intelligence called supervised machine learning. In our cross-national study,

* Université Grenoble Alpes, FR

† Rijksuniversiteit Groningen, NL

‡ Tilburg University, NL

§ Middle East Technical University, TR

|| Universität Kassel, DE

¶ Universidad de Chile, CL

** Virginia Commonwealth University, US

†† University of Belgrade, RS

‡‡ Afyon Kocatepe University, TR

§§ SWPS University of Social Sciences and Humanities, PL

||| University of Oxford, UK

¶¶ University of California, Santa Barbara, US

*** Tsinghua University, CN

††† Singapore Management University, SG

‡‡‡ ISPA Instituto Universitário, Lisbon, PT

§§§ University of Oslo, NO

||| Lusófona University, PT

¶¶¶ University of Zürich, CH

†††† Pontificia Universidad Católica de Chile, CL

†††† Universidad del Desarrollo, CL

†††† Instituto Universitário de Lisboa, ISCTE-IUL, PT

§§§§ Otto-Friedrichs-Universität Bamberg, DE

Corresponding author: Hans IJzerman

(h.ijzerman@gmail.com)

we again used the same exploratory method of supervised machine learning. We then conducted a split-half cross-validation (“training” a mediation in one half of the data, which was then tested in the second half of the data) of a path model to assess *how* the earlier identified variables relate to core body temperature. Our machine learning results and path model both provide support for a strong relationship between people’s environmental temperature (operationalized as distance from the equator), their levels of social integration across different relationships, and their core body temperatures.

Social Thermoregulation as Key Facet of Human Social Attachments

Having high-quality social relationships is one of the biggest predictors of one’s health (Holt-Lunstad et al., 2010). Although scholars dating back to Hippocrates have understood that disturbances in health closely relate to dysregulated body temperature (Benzinger, 1969; Minard et al., 1964), evidence for the link between thermophysiology and social relationships in humans has only just begun to accumulate (for a review, see IJzerman & Hogerzeil, 2018). In its most elementary form endotherms (i.e., animals that generate their own heat) maintain temperature at homeostatic levels in various ways, such as yawning, panting, or sweating when temperatures increase or shivering when temperatures decrease (Gallup & Gallup, 2008; Jansky, 1973).

Organisms also turn to others to help with temperature regulation. Suggestive evidence for these ideas can be found in studies across nonhuman endothermic animals. Amongst rodents, social thermoregulation is likely one of the most important motivating forces behind group living, especially when temperatures drop (Ebensperger, 2001). Experimental research has shown that the *Octodon Degus* (a Chilean rodent) uses 40% less energy and achieves a higher surface temperature when housed with three or five others vs. alone (Nuñez-Villegas et al., 2014). Studies of vervet monkeys display somewhat more complex mechanisms: Larger social networks do buffer core temperature from the cold and even grooming a dead vervet monkey’s pelt insulates against temperature variations (McFarland, Henzi et al., 2015; McFarland, Fuller et al., 2015). The reliance on conspecifics seems remarkably asymmetrical: Coping with elevated temperatures is typically accomplished by the organism itself (e.g., through internal regulation like yawning or through behavioral thermoregulation like getting into colder water) because overheating can be immediately threatening for survival. In contrast, because temperature decreases are not immediately dangerous, regulation back to homeostasis is often “outsourced” to conspecifics through huddling.

In humans, *social thermoregulation* extends beyond huddling. More specifically, social thermoregulation theory explicates how what English speakers intuitively know as “social warmth”, that is, trustworthiness and social predictability, relies on and grows out of more ancient needs for physical warmth. This idea is supported by recent findings by Vergara et al. (2017), who find that

attachment avoidance is negatively correlated (at $r = -.32$; $N = 1504$), with habits related to social thermoregulation (e.g., “I prefer to warm up with someone rather than with something”). This repurposing has likely happened to avoid redundancy: Earlier in evolutionary history regulating temperature was crucial for survival, so it was efficient to “reuse” similar brain regions when social thermoregulation evolved. Evidence for overlap between neural areas involved in social thermoregulation and social interaction supports this possibility (Anderson, 2010; Satinoff, 1982, 1983). The connection between social behavior and thermoregulation is likely not incidental: Like other homeotherms, ancient *Homo Sapiens* simply needed to stay physically proximate to stay warm.¹

From that perspective, it may not come as a surprise that the aggregate evidence has come to favor this evolved relation between social and physical warmth (IJzerman & Hogerzeil, 2018; IJzerman, Janssen, et al., 2015; Schilder et al., 2009). Evolutionary pressures related to infant survival (e.g., feeling cold and wanting to be held) likely form the basis for an evolved template for mental (attachment-like) models concerning the relationship between physical and social warmth. Furthermore, the priming of (lack of) trust is as asymmetrical as the underlying physiological systems: Priming trust leads to higher temperature perceptions when temperatures are low, but not when temperatures are high (Ebersole et al. 2016; IJzerman et al., 2016).

This relationship can also be observed in its most basic form in the “Strange Situation,” in which researchers observe an infant’s behavior in response to separation of the mother. When the mother leaves the room in this ethological observation, infants’ skin temperature drops, and peripheral temperature only returns to baseline once the mother returns (Mizukami et al., 1990). Similar effects can be observed in adults: Students’ peripheral temperatures drop when they feel socially excluded (IJzerman et al., 2012). Nevertheless, doubts can exist in regard to what extent social thermoregulation is still important for modern day humans. After all, we regulate our temperatures in many ways (clothes, heaters) that do not involve other people. However, the availability of modern conveniences to regulate temperature has been so brief that the evolved link between physical and social warmth is likely to *still* lead to strategies that help buffer individuals from the cold through their social networks – even via relationships that do not typically permit physical touch.

From prior (pre-registered) research, we learnt that feeling cold increases the need to socially connect (Van Acker et al., 2016). There is only one pilot study demonstrating a relationship between network quality and higher core body temperatures in humans, showing that having greater feelings of social connection is positively correlated to core body temperature (Inagaki et al., 2016). But on the basis of the existing literature, it is not at all clear *whether* social connections protect against the cold, *which* aspects of social contact protect against the cold, and *whether* social contacts are more prominent in predicting core body temperature than other known variables. We therefore first explore which variables are

crucial in predicting core body temperatures. Although we used a data-driven analysis, principles from existing theories were used to determine the variables to include in our study.

Analytical Approach – Supervised Machine Learning and Split-Half Cross-Validation

To achieve our goals, we first used a data-driven machine learning approach. This method has not been widely used in social psychological research, which has typically focused on testing theoretically derived hypotheses. Several researchers have now argued that psychological science's focus on identifying complex prediction-focused models through hypothesis testing, together with flexibility in data analyses have led to what is called "overfitting" (mistaking noise for a real signal by fitting an overly complex model to existing data).

This process of overfitting is an important cause of the reproducibility crisis (for overviews, see IJzerman, Pollet, Ebersole, & Kun, 2016; Yarkoni & Westfall, 2017). Recently, multiple researchers have called for trying to rely on more accurate predictive models before moving on to mechanistic explanations of human behavior (IJzerman et al., 2017; Yarkoni & Westfall, 2017). Indeed, valid scientific discoveries require separating exploratory and confirmatory research (Wagenmakers, Wetzels, Borsboom, Van der Maas, & Kievit, 2012). We combine an exploratory approach with a partially confirmatory approach in our analyses by first using a powerful exploratory analysis method called *supervised machine learning*, after which we move on to *split-half cross-validation*.

Supervised machine learning is an approach that generates solutions from data, and relies on flexibility in data analysis to specify predictors (predictors, in this case, does not imply causality). It is a powerful way to explore data, which helps formulate predictions for out-of-sample testing in a relatively robust way and which does not presume a specific relationship (e.g., positive or negative, linear or nonlinear). A popular and widely tested approach to supervised machine learning is a method called "random forest" (Breiman, 2001). This method allows for measuring the relative contribution of each specific variable, considering all the variables used in a dataset to predict the outcome of a "signal" (i.e., a dependent variable) through a bootstrapping-type method, yielding a highly predictive accuracy and great deal of precision over the specification of control variables.

In unsupervised machine learning, the algorithm infers a hidden function or pattern from the data, without regard to such a "signal" (which we typically refer to as the dependent variable). Supervised machine learning on the other hand differs from unsupervised machine learning in that the data patterns are derived by a "supervisory signal" (an outcome variable). As the name implies, the "forest" consists of many such "trees". The method relies on "out of bag estimates" (bagging), which involves repeated sampling to form training datasets from an original dataset (Breiman, 1996; Bylander & Hanzlik, 1999).

The rest of the datasets in each case (the test datasets) are used to evaluate the predictive power of the variable

importance and trees trained on the training dataset. The forest is then aggregated with each tree getting a "vote", which constitutes a weight in the ensembled model that summarizes all information from the trees. The outcome of this repeated sampling is captured in a variable importance list that indicates which variables are very likely to predict the outcome variable (for more technical discussions see Breiman, 2001; IJzerman, Pollet, et al., 2016; Jones & Lindner, 2015; Yarkoni & Westfall, 2017). Because the analyses are exploratory, it does not provide an effect size estimate but instead shows the relative importance of one variable over the other within the model, and from the variables that predict compared to a model built of random noise.

The type of supervised machine learning we used here 1) allows for non-linearity (which standard regressions cannot do without a priori specification), 2) does not presume direction (a positive or negative relationship), 3) has much less problems with collinearity, and 4) is agnostic on what type of variable predicts the outcome, thus allowing the researcher to *classify before regressing* onto the signal (i.e., what psychologists typically refer to as the dependent variable).

The type of machine learning we used (*conditional random forest*) improves its predictive power throughout each iteration of the analyses and has a higher explained variance than a regular random forest. This method has the potential to reduce bias in analyses and is particularly useful when multiple predictor variables are measured. Overall, it is useful as an exploratory approach to identify variables for further confirmatory testing. For us that meant that we were able to pinpoint which variables were most relevant in predicting core body temperature in Phase I of Study 1 and 2.

Chances for overfitting are much lower than in the case of fitting hypotheses to data a posteriori because of the repeated sampling of the variables. Furthermore, chances for overfitting are also lower for the type of supervised machine learning we utilize, as conditional random forests reduce the error throughout each model (see e.g., Strobl, Malley, & Tutz, 2009). Overfitting is still possible, however, and can be further reduced by 1) replicating the study (which we do from the pilot to main study), 2) replicating the analyses through different seeds (which we do, one time in our pilot and three times in our main study), and 3) by supplementing them with different analyses that converge with the supervised machine learning analyses (which we do through split-half cross-validation).

Thus, we combine our approach with a partially confirmatory approach in Phase II, where we split the data in two, exploring a path model (based on variables identified through supervised machine learning) in the first half of the data, which we then confirm with the second half of the data (a method called "split-half cross-validation"). For these analyses – based on Phase I and our theoretical reasoning – we specified a path model (an approach that may be better known to readers). Altogether, this allowed us to first explore the data in Phase I (through supervised machine learning), and to specify a mechanism

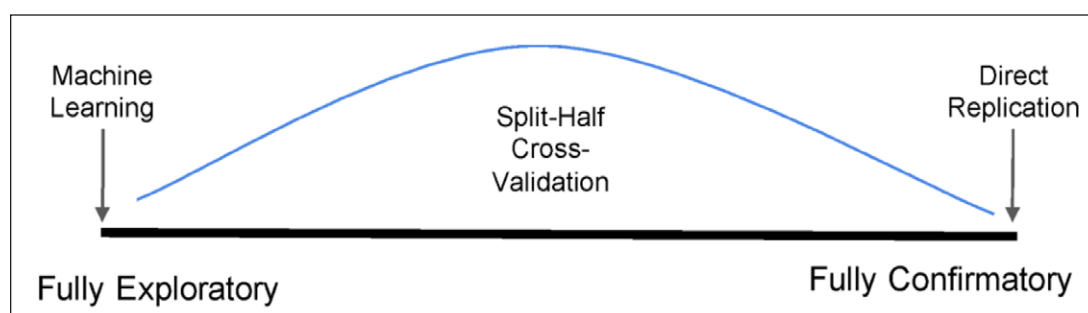


Figure 1: Depiction of Certainty of Exploratory versus Confirmatory Approaches. A depiction of the research process with greater uncertainty to the left and greater certainty to the right. Fully unsupervised machine learning (where no relationship between variables is clear) should be placed entirely on the left. Supervised machine learning, in which the “dependent variable” (or signal) is clear and certain predictor variables are likely well specified (such as in our case) fall somewhat to the right of that. Split-half cross-validation (where datasets are split in two and explored and confirmed) fall yet further to the right. The “most confirmatory approach” is a “close” or direct replication.

in Phase II. Epistemologically, relying on exploratory approaches followed by split-half cross-validation pushes the certainty of our findings toward near-confirmatory status (but not as certain as, for example, a replication. See also **Figure 1** for a depiction of the continuity between exploratory and confirmatory research approaches).

To summarize: We were relatively certain about the types of variables that should be included (those related to the quality of the social network and core body temperature), but were less certain about the exact variable that would predict core body temperature and about the exact variables we should specify as controls. Our supervised machine learning allowed us to discover these features from the data, which we then tested through more traditional methods.

The Human Penguin Project Overview

We accomplished our goals in two studies. We first ran an online pilot study ($N = 232$) and then ran a large, cross-national study (12 countries; $N = 1523$) to identify which variables are most accurate in predicting people's core body temperature. We measured a number of known correlates of core body temperature using a questionnaire and a number of social relationship variables that – based on prior research – should logically be related to core body temperature (e.g., nostalgia (Zhou et al., 2012) or attachment to homes (Van Acker et al., 2016)). We included a number of variables that have been found to relate to body temperature (stress; Marazziti et al., 1992; whether participants use medication) or those that have been known to relate to environmental temperature variations (nostalgia; Zhou et al., 2012; attachment to homes; Van Acker et al., 2016) or to metabolism and social network quality (like daily diet/sugary drinks consumption; Henriksen et al., 2014).

In selecting our variables, we were over- rather than under-inclusive, as our first priority was to identify *which* variables were the most prominent predictors of core body temperature. We also asked questions that relate to the regulation of stress (and could thus relate to body temperature) like self-control (Tangney et al., 2004), attachment (Fraley et al., 2000), and access to one's own feelings and bodily states (alexithymia; Kooiman

et al., 2002). In our second (main) study, we again first relied on supervised machine learning, after which we specified our path model through a split-half cross-validation method.

Method Pilot Study

Participants and Procedure. Our questionnaire was programmed into the online platform Qualtrics and we collected data from mTurk ($N = 143$) and Prolific Academic ($N = 148$). Participants (all from US and UK; $M_{age} = 49.45$ ($SD_{age} = 7.01$; 37.9% male, 61.3% female) were asked to complete the survey between 9–11am, not eat or drink anything warm or cold for 10 minutes preceding the survey, and not exercise an hour preceding the survey. In our pilot study (where we could not easily control for other variables), we excluded all participants who did not adhere to these guidelines (mTurk $N = 3$; PA $N = 56$). When we used mTurk and Prolific Academic, mTurk had parameters that could be set to control the quality of participants (so called “qualification requirements”) which Prolific Academic did not have at the time that we ran our study. We had many more participants at Prolific Academic than mTurk that did not follow our instructions, which can likely be explained by the qualification requirements we could set a priori for mTurk. The total remaining N for the pilot study was 232 ($M_{age} = 49.5$, $SD = 7.04$; 37.5% male, 62.5% female).

Core body temperature was measured with an oral thermometer at the beginning of the questionnaire (Measurement 1) and at the end of the questionnaire (Measurement 2) by participants. To authenticate the temperature reading, participants took a picture of the thermometer (with date, time, and Measurement – 1 or 2 – included; for an example photo uploaded by our participants, see **Figure 2**).²

Survey details. Our dataset included a number of scales relevant to thermoregulation. In order to assess the importance of the quality of social networks (in relation to climate), we measured known correlates of core body temperature (from here on CBT) or behavior in response to temperature fluctuations, like self-reported stress (“stress” in the forest plot; Cohen & Wills, 1985), nostalgia (“nostalgia”; Routledge et al., 2008), attachment

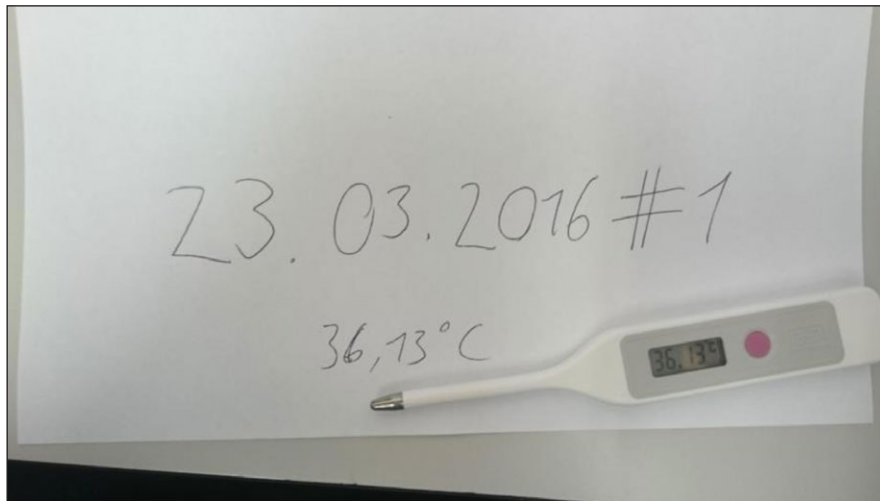


Figure 2: HPP Thermometer. An example picture of a thermometer uploaded to Qualtrics.

to homes (“attachhome”; Harris et al., 1996), daily sugary drinks consumption (“gluctot”; Henriksen et al., 2014) and diet drinks consumption (“artgluctot”; Henriksen et al., 2014), known benchmarks of CBT, like sex (“sex”), height (“height”), weight (“weightkg”), and whether they used medication (“meds”; Bergmann, 1847; Hill & Rahimtulla, 1965; Peters, 1986).

We also included variables that potentially influence CBT and quality of the social network in other ways, like feelings of agency, measured through self-control (“selfcontrol”; Tangney et al., 2004), attachment (“avoidance” and “anxiety”; Fraley et al., 2000), and access to one’s own feelings and bodily states (alexithymia subscales “EOT” and “DIDF”; Kooiman et al., 2002). In our pilot study, our exploratory analyses reliably showed that the quality of social networks was an important predictor of CBT. The quality of social networks was measured through a measure of people’s social networks (“networksize”, “socialembedded”, and a measure of Complex Social Integration (from here on CSI); Cohen et al., 1997).

We found that one of the most important predictors of CBT was CSI. CSI is an inventory of the frequency of contact that people have with various important people in their lives. More specifically, it includes an inventory of the following ties: Relationships with spouse, parents, parents-in-law, children, other close family members, neighbors, friends, workmates, schoolmates, fellow volunteers (e.g., charity or community work), members of groups without religious affiliations (e.g., social, recreational, professional), and members of religious groups. One point was assigned for participation in each kind of relationship for which respondents reported that they spoke (in person or on the phone) to someone in that relationship at least once every 2 weeks. At the end of the survey, participants were thanked for their participation and debriefed.

The complete scales, reliabilities, and averages per scale per site can be accessed on our project page (<https://osf.io/2w46c/>). Finally, we looked up the minimum temperature (“mintemp”) and average humidity on the day (“avghumidity”) participants completed the survey based on their IP address by using a weather history site

(<http://www.wunderground.com/history/>), which bases weather on the nearest airport.

Analyses and Results Pilot Study

Our method, conditional random forest, consists of an algorithm that classifies data points by weighting “votes” of predictions of a potential hypothesis underlying the relationship between variables. It then creates “decision trees” that indicate which variables get the most weight in relevance for the “signal” (i.e., a dependent variable). The order in which these decisions are taken are represented by the “levels” of the tree. The path from the root of the tree to a node is a series of decisions and the node is then tagged by the prediction power of such a path. Given enough data (and enough predictors) these decision trees are very flexible, as the algorithm explores all possible relationships between predictor variables and signal that could be generated from the data. Each of these decision trees then allow for specification of the strength of predictor of variables based on a “desired outcome value” (the “supervisory signal”, which is basically the dependent variable). Conditional random forests immediately engage in error-correction during the process, which means that, as a result, no training vs. test dataset is needed as happens in typical random forests. The forest is then aggregated with each tree getting a “vote”, which constitutes a weight in the ensembled model that summarizes all information from the trees. This then creates a *permutation variable importance* (i.e., the list with relative importance of each variable in predicting the outcome variable).

For creating the classification trees, we relied on the R packages: tree (Ripley, 2016), lattice (Sarkar, 2017), plyr (Wickham, 2016), stargazer (Hlavac, 2015), and summarytools (Comtois, 2016). MTry is the numbers of variables (out of the total list of variables) sampled at each split. MTry is recommended to be the square root of the total number of predictors. For our pilot study, we ran the analyses four times, with two different versions of mtry and a replication of each seed (with mtry = 4 and 5, trees = 1000, for seeds 1 (original) and 2 (replication); link to script: <https://osf.io/ahks6/> and to data: <https://osf.io/x386r/>). The chance for overfitting is further reduced

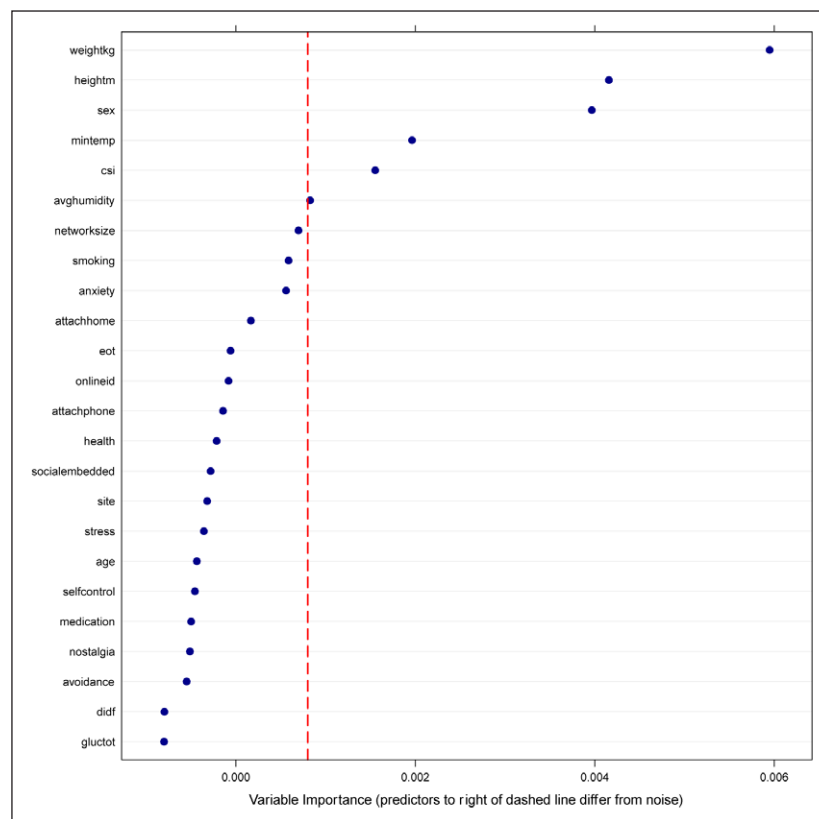


Figure 3: HPP Pilot Dotplot. Permutation variable importance of predictors of Core Body Temperature from our supervised machine learning analyses in our pilot study. Variables exceeding the red line are very unlikely random noise.

by examining the stability of the two analyses through a Spearman Rank correlation between the forests. In this case, the model was very stable, with Spearman Rank r ranging between .82–.93 (for more details about the procedure, see IJzerman, Pollet et al., 2016; for link to variable orders, see <https://osf.io/z3v7h/> and for output, see <https://osf.io/rrhhh/files/>; see **Figure 3** for one of the four dotplots).

The outcome of our conditional random forest analyses in our pilot study was that core body temperature is best predicted by (in order of importance):

participants' weight (weightkg) > participants' height (heightm) > sex > minimum temperature of that day (mintemp) > complex social integration (CSI; see **Figure 3**).

These variables exceed the “random noise threshold” in our forest, suggesting that these variables (and not others) differ from random noise in our dataset (as indicated by them not exceeding the red line that defines what differs from random noise in a dataset; Strobl et al., 2009).

Discussion Pilot Study

Our first study showed that the best predictors of CBT in our online samples of Prolific Academic and mTurk were height, CSI, weight, minimum temperature, and sex. Beyond known benchmarks like height, weight, and sex, we discovered that CSI was one of the most important predictors of CBT (with CSI positively relating to CBT). Why could this be so? It is likely that there are mechanisms in place to achieve higher CBT through diverse social contacts.

Recall that we already know that feeling cold increases the need to socially connect (even via email and phone; Van Acker et al., 2016) and that one pilot study has shown a relationship between feeling more connected socially and higher core temperatures in humans (Inagaki et al., 2016). Furthermore, we know that people project their relationships even onto inanimate objects like consumer products (Hadi et al., 2012; IJzerman, Janssen, et al., 2015; Rotman et al., 2016). It is also known that neural areas related to thermoregulation overlap considerably with those related to social behaviors (Satinoff, 1982, 1983).

Given these diverse findings that point in one direction, the most important reason for the relation between CSI and CBT may be that a low variety of relationships is a risk factor for poor life outcomes. In our complex social world, rejection and loneliness are common experiences because of the tendency to compare ourselves to others and because of the fragility of many relationships in life – from work to recreation to home to friendships. Given the unknown stability of any particular domain of life, past research has found that having a wider-range of strong social ties – or, higher levels of CSI – is particularly important for health and well-being (Cohen et al., 1997; Seeman, 1996). Thus, staking too much in any particular type of relationship is risky and potentially isolating if conflict arises or we do not feel we are meeting the standards in a particular life domain (Crocker et al., 2003; Deci & Ryan, 1995; Kernis, 2003; Steverink & Lindenberg, 2008).

This is likely the reason why some of the strongest evidence for the buffering effects of a web of social ties (versus putting too much weight in the strength of particular social tie) comes from studies that have

assessed levels of CSI. This measure asks questions related to whether people are in regular contact with people in multiple facets of their lives (parents, relatives, close friends, colleagues, and so forth). The fact that having higher levels of CSI is also the most established buffer against loneliness (Holt-Lunstad et al., 2010), and that previous work suggests that loneliness is related to (peripheral) body temperature (IJzerman et al., 2012), is an additional reason why CSI might be positively associated with core body temperature.

The Human Penguin Main (Cross-National) Study

Research Summary

In our pilot study, we found that CSI was one of the most potent predictors of CBT. In our second, larger and cross-national project, we investigated whether CSI indeed protects against the cold. For the main study, we crowdsourced our data collection, as having samples from different locations around the world was crucial for our research hypothesis. We suspected that people who live in colder climates (i.e., further from the equator) rely more on more varied social contacts to keep warm, and as a result, climate would moderate an association between CSI and CBT. We again relied on supervised machine learning to identify predictors for CSI and again for CBT, using the same exploratory method before proceeding with a split-half cross-validation analysis that could help us specify a path model. We expected that CSI would again turn up as an important predictor of CBT, and we explored whether CSI relates to higher CBT especially when climates are colder.

Participants

In our cross-national study, we tested the interrelationship between climate, CSI, and CBT on a fairly large scale, including 12 different countries on 3 different continents, and 1,507 participants. We report all our exclusions in our data handling section. We also report all of the variables we measured in our study, except for two questionnaires intended for the development and validation of scales separate from this study.³ The first author recruited collection sites through personal contacts and through “the ManyLab” (<https://osf.io/89vqh/>). Labs received a description about the study, information about what was required for participating labs (150 participants at a minimum, with 200 as ideal), what they would get in exchange for providing data and translations, and how they could join the project. Because data collection was more difficult than anticipated, minimum participant number was relaxed to a 100-participant minimum. Co-authors from labs that did not achieve the target sample were dropped to a second-tier authorship (=later in the authorship list). Participants again completed a variety of online questionnaires at home or in the lab (depending on site). Answering the questionnaire took approximately 35 minutes in total.

Method Cross-National Study

Samples. We collected data via University of Oxford (UK; $N = 137$, 56.2% female, $M_{\text{birthyear}} = 1985.43$; $SD_{\text{birthyear}} = 13.51$), University of Belgrade (Serbia; $N = 164$,

80.5% female, $M_{\text{birthyear}} = 1993.73$; $SD_{\text{birthyear}} = 4.91$), Singapore Management University ($N = 135$, 56.2% female, $M_{\text{birthyear}} = 1993.80$; $SD_{\text{birthyear}} = 1.54$), Tsinghua University (China; $N = 174$, 62.2% female, $M_{\text{birthyear}} = 1993.68$; $SD_{\text{birthyear}} = 6.41$), University of Zürich (Switzerland; $N = 37$, 72.5% female, $M_{\text{birthyear}} = 1987.57$; $SD_{\text{birthyear}} = 8.72$), Virginia Commonwealth University (United States; $N = 150$, 78.8% female, $M_{\text{birthyear}} = 1992.62$; $SD_{\text{birthyear}} = 4.70$), University of Kassel (Germany; $N = 105$, 69.8% female, $M_{\text{birthyear}} = 1990.31$; $SD_{\text{birthyear}} = 7.82$), University of California, Santa Barbara (United States; $N = 108$, 63.8% female, $M_{\text{birthyear}} = 1995.82$; $SD_{\text{birthyear}} = 1.71$), Lusófona University, Lisbon (Portugal; $N = 18$, 33.3% female, $M_{\text{birthyear}} = 1984.12$; $SD_{\text{birthyear}} = 11.92$), University of Chile (Chile; $N = 34$, 62.9% female, $M_{\text{birthyear}} = 1979.33$; $SD_{\text{birthyear}} = 13.16$), University of Southampton (United Kingdom; $N = 6$, 50.0% female, $M_{\text{birthyear}} = 1992.17$; $SD_{\text{birthyear}} = 1.60$), Otto-Friedrichs-Universität Bamberg (Germany; $N = 40$, 69.0% female, $M_{\text{birthyear}} = 1982.11$; $SD_{\text{birthyear}} = 14.67$), Middle East Technical University (Turkey; $N = 181$, 65.7% female, $M_{\text{birthyear}} = 1992.42$; $SD_{\text{birthyear}} = 5.00$), University of Oslo (Norway; $N = 85$, 69.4% female, $M_{\text{birthyear}} = 1992.31$; $SD_{\text{birthyear}} = 6.42$), and SWPS University of Social Sciences and Humanities (Poland; $N = 133$, 86.6% female, $M_{\text{birthyear}} = 1986.18$; $SD_{\text{birthyear}} = 8.89$). Our total sample (after data exclusions) consisted of $N = 1507$ (68.9% female, $M_{\text{birthyear}} = 1990.95$; $SD_{\text{birthyear}} = 8.45$). We collected birth year instead of age due to an oversight. Data were collected in 2016.

Procedure. We created one central survey, which was translated and back translated keeping in mind loyalty to the original meaning (cf., Brislin, 1970). All surveys were programmed into the online survey platform Qualtrics. Participants were run online or in the lab across our different sites. Participants were again requested to complete the survey between 9–11am in their local time zone, not to eat or drink anything warm or cold for 10 minutes preceding the survey, and not to have exercised an hour preceding the survey. To be sure, we again asked whether they did eat or drink anything warm or cold 10 minutes before the study (“eatdrink”) or whether they had exercised an hour preceding the study (“exercise”). At the beginning and end of the task, participants again measured their own temperature with an oral thermometer of which they took a picture and uploaded this to our online platform (for an example, see **Figure 2**; descriptives, analysis script, and details of how the study was conducted at each site are available on our OSF project page; <https://osf.io/mc5gu/>).⁴ In our main study, the range of CSI was 0–12 and the average was 6.63.

Survey details. We used the same questionnaires as our pilot study, but added a few questions that seemed relevant for understanding the nature and structure of the relationship between CSI and CBT, like whether people are in a romantic relationship or not (“romantic”), how monogamous they perceive themselves to be (“monogamous”), and questions that pertain to the size of their online social networks (strength to one’s online social identity; “onlineid” and strength of attachment to one’s smartphone; “attachphone”; Yildirim & Correia, 2015). We also recorded participants’ longitude and latitude via

a standard option available in Qualtrics (“longitude”; we calculated latitude into equator distance “DEQ”). Finally, as the number of social contexts in which people are socially engaged may differ widely between cultures and language coding for “warm” and “cold” (Koptjevskaja-Tamm, 2015), we also included proxies for cultural influences with dummies for “language family” (Indo-European, Sino-Tibetan, and Uralic). Because cultural influences may be similarly large within a language family when the same language is spoken in highly different longitudinal locations (such as English spoken in the US versus that spoken in Singapore), we also included degrees longitude. At the end of the survey, participants were thanked for their participation and debriefed. We again looked up minimum temperature of that day and average humidity of that day through their IP address and the weather history site.

Data Handling. Before analyzing the data, for each scale variable, we checked the questionnaire’s reliability, and corrected labeling differences between sites where necessary (a complete file with all alterations can be requested from the first author). We then created a final “raw” dataset. Next, we reviewed all pictures that participants uploaded to our Qualtrics platform. We made 193 (mostly small) corrections to the CBT values, based on the picture participants uploaded. We also deleted 13 participants, as these participants uploaded either generic pictures or pictures that were irrelevant for our study.

When no picture was uploaded, we kept the participant in our dataset. We also deleted participants from our dataset who reported temperature values (“CBT” variable in the dataset) lower than 34.99 degrees Celsius, and participants that reported very unlikely temperature values (e.g., 100 degrees Celsius). Our final sample consisted of 1523 participants. Because we had a far larger N than our pilot study, we were somewhat more liberal with our inclusion on the basis of time of day, and left participants in even when they were not within our requested time frame. Instead, we included the time of day at which they ended the survey (“endtime”) as a control in the random forest and then in our mediation analyses.

Analyses and Results Cross-National Study

Degrees of freedom or sample size may differ throughout due to missing values for specific variables. We do not outline them here, but point to the data available from our project page. For our Cross-National Study, we split our analyses in two phases. In the first phase, we again relied on conditional random forests to specify variable importance, but now with both CSI and CBT as supervisory signals in two separate analyses. For both supervisory signals, we ran 8 versions (1 original and 7 replications; 4 different seeds, and 2 different levels of mtry), which ensured that we obtained the most stable model possible. The range Spearman Rank for CSI was $r = .984-.992$ and for CBT was $r = .939-.957$ (scripts, data,

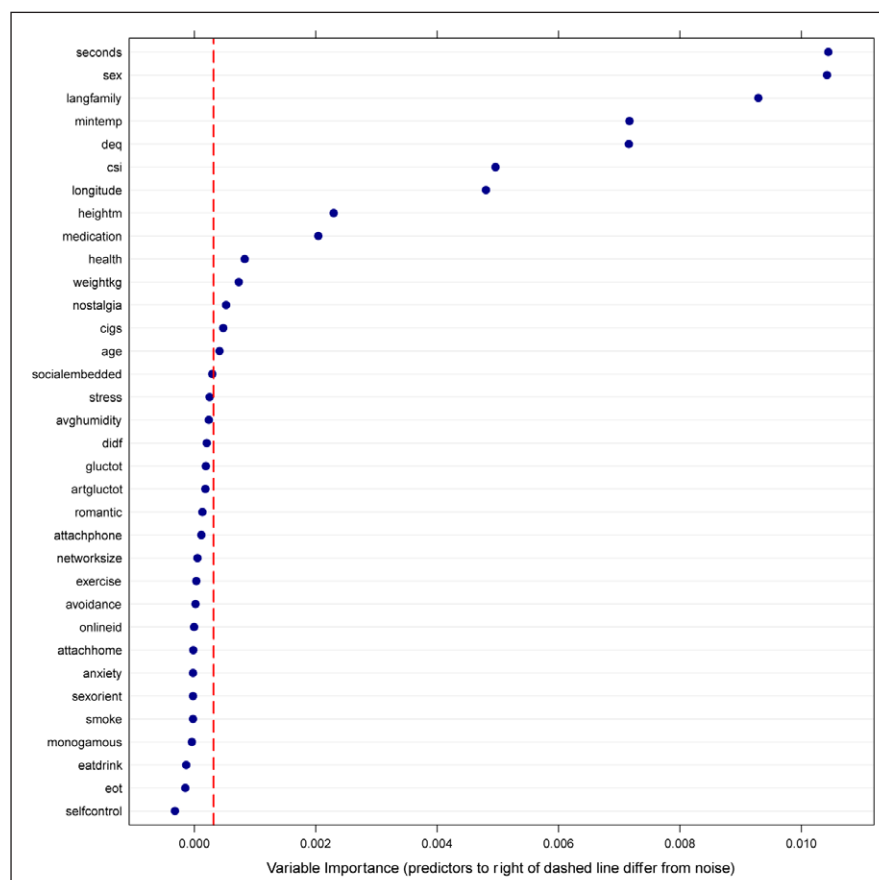


Figure 4: HPP Cross-National Study Dotplot for CSI. Permutation variable importance of predictors of Complex Social Integration from our supervised machine learning analyses in our pilot study. Variables exceeding the red line are very unlikely random noise.

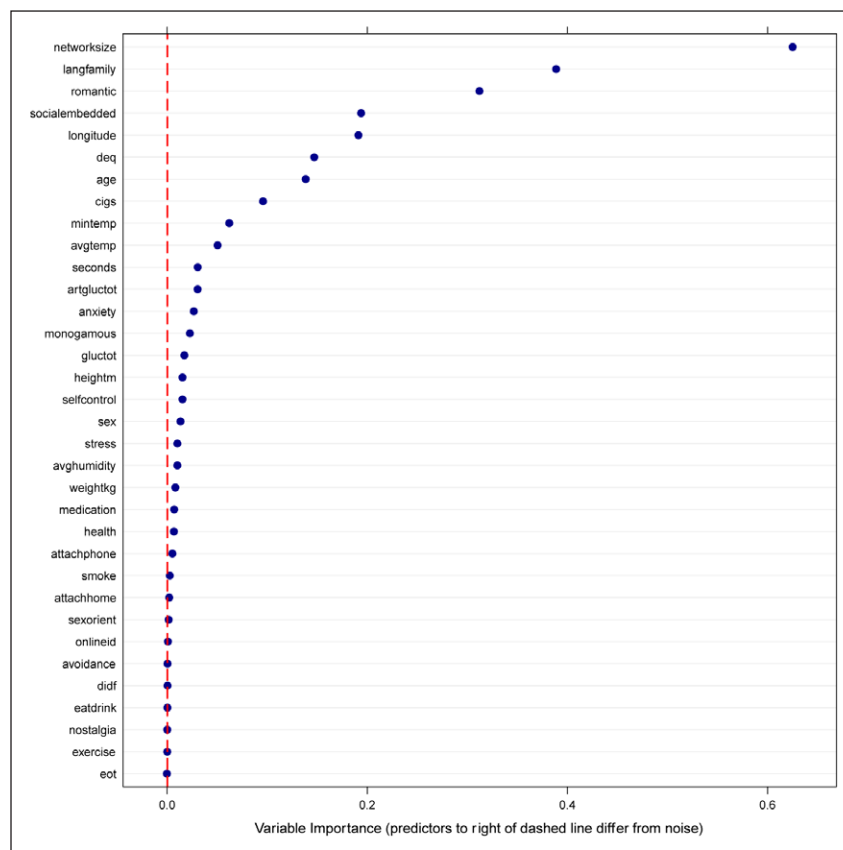


Figure 5: HPP Cross-National Study Dotplot for CBT. Permutation variable importance of predictors of Core Body Temperature from our supervised machine learning analyses in our pilot study. Variables exceeding the red line very likely differ from random noise.

and results are all available on our project page: <https://osf.io/mc5gu/files/>).

The below variables were the most important predictors of CBT (**Figure 4**; in the following order):

Time of day (seconds) > sex > language family of participant's language (langfamily) > minimum temperature of that day (mintemp) > distance from the equator (deq) > complex social integration (CSI) > longitude > participant's height (heightm) > whether participants took medication (meds)⁵ > participants' self-reported health (health) > participant's weight (weightkg).

The following were the most important predictors of CSI (**Figure 5**; in that order):

The size of participant's network > language family of participant's language (langfamily) > whether participants were in a romantic relationship (romantic) > Participants' social embeddedness in their network > longitude > distance from the equator (deq) > age > how many cigarettes the participant smokes (cigs) > minimum temperature of that day (mintemp) > core body temperature (avgtemp).

Our results thus show that distance from the equator (DEQ) and complex social integration (CSI) are amongst

the most important predictors of core body temperature (CBT), close to being as important as sex and "language family", and more important than known benchmarks like height, weight, and stress. Furthermore, DEQ is an important predictor of CSI (**Figure 5**). Because DEQ and mintemp correlate highly, we chose to retain DEQ in our further analyses (but comparable results are obtained when using mintemp).

Mediation by CSI. Based on these initial results and based on the relevance of distance from the equator and social integration, we decided to test a mediation hypothesis through a split-half cross-validation method. We created a training and test dataset through a random number generator in SPSS. We explored our predictions in our training dataset, which we then confirmed in our testing dataset.

For our mediation model, we were guided by our machine learning results. We hypothesized that distance from the equator would (positively) predict CSI, which in turn should (positively) predict CBT. CSI should then repress the (negative) relationship between DEQ and CBT. We conducted more traditional regression analyses to further understand the relationship between our most important variables. We did so again in the most robust way possible, by creating a training dataset (<https://osf.io/6v9d7/>) and a test dataset (<https://osf.io/qs2pb/>) and examining which hypotheses survived analyses in both datasets (splitting code can be found here: <https://osf.io/q5tga/>). To provide the highest informational value

possible, we report here the analyses over the entire dataset (<https://osf.io/nxuev/>; for the training mediation analyses see <https://osf.io/p9yj6/> and for the test mediation analyses see <https://osf.io/89juh/>).

In a regression, DEQ shows a robust relation with CSI (with $B = .014$, $t = 5.05$, $p < .001$, 95% CI [0.0083, 0.0188]).⁶ This means that for each degree increase in distance from the equator, the number of high-contact roles increases by .014. In turn, CSI is a positive predictor of CBT (in a regression; $B = .066$, $t = 8.40$, $p < .001$, 95% CI [0.051, 0.082]), meaning that for each point increase in high-contact roles there is a .066 degrees Celsius increase in core body temperature. DEQ on the other hand is a negative predictor of CBT ($B = -.0053$, $t = -6.56$, $p < .001$, 95% CI [-0.069, -0.0037]), meaning that each degree increase in distance from the equator lowers core body temperature by .0053 degree Celsius. These results thus show a robust relationship between DEQ, CSI, and CBT: Having a more varied active set of social relations relates to a higher CBT, while distance from the equator decreases it. People further away from the equator also have a more varied set of social relations.

Romantic relationships. Because our data are cross-sectional and thus only allow indirect causal statements, we conducted a number of extra analyses to explore the interrelationship between DEQ, CSI, and CBT. These again followed the logic of relying on using first a training and then a test dataset. First of all, because having a romantic relationship was a key predictor of CSI, we explored the influence of romantic relationships and found in our conditional random forests that the effects differed for those who do and those who do not have a romantic relationship. There are different possibilities for the mechanisms involved. Having a romantic relationship could provide the individual with an initial safe haven, making him or her less inhibited to explore and connect closely to others in various social contexts that can help protect core body temperature, as one would predict

on the basis of attachment theory (Bowlby, 1969). By contrast, having a romantic relationship could be a proxy for a *reduced* urgency to derive social warmth from CSI, in which case DEQ would be a weaker predictor of CSI for those with a romantic relationship.

We saw the former conjecture (i.e., explore and connect) supported in our conditional random forests and further confirmed in our mediation analyses. The mediation showing the relationship between DEQ, CSI, and CBT was moderated by having a romantic relationship or not in our training dataset (<https://osf.io/keqdu/>) and in our testing dataset (<https://osf.io/d2jhz/>). In our full dataset, having a romantic relationship moderated the link between DEQ and CSI ($B = -.034$, $t = -6.85$, $p < .001$, 95% CI [-0.044, -0.025]), as was predicted by our conditional random forests.

These analyses showed that both for people with and without a romantic relationship, this relationship was significant in the full dataset. However, it was only significant in the training *and* test dataset for those with a romantic relationship, and thus, that is the only one we consider robust. For people with a romantic relationship, the link of DEQ with CBT is to a significant degree mediated by CSI (such that people who live further from the equator score higher on CSI; see **Figure 6**). For those without a romantic relationship, the effect appeared to be in the opposite direction (the further away from the equator, the lower the score on CSI), but this effect was, as stated, unstable (see also **Figure 7**). We thus infer that CBT is buffered through CSI (i.e., that CSI raises CBT when ambient temperatures drop), and having a romantic relationship seems indeed to indicate an ability to engage in, and extend, the social network to generate warmth. That the mediation is not complete suggests that other regulation mechanisms also play a part in buffering core body temperature.

Network Size. There is yet another way to test for the role of CSI, and how humans may be distinct from the vervet monkeys whose core temperature is protected by

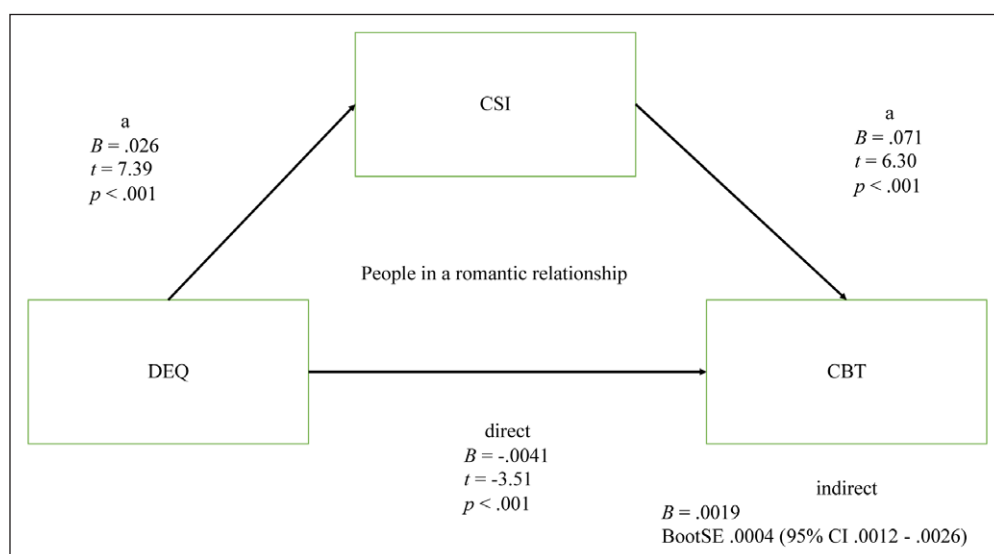


Figure 6: HPP Mediation Model for people in a romantic relationship. Mediation analyses showing how Complex Social Integration (CSI) protects the core temperatures (CBT) of people in a romantic relationship from colder climates (DEQ), c = significance with, and c' = significance without mediation by CSI.

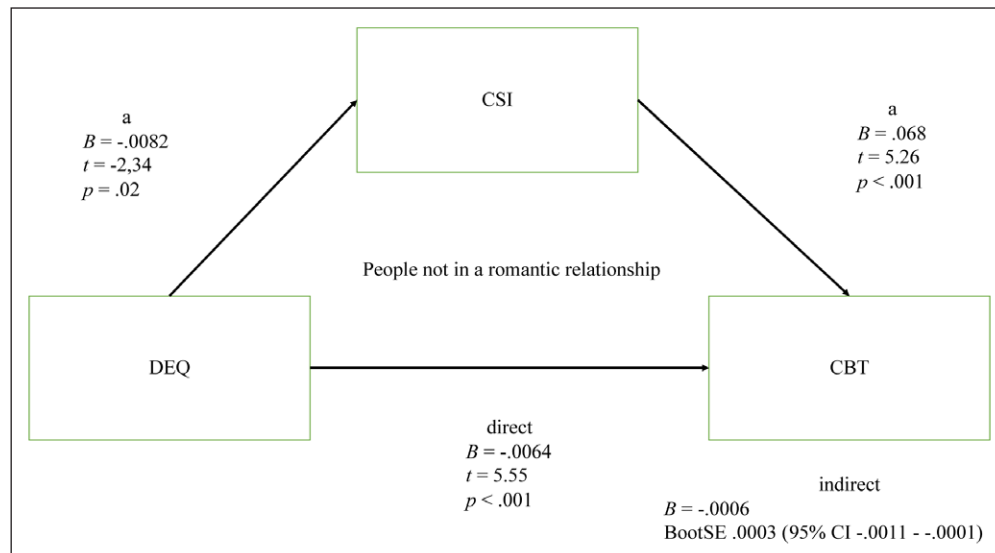


Figure 7: HPP Mediation Model for people not in a romantic relationship. Mediation analyses showing that Complex Social Integration (CSI), although significant, the model is instable (as we could not detect it in the training and testing datasets). We thus conclude that CSI does not protect the core temperatures (CBT) of people without a romantic relationship from colder climates (DEQ), c = significance with, and c' = significance without mediation by CSI.

the size of the social network. On the basis of the existing literature (Cohen & Lemay, 2007; Holt-Lunstad et al., 2010) and our machine learning results, we suspected that, with regard to physical effects of social warmth in humans, the *quality* of one's networks (i.e., CSI) is superior to the sheer size of people's networks. Of course, CSI correlates considerably with the size of people's social network ($r = .495, p < .001$). Yet, if it is indeed CSI rather than the sheer size of the network, then network size should not mediate the relationship between DEQ and CBT. Indeed, the mediation did not even survive our first exploratory analyses, as network size was not predicted by DEQ ($B = .03, t = 0.92, p = .36, 95\% \text{ CI } [-0.030, 0.084]$) nor did network size predict CBT ($B = .0005, t = .36, p = .72, 95\% \text{ CI } [-.0024, .0035]$). It is thus quality of the network, and not size, that matters.

In short, we infer that maintaining one's core body temperature is an important driver for CSI and thereby also has consequences for physical, social, and emotional functioning. At least for relational motivations that are derived from social thermoregulation, the results suggest that being closer to the equator makes one feel warmer and thus leads to less urgency to engage in CSI. Being further away from the equator influences the degree to which one engages in a higher level of CSI.

General Discussion

Our study finds robust evidence that maintaining one's core body temperature is an important driver for complex social integration and thereby also has consequences for physical, social, and emotional functioning. At least for relational motivations rooted in social thermoregulation, the results suggest that being closer to the equator makes one feel warmer and thus leads to less urgency to engage in complex social integration. Being further away from the equator influences the degree to which one engages in a higher level of complex social integration.

To our knowledge, the Human Penguin Project (HPP) is the first large scale study to empirically investigate the interrelation of distance from the equator (DEQ), complex social integration (CSI) and core body temperature (CBT). In a pilot study and a main, cross-national study spanning 12 countries and 3 continents, with various distances from the equator, we find (a) a considerable association between distance from the equator and CSI, and (b) a significant association between CSI and people's CBT for those who are seemingly not inhibited to socially connect (i.e., in our study: those with a romantic relationship). The data are very clear: CSI is closely intertwined with thermoregulation. We infer that many of our "older" physiological systems (like body temperature regulation) are crucial in shaping our modern ways of connecting with each other. What is also very clear from our data is that DEQ and CBT in relation to CSI are an important part of the story, but not the entire story. Culture (language family) plays a role that may be even more important for CSI, opening up the door to investigate interrelationships between socio-economic development, level of close-knittedness in cultures (Van de Vliert & Lindenberg, 2006), and linguistic structures (Koptjevskaja-Tamm, 2015) with complex social integration and temperature regulation.

We need to be clear: our studies and mediation analyses do not allow for direct causal statements. But as Bollen and Pearl (2013) suggest, it is possible to make some causal inferences about cross-sectional data, as "[causal] assumptions derive from prior studies, research design, scientific judgment, logical arguments, temporal priorities, and other evidence that the researcher can marshal in support of them". In our case, we cannot derive causality from our research design. Theoretically and intuitively however it is much less likely that people's core body temperature drives how far they live from the equator. It is even less likely that people take their entire social network further away from the equator (instead, distance from the

equator likely drives the extent to which one participates in various social networks). Based on prior studies, it is also relatively likely that the level of social integration is predictive of core body temperatures (although more stable core body temperatures could – in the long run – allow for greater exploration of the social network due to the possibility to explore, as would be predicted by social thermoregulation theory). We therefore make the inference that complex social integration protects from the cold. Our model (DEQ->CSI->CBT) should be targeted for replication research.

Although we derived our results from a large sample with very robust methods, it is worth mentioning that some effects in this literature have failed to replicate. This is not surprising, as published studies across scientific disciplines are heavily “underpowered” (i.e., too small to be able to support the tested hypothesis; Open Science Collaboration, 2015). It may then also not come as a surprise that also in the field of thermoregulation some effects failed to replicate (Vess, 2012; LeBel & Campbell, 2013). Yet, other effects did replicate (Schilder et al., 2014; Ebersole et al., 2016; IJzerman & Semin, 2009; IJzerman et al., 2016; Inagaki, Irwin, & Eisenberger, 2015), while some effects have been obtained with considerably larger samples, like those between 100 and 500 (e.g., IJzerman, Janssen, et al., 2015; Van Acker et al., 2016), or even around 30,000 (Hong & Sun, 2012) and above 6 million (Zwebner et al., 2013). Finally, extensive converging evidence exists with other species and human biological functioning that attests to the importance of social thermoregulation (for reviews, see IJzerman & Hogerzeil, 2018; IJzerman, Coan et al., 2015; Terrien et al., 2011). We suspect that data-driven approaches can help establish formal theoretical models outlining specific mechanisms and formulating specific predictions related to social thermoregulation theory.

The results from our work are robust, but the mechanisms via which people arrive at a higher core body temperature in colder environments are not yet clear. Why do complex social networks protect our bodies from the cold? We have answered this question indirectly in our theoretical introduction: People's modern forms of relationships probably grow out of more ancient relationships. That means that people used to huddle with each other, and that in modern relationships we still “track” people's trustworthiness and predictability by gauging whether they are cold or warm. This is confirmed in research showing that lower temperatures increase our desire to more frequently email or call loved ones (Van Acker et al., 2016), and research showing that “priming” people with loneliness lead them to estimate ambient temperature as lower (IJzerman & Semin, 2010).

But the more proximate mechanisms are not yet clear. We strongly suspect that direct co-thermoregulatory mechanisms exist. There are some indications that mothers increase their peripheral temperatures when their infants are in distress (Vuorenkoski et al. 1969). In adults, Wagemans and IJzerman (2014) found that people respond with peripheral temperature increases when seeing their sad partner, arguably to co-regulate their partner. The relationship literature is further replete

with suggestions that people physiologically co-regulate in the service of homeostasis, which we have argued to include temperature homeostasis (IJzerman, Heine, et al., 2017). From this perspective, temperature regulation has become implicated in attachment processes, which, in turn, form the basis for how people form predictions about others.

These manifest in individual differences in attachment (and thus how one forms one's social networks), which should be shaped according to the demands of one's physical (and social) environment. A meta-analysis of the Strange Situation has suggested that in Western European countries insecure avoidant people are relatively more prevalent, while in Israel and Japan the insecure ambivalent/resistant classification emerged as relatively more frequent (Van IJzendoorn & Kroonenberg, 1988). These personality differences are partly shaped by what is called “the adaptive landscape” (Buss, 2010), which includes disease prevalence (Schaller & Murray, 2008) and temperature. And temperature has been linked to personality differences. More clement climates (closer to 22 degrees Celsius), for example, are linked to greater conscientiousness, greater openness to experience, greater extraversion, greater agreeableness, and less neuroticism (Wei et al., 2017).

Future research should also focus more on the moderating role of romantic relationships between DEQ->CSI->CBT. So far, we had concluded that “having a romantic relationship could provide the individual with an initial safe haven, making her less inhibited to explore and connect closely to others in various social contexts that can help protect core temperature.” It is very possible that a third, latent, variable drives this effect, such as a secure attachment style. In our sample however, attachment security did not drive our results, as questions related to self-reported security in relationships do not map on well onto social thermoregulatory mechanisms. In new research, we have been developing a questionnaire assessing social thermoregulatory mechanisms (e.g., “I prefer to warm up with someone rather than with something”; Vergara et al., 2017) and find that these map onto attachment behaviors. We suspect that this questionnaire can help us assess whether the moderation was driven by a third, latent variable.

Conclusion

Although social thermoregulation is a hotly debated topic, the results from the Human Penguin Project (HPP) proved to be robust and open up new perspectives to investigate such co-thermoregulatory dynamics. We anticipate our study to be a starting point for other larger scale studies on the connection between temperature regulation, relationship quality, social integration, and health. In order to better understand the role of temperature in relationship (co-)regulation, future studies should assess social thermoregulation itself in more detailed ways, for example through longitudinal studies relying on modern sensor and actuator technologies. For this future work, our HPP study has made a crucial first step towards understanding how human social thermoregulation affects

the relationship between complex social integration and a key factor for our health: Our core body temperature.

Label legend for Dotplot Figures

1. Age = Participant's age.
2. Anxiety = Attachment anxiety (Fraley et al., 2000).
3. Artglucot = Artificial glucose intake based on diet drinks (Henriksen et al., 2014).
4. Avoidance = Attachment avoidance (Fraley et al., 2000).
5. Attachhome = Attachment to one's home (Harris et al., 1996).
6. Attachphone = Attachment to one's smartphone (Yildirim & Correia, 2015).
7. Avghumid = Average humidity of day of participation of participant's location.
8. CBT = Core body temperature
9. CSI = Number of high-contact roles (Cohen et al., 1997).
10. DIDF = Difficulty identifying feelings subscale of alexithymia (Kooiman et al., 2002).
11. Eatdrink = Did participant eat or drink something warm or cold 10 minutes before the study started?
12. EOT = Externally-oriented thinking subscale of alexithymia (Kooiman et al., 2002).
13. Equatordistance = Distance from the equator (calculated based on latitude).
14. Exercise = Did the participant exercise 60 minutes before the study started? (Yes/No).
15. Glucot = Glucose intake based on sugary drink intake (Henriksen et al., 2014).
16. Health = Participant's self-reported health (one-item).
17. Heightm = Height in meters.
18. Langfamily = Language family of participant's native language (Indo-European, Sino-Tibetan, or Uralic).
19. LocationLongitude = Longitude of participant's location.
20. Medication = Did participant use medication (yes vs. no).
21. Mintemp = Minimum temperature of day of participation of participant's location.
22. Monogamous = One item self-rated identification with monogamy.
23. Networksize = Number of people in the social network (Cohen et al., 1997).
24. Nostalgia = Nostalgia Proneness (Routledge et al., 2008).
25. Onlineid = Attachment to one's online identity (Yildirim & Correia, 2015).
26. Romantic = Participant's relationship status (In a relationship yes or no).
27. Seconds = Time of the day that participant finished the study (in seconds).
28. Selfcontrol = Self-control (Tangney et al., 2004).
29. Sex = Participant's sex.
30. Sexpref = Participant's sexual orientation.
31. Site = Platform where participant took part in pilot (mTurk vs. Prolific Academic).

32. Smoke = Does participant smoke? (Yes/No)

33. Socialembded = Number of embedded networks (Cohen et al., 1997).

34. Stress = Self-reported stress levels (Cohen & Wills, 1985).

35. Weightkg = Weight in kilograms.

Data Accessibility Statement

Data and materials are linked throughout the paper. Please note that only the shareable, deidentified versions are included (leaving out variables like age, sex, longitude, and sexual orientation). For researchers who want to have access to the full dataset, please contact the corresponding author.

Data, materials (including translated versions of all scales used in the project), and analyse code can be found at <https://osf.io/2rm5b/>.

Notes

¹ Past research (for example by IJzerman & Semin, 2009) has attributed thermoregulatory effects to metaphor theories (Lakoff & Johnson, 1983, 1999). We now believe this is wrong. Metaphor theories posit that connections between concrete experience and abstract concepts occur because activation in one area (e.g., for social interaction) becomes associated with superficially related areas (e.g., for physical warmth) in the brain. Metaphor theories also suggest unidirectionality (e.g., manipulating warmth should lead to closeness, but manipulating closeness should not lead to warmth) and cross-cultural universality (the metaphor should hold the world over).

It is now clear however that the central predictions of metaphor theories (uni-directionality and cross-cultural universality) have been falsified. Manipulations of loneliness do lead to changes in temperature perceptions (e.g., IJzerman & Semin, 2010) while linguistic metaphors related to warmth and affection are not universal, with many languages around the equator not showing the WARMTH is AFFECTION metaphor (Koptjesvkaja-Tamm, 2015). Furthermore, while some effects appear to support metaphor theories by showing an overlap of insular cortex activation for social and physical warmth (e.g., Inagaki & Eisenberger, 2013), other research has clarified that it is *subjective* temperature changes that lead to insular cortex activation (Craig et al., 2000) while many social and thermoregulatory effects already happen at much lower levels, like at the level of the medial pre-optic area of the hypothalamus (Boulant, 2000). The more appropriate way to conceptualize the underlying organization of neural systems related to social thermoregulation is as a "hierarchical prediction machine", where higher order areas (e.g., those related to insular cortex activation and monitoring social contact) help foster more efficient activity at lower levels (e.g., regulating temperature and social contact at the hypothalamic level; Clark, 2013; Friston, 2008).

² Clinicians typically consider oral temperature as a proxy for core temperature. Unfortunately, a recent

meta-analysis found that any thermometer other than central thermometers (such as rectal thermometers) are less reliable than desired for clinical purposes (Niven et al., 2015). Of course, for our (non-clinical) purposes, applying central thermometers (such as rectal ones) would have been highly impractical. Niven and colleagues (2015) recommended in such cases using an electronic oral thermometer. This is what the majority of our participants used ($N = 1119$). In addition, to ensure greater accuracy, we had participants measure their own oral temperature twice. Although we are aware that our approach introduced some noise, with our large sample size, two measurement points, and the second-best alternative to central thermometers, we are confident our oral temperature measures are sufficiently solid for our conclusions and sufficiently reflect core temperature.

³ There were two questionnaires that were included solely intended for scale development. These were the *Kama Muta Frequency Scale* and the *Social Thermoregulation and Risk Avoidance Questionnaire*. The analyses of these questionnaires are conducted in other projects.

⁴ There was one exception to the usage of the oral thermometer: Participants at UCSB used a temporal artery thermometer. To be sure, we ran the analyses with and without participants from UCSB. The effects for the full mediation model remained the same: There was a mediation for participants with a relationship (95% CI [.0005, .0015]), but not for participants without a relationship (95% CI [−.0001, .0004]), with a significant interaction between DEQ and having a romantic relationship or not onto CSI ($B = -.02$, $t = -5.05$, $p < .01$, 95% CI [−.03, −.01]). For analyses excluding UCSB sample, see <https://osf.io/b6r9v/>.

⁵ Although our model was stable, one reviewer noted that we should have included height and weight as controls in our analyses, based on our pilot study. We disagree. Note that both height and weight became a much less potent predictor of core body temperature in our cross-national project. Height (8th) and weight (11th) dropped as predictors of our model in our main study. In hindsight, that the effect becomes smaller should not be surprising. The main reason for the lesser importance of height and weight is because there were correlations between height (r (full dataset) = .146, $p < .001$) and weight (r (full dataset) = .144, $p < .001$) with DEQ in our sample. This is well-known: Across endotherms, within the same taxa, body size correlates with distance from the equator (something that has become known as “Bergmann’s rule” (Bergmann, 1847), which we thus also found in our data. Larger animals have a lower surface to body ratio, making them better able to stay warm in colder climates (something that is also true for modern humans; Foster & Collard, 2013). This does not mean that height and weight are not important in protecting from the cold, on the contrary. However, as we were mostly interested in social variables (and the machine learning analyses showed they had a separate and superior effect to height and weight in

predicting core body temperature) we did not include them in our analyses, at risk for overfitting. After all, height (16th) and weight (21th) were far less important for predicting CSI.

⁶ Because of missing values, degrees of freedom differ per analysis. The exact degrees of freedom can be obtained from our analyses output on our OSF project page (<https://osf.io/2rm5b/>).

Ethics and Consent

This research was approved under an “umbrella” ethics proposal at Vrije Universiteit, Amsterdam, and at each site where there was an ethics board (all ethics approvals can be downloaded from the project page from the individual site: <https://osf.io/2rm5b/>). We complied with ethics code outlined in the Declaration of Helsinki.

Acknowledgements

The authors thank Randy McCarthy (the editor) and John Edlund (reviewer) for helpful comments in revising the manuscript.

Funding Information

This research was supported by a Veni grant of the Netherlands Organization for Scientific Research (NWO) (016.145.049) and a French National Research Agency “Investissements d’avenir” program grant (ANR-15-IDEX-02) both awarded to Hans IJzerman.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

HIJ conceived of the idea (together with SL) and organized the project (together with RCV, ID, and SCW). HIJ wrote the paper and SL provided feedback on all steps of the process and wrote part of the paper. HIJ, ID, SCW, RCV, AHC, MVC, PD, NF, RH, CJH, YH, JJG, DL, LBL, MP, KGR, DR, SS, CS, MS, DS, JKB, RB, MH, FJG, TWB, AS, BS, and JHZ facilitated local data collection (which included but was not limited to translation and backtranslation of questionnaires, coordination of lab facilities or subject pools, and coordination with project leads about data collection). All authors provided comments on the paper at some stage.

References

- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33, 245–266. DOI: <https://doi.org/10.1017/S0140525X10000853>
- Benzinger, T. H. (1969). Clinical temperature: New physiological basis. *JAMA*, 209, 1200–1206. DOI: <https://doi.org/10.1001/jama.1969.03160210032008>
- Bergmann, C. (1847). Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien*, 3, 595–708.
- Bollen, K. A., & Pearl, J. (2013). Eight myths about causality and structural equation models. In: Morgan, S. (Ed.), *Handbook of Causal Analysis for Social Research*. Dordrecht, the Netherlands: Springer. DOI: https://doi.org/10.1007/978-94-007-6094-3_15

- Boulant, J. A.** (2000). Role of the preoptic-anterior hypothalamus in thermoregulation and fever. *Clinical Infectious Diseases*, 31, S157–S161. DOI: <https://doi.org/10.1086/317521>
- Bowlby, J.** (1969). *Attachment and Loss*. London: Basic Books.
- Breiman, L.** (1996). Bagging predictors. *Machine Learning*, 24, 123–140. DOI: <https://doi.org/10.1007/BF00058655>
- Breiman, L.** (2001). Random forests. *Machine Learning*, 45, 5–32. DOI: <https://doi.org/10.1023/A:1010933404324>
- Brislin, R. W.** (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185–216. DOI: <https://doi.org/10.1177/135910457000100301>
- Buss, D. M.** (2010). Personality and the adaptive landscape: The role of individual differences in creating and solving social adaptive problems. In: Buss, D. M., & Hawley, P. H. (Eds.), *The Evolution of Personality and Individual Differences*, 29–57. Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195372090.003.0002>
- Bylander, T., & Hanzlik, D.** (1999). Estimating generalization error using out-of-bag estimates. *AAAI-99 Proceedings*. Available at: <http://www.aaai.org/Papers/AAAI/1999/AAAI99-047.pdf>
- Clark, A.** (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204. DOI: <https://doi.org/10.1017/S0140525X12000477>
- Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr.** (1997). Social ties and susceptibility to the common cold. *JAMA*, 277, 1940–1944. DOI: <https://doi.org/10.1001/jama.1997.03540480040036>
- Cohen, S., & Lemay, E. P.** (2007). Why would social networks be linked to affect and health practices? *Health Psychology*, 26, 410. DOI: <https://doi.org/10.1037/0278-6133.26.4.410>
- Comtois, D.** (2016). Package 'summarytools'. Available at: <https://cran.r-project.org/web/packages/summarytools/summarytools.pdf>
- Craig, A. D., Chen, K., Bandy, D., & Reiman, E. M.** (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, 3, 184–190. DOI: <https://doi.org/10.1038/72131>
- Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A.** (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, 85, 894–908. DOI: <https://doi.org/10.1037/0022-3514.85.5.894>
- Deci, E. L., & Ryan, R. M.** (1995). Human autonomy. In: *Efficacy, Agency, and Self-esteem*, 31–49. Springer US. DOI: https://doi.org/10.1007/978-1-4899-1280-0_3
- Ebensperger, L. A.** (2001). A review of the evolutionary causes of rodent group-living. *Acta Theriologica*, 46, 115–144. DOI: <https://doi.org/10.4098/AT.arch.01-16>
- Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Brown, E. R., et al.** (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, 67, 68–82. DOI: <https://doi.org/10.1016/j.jesp.2015.10.012>
- Foster, F., & Collard, M.** (2013). A reassessment of Bergmann's rule in modern humans. *PloS One*, 8, e72269. DOI: <https://doi.org/10.1371/journal.pone.0072269>
- Fraley, R. C., Waller, N. G., & Brennan, K. A.** (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, 78, 350–365. DOI: <https://doi.org/10.1037/0022-3514.78.2.350>
- Friston, K.** (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4, e1000211. DOI: <https://doi.org/10.1371/journal.pcbi.1000211>
- Gallup, A. C., & Gallup, G. G.** (2008). Yawning and thermoregulation. *Physiology & Behavior*, 95, 10–16. DOI: <https://doi.org/10.1016/j.physbeh.2008.05.003>
- Hadi, R., Block, L., & King, D.** (2012). Mental thermoregulation: Affective and cognitive pathways for non-physical temperature regulation. *NA-Advances in Consumer Research*, 40. Abstract available online: <http://www.acrwebsite.org/volumes/1014152/volumes/v10e/E-10>
- Harris, P. B., Brown, B. B., & Werner, C. M.** (1996). Privacy regulation and place attachment: Predicting attachments to a student family housing facility. *Journal of Environmental Psychology*, 16, 287–301. DOI: <https://doi.org/10.1006/jev.1996.0025>
- Henriksen, R. E., Torsheim, T., & Thuen, F.** (2014). Loneliness, social integration and consumption of sugar-containing beverages: Testing the social baseline theory. *PloS one*, 9, e104421. DOI: <https://doi.org/10.1371/journal.pone.0104421>
- Hill, J. R., & Rahimtulla, K. A.** (1965). Heat balance and the metabolic rate of newborn babies in relation to environmental temperature and the effect of age and of weight on basal metabolic rate. *The Journal of Physiology*, 180, 239–265. DOI: <https://doi.org/10.1113/jphysiol.1965.sp007701>
- Hlavac, M.** (2015). Stargazer. Available at: <https://cran.r-project.org/web/packages/stargazer/stargazer.pdf>
- Holt-Lunstad, J., Smith, T. B., & Layton, J. B.** (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Med*, 7, e1000316. DOI: <https://doi.org/10.1371/journal.pmed.1000316>
- Hong, J., & Sun, Y.** (2012). Warm it up with love: The effect of physical coldness on liking of romance movies. *Journal of Consumer Research*, 39, 293–306. DOI: <https://doi.org/10.1086/662613>
- Ijzerman, H., Coan, J. A., Wagemans, F. M., Missler, M. A., van Beest, I., Lindenberg, S., & Tops, M.** (2015). A theory of social thermoregulation in human primates. *Frontiers in Psychology*, 6, 464. DOI: <https://doi.org/10.3389/fpsyg.2015.00464>
- Ijzerman, H., Gallucci, M., Pouw, W. T., Weißgerber, S. C., Van Doesum, N. J., & Williams, K. D.** (2012). Cold-blooded loneliness: Social exclusion leads to lower skin temperatures. *Acta Psychologica*, 140, 283–288. DOI: <https://doi.org/10.1016/j.actpsy.2012.05.002>
- Ijzerman, H., Heine, E. C., Nagel, S. K., & Pronk, T. M.** (2017). Modernizing relationship therapy

- through Social Thermoregulation Theory: Evidence, hypotheses, and explorations. *Frontiers in Psychology*, 8. DOI: <https://doi.org/10.3389/fpsyg.2017.00635>
- IJzerman, H., & Hogerzeil, L. J.** (2018). People as penguins: Thermoregulation as part of the human essence. In: van Zomeren, M., & Dovidio, J. (Eds.), *Oxford Handbook of Human Essence*. Oxford University Press. Preprint available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2844963.
- IJzerman, H., Janssen, J. A., & Coan, J. A.** (2015). Maintaining warm, trusting relationships with brands: Increased temperature perceptions after thinking of communal brands. *PloS One*, 10, e0125194. DOI: <https://doi.org/10.1371/journal.pone.0125194>
- IJzerman, H., Pollet, T., Ebersole, C., & Kun, D.** (2016). What predicts Stroop performance? A conditional random forest approach. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2805205.
- IJzerman, H., & Semin, G. R.** (2009). The thermometer of social relations: Mapping social proximity on temperature. *Psychological Science*, 20, 1214–1220. DOI: <https://doi.org/10.1111/j.1467-9280.2009.02434.x>
- IJzerman, H., & Semin, G. R.** (2010). Temperature perceptions as a ground for social proximity. *Journal of Experimental Social Psychology*, 46, 867–873. DOI: <https://doi.org/10.1016/j.jesp.2010.07.015>
- IJzerman, H., Szymkow, A., & Parzuchowski, M.** (2016). Warmer hearts, and warmer, but noisier rooms: Communality does elicit warmth, but only for those in colder ambient temperatures—Commentary on Ebersole et al. *Journal of Experimental Social Psychology*, 67, 88–90. DOI: <https://doi.org/10.1016/j.jesp.2015.12.004>
- Inagaki, T. K., & Eisenberger, N. I.** (2013). Shared neural mechanisms underlying social warmth and physical warmth. *Psychological Science*, 24, 2272–2280. DOI: <https://doi.org/10.1177/0956797613492773>
- Inagaki, T. K., Irwin, M. R., & Eisenberger, N. I.** (2015). Blocking opioids attenuates physical warmth-induced feelings of social connection. *Emotion*, 15, 494. DOI: <https://doi.org/10.1037/emo0000088>
- Inagaki, T. K., Irwin, M. R., Moieni, M., Jevtic, I., & Eisenberger, N. I.** (2016). A pilot study examining physical and social warmth: Higher (non-febrile) oral temperature is associated with greater feelings of social connection. *PloS One*, 11, e0156873. DOI: <https://doi.org/10.1371/journal.pone.0156873>
- Janský, L.** (1973). Non-shivering thermogenesis and its thermoregulatory significance. *Biological Reviews*, 48, 85–132. DOI: <https://doi.org/10.1111/j.1469-185X.1973.tb01115.x>
- Jones, Z., & Linder, F.** (2015). Exploratory data analysis using random forests. Prepared for the 73rd annual MPSA conference. Available online at: http://zmjones.com/static/papers/rfss_manuscript.pdf.
- Kernis, M. H.** (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, 14, 1–26. DOI: https://doi.org/10.1207/S15327965PLI1401_01
- Kooiman, C. G., Spinhoven, P., & Trijsburg, R. W.** (2002). The assessment of alexithymia: Acritical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20. *Journal of Psychosomatic Research*, 53, 1083–1090. DOI: [https://doi.org/10.1016/S0022-3999\(02\)00348-3](https://doi.org/10.1016/S0022-3999(02)00348-3)
- Koptjevskaja-Tamm, M.** (2015). *The linguistics of temperature*. Amsterdam: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/tsl.107>
- LeBel, E. P., & Campbell, L.** (2013). Heightened sensitivity to temperature cues in individuals with high anxious attachment. Real or elusive phenomenon? *Psychological Science*, 24, 2128–2130. DOI: <https://doi.org/10.1177/0956797613486983>
- Marazziti, D., Di Muro, A., & Castrogiovanni, P.** (1992). Psychological stress and body temperature changes in humans. *Physiology & Behavior*, 52, 393–395. DOI: [https://doi.org/10.1016/0031-9384\(92\)90290-I](https://doi.org/10.1016/0031-9384(92)90290-I)
- McFarland, R., Fuller, A., Hetem, R. S., Mitchell, D., Maloney, S. K., Henzi, S. P., & Barrett, L.** (2015). Social integration confers thermal benefits in a gregarious primate. *Journal of Animal Ecology*, 84, 871–878. DOI: <https://doi.org/10.1111/1365-2656.12329>
- McFarland, R., Henzi, S. P., Barrett, L., Wanigaratne, A., Coetzee, E., Fuller, A., Maloney, S. K., et al.** (2015). Thermal consequences of increased pelt loft infer an additional utilitarian function for grooming. *American Journal of Primatology*, 78, 456–461. DOI: <https://doi.org/10.1002/ajp.22519>
- Minard, D., Copman, L., & Dasler, A. R.** (1964). Elevation of body temperature in health. *Annals of the New York Academy of Sciences*, 121, 12–25. DOI: <https://doi.org/10.1111/j.1749-6632.1964.tb13680.x>
- Mizukami, K., Kobayashi, N., Ishii, T., & Iwata, H.** (1990). First selective attachment begins in early infancy: A study using telethermography. *Infant Behavior and Development*, 13, 257–271. DOI: [https://doi.org/10.1016/0163-6383\(90\)90034-6](https://doi.org/10.1016/0163-6383(90)90034-6)
- Niven, D. J., Gaudet, J. E., Laupland, K. B., Mrklas, K. J., Roberts, D. J., & Stelfox, H. T.** (2015). Accuracy of peripheral thermometers for estimating temperature: A systematic review and meta-analysis. *Annals of Internal Medicine*, 163, 768–777. DOI: <https://doi.org/10.7326/M15-1150>
- Núñez-Villegas, M., Bozinovic, F., & Sabat, P.** (2014). Interplay between group size, huddling behavior and basal metabolism: An experimental approach in the social degu. *Journal of Experimental Biology*, 217, 997–1002. DOI: <https://doi.org/10.1242/jeb.096164>
- Open Science Collaboration.** (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716. DOI: <https://doi.org/10.1126/science.aac4716>
- Peters, R. H.** (1986). *The ecological implications of body size*. Cambridge University Press.
- Ripley, B.** (2016). Package ‘tree’. Available at: <https://cran.r-project.org/web/packages/tree/tree.pdf>.
- Rotman, J. D., Lee, S. H. M., & Perkins, A. W.** (2016). The warmth of our regrets: Managing regret through physiological regulation and consumption. *Journal of Consumer Psychology*, 2, 160–170.

- Routledge, C., Arndt, J., Sedikides, C., & Wildschut, T.** (2008). A blast from the past: The terror management function of nostalgia. *Journal of Experimental Social Psychology*, 44, 132–140. DOI: <https://doi.org/10.1016/j.jesp.2006.11.001>
- Sarkar, D.** (2017). Lattice: Trellis Graphics for R. Available at: <https://cran.r-project.org/web/packages/lattice/index.html>.
- Satinoff, E.** (1982). Are there similarities between thermoregulation and sexual behavior? In: *The physiological mechanisms of motivation*, 217–251. Springer New York. DOI: https://doi.org/10.1007/978-1-4612-5692-2_8
- Satinoff, E.** (1983). A reevaluation of the concept of the homeostatic organization of temperature regulation. In: *Motivation*, 443–472. Springer US. DOI: https://doi.org/10.1007/978-1-4684-4286-1_11
- Schaller, M., & Murray, D. R.** (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, 95, 212–221. DOI: <https://doi.org/10.1037/0022-3514.95.1.212>
- Schilder, J. D., IJzerman, H., & Denissen, J. J.** (2009). Physical warmth and perceptual focus: A replication of IJzerman and Semin. *PloS One*, 9, e112772. DOI: <https://doi.org/10.1371/journal.pone.0112772>
- Seeman, T. E.** (1996). Social ties and health: The benefits of social integration. *Annals of Epidemiology*, 6, 442–451. DOI: [https://doi.org/10.1016/S1047-2797\(96\)00095-6](https://doi.org/10.1016/S1047-2797(96)00095-6)
- Steverink, N., & Lindenberg, S.** (2008). Do good self-managers have less physical and social deficits and more well-being in later life? *European Journal of Ageing*, 5, 181–190. DOI: <https://doi.org/10.1007/s10433-008-0089-1>
- Strobl, C., Malley, J., & Tutz, G.** (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14, 323–348. DOI: <https://doi.org/10.1037/a0016973>
- Tangney, J. P., Baumeister, R. F., & Boone, A. L.** (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271–324. DOI: <https://doi.org/10.1111/j.0022-3506.2004.00263.x>
- Terrien, J., Perret, M., & Aujard, F.** (2011). Behavioral thermoregulation in mammals: A review. *Frontiers in Bioscience*, 16, 1428–1444. DOI: <https://doi.org/10.2741/3797>
- Van Acker, B., Kerselaers, K., Pantophlet, J., & IJzerman, H.** (2016). Homelike thermoregulation: How physical coldness makes an advertised house a home. *Journal of Experimental Social Psychology*, 67, 20–27. DOI: <https://doi.org/10.1016/j.jesp.2015.10.008>
- Van de Vliert, E., & Lindenberg, S.** (2006). Wealth, climate, and framing: Cross-national differences in solidarity. In: Fetchenhauer, D., Flache, A., Buunk, A. P., & Lindenberg, S. (Eds.), *Solidarity and Prosocial Behavior: An Integration of Sociological and Psychological Perspectives*. New York: Springer. DOI: https://doi.org/10.1007/0-387-28032-4_13
- Van IJzendoorn, M. H., & Kroonenberg, P. M.** (1988). Cross-cultural patterns of attachment: A meta-analysis of the Strange Situation. *Child Development*, 59, 147–156. DOI: <https://doi.org/10.2307/1130396>
- Vergara, R. C., Hernández, C., Jaume-Guazzini, F., Lindenberg, S., & IJzerman, H.** (2017). Development and validation of the Social Thermoregulation and Risk Avoidance Questionnaire (STRAQ-1). Unpublished dataset.
- Vess, M.** (2012). Warm thoughts: Attachment anxiety and sensitivity to temperature cues. *Psychological science*, 23, 472–474. DOI: <https://doi.org/10.1177/0956797611435919>
- Vuorenkoski, V., Wasz-Hockert, O., Koivisto, E., & Lind, J.** (1969). The effect of cry stimulus on the temperature of the lactating breast of primipara. A thermographic study. *Experientia*, 25, 1286–1287. DOI: <https://doi.org/10.1007/BF01897502>
- Wagemans, F. A. M., & IJzerman, H.** (2014). Social thermoregulation in close relationships. *Poster presented at SPSP*, available at Figshare: https://figshare.com/articles/Social_thermoregulation_in_close_relationships/926484.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A.** (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7, 632–638. DOI: <https://doi.org/10.1177/1745691612463078>
- Wei, W., Lu, J. G., Galinsky, A. D., Wu, H., Gosling, S. D., Rentfrow, P. J., Gui, W., et al.** (2017). Regional ambient temperature is associated with human personality. *Nature Human Behaviour*, 1, 890. DOI: <https://doi.org/10.1038/s41562-017-0240-0>
- Wickham, H.** (2016). Package ‘plyr’. Available at: <https://cran.r-project.org/web/packages/plyr/index.html>.
- Yarkoni, T., & Westfall, J.** (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspective on Psychological Science*. DOI: <https://doi.org/10.1177/1745691617693393>
- Yildirim, C., & Correia, A. P.** (2015). Exploring the dimensions of nomophobia: Development and validation of a self-reported questionnaire. *Computers in Human Behavior*, 49, 130–137. DOI: <https://doi.org/10.1016/j.chb.2015.02.059>
- Zhou, X., Wildschut, T., Sedikides, C., Chen, X., & Vingerhoets, A. J.** (2012). Heartwarming memories: Nostalgia maintains physiological comfort. *Emotion*, 12, 678–684. DOI: <https://doi.org/10.1037/a0027236>
- Zwebner, Y., Lee, L., & Goldenberg, J.** (2013). The temperature premium: Warm temperatures increase product valuation. *Journal of Consumer Psychology*, 24, 251–259. DOI: <https://doi.org/10.1016/j.jcps.2013.11.003>

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments are available at: <http://doi.org/10.1525/collabra.165.pr>

How to cite this article: IJzerman, H., Lindenberg, S., Dalğar, İ., Weissgerber, S. C., Vergara, R. C., Cairo, A. H., Čolić, M. V., Dursun, P., Frankowska, N., Hadi, R., Hall, C. J., Hong, Y., Hu, C.-P., Joy-Gaba, J., Lazarević, D., Lazarević, L. B., Parzuchowski, M., Ratner, K. G., Rothman, D., Sim, S., Simão, C., Song, M., Stojilović, D., Blomster, J. K., Brito, R., Hennecke, M., Jaime-Guazzini, F., Schubert, T. W., Schütz, A., Seibt, B., & Zickfeld, J. H. (2018). The Human Penguin Project: Climate, Social Integration, and Core Body Temperature. *Collabra: Psychology*, 4(1): 37. DOI: <https://doi.org/10.1525/collabra.165>

Senior Editor: Simine Vazire

Guest Editor: Randy McCarthy

Submitted: 27 April 2018 **Accepted:** 24 September 2018 **Published:** 19 October 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.