The behavioral and energetic responses of montane grasshoppers to climate change and the effects of competition in introductory college STEM classrooms

Julia Smith

A dissertation

Submitted in partial fulfillment of

Requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Lauren Buckley, Chair

Elli Theobald, Chair

Berry Brosi

Program Authorized to Offer Degree:

Biology

University of Washington

Abstract

The behavioral and energetic responses of montane grasshoppers to climate change and the effects of competition in introductory college STEM classrooms

Julia Smith

Chair of the Supervisory Committee:

Lauren Buckley and Elli Theobald

Department of Biology

Understanding thermal physiology can help us make mechanistic predictions about the organismal impacts of climate change. In my two ecology chapters, I explore the consequences of thermal physiology and behavior for the energy gain of montane grasshoppers. My study species – *Melanoplus boulderensis* and *Melanoplus sanguinipes* – are found along an elevational gradient (1740-3515m) in Colorado, and I compare populations at different sites along this gradient. I measure the temperature sensitivity of digestion (Chapter 1) and diel patterns in field body temperature, activity, gut fullness, and preferred body temperature (Chapter 2). I interpret my findings using a simple energy budget, accounting for the energy gained through digestion of food and the energy expended through metabolism (both temperature-dependent processes). Past work suggests that these grasshoppers optimize their digestion at higher temperatures (~40°C) than their locomotion (~25°C) and most environmental temperatures they encounter. This led me to hypothesize that this difference in thermal optima could be adaptive if grasshoppers are using microhabitat selection and temporal partitioning of activities. Specifically, I expected that grasshoppers would feed in the morning, filling their guts by the afternoon. At these peak afternoon temperatures, they would bask in a warm microclimate to optimize their digestion. In Chapter 1, I find that grasshoppers do indeed optimize their digestion – and consequently energy gain – at high temperatures ~40°C, with minimal differences between populations in thermal optima but a of higher peak digestion for populations at higher elevations. Consequently, both climate and physiology affect population energy balances, with both species exhibiting counter-gradient variation. I find that, since metabolic costs have a minimal effect on net energy gain assuming abundant food, climate change has led to an increase in net energy gain of grasshoppers. In Chapter 2, I find that

field body temperatures are elevated in cool conditions. *M. sanguinipes* individuals select a higher body temperature in a thermal gradient in the afternoon than they do in the morning, but feeding state (unfed or fed) does not significantly affect thermal preference. I do not find evidence for diel patterns in gut fullness and the best model for predicting activity includes only the operative temperature and not time of day. Unlike I predicted, grasshoppers are most active at high temperatures, which generally correspond to midday, and guts are not fullest at midday. However, the evidence I found for behavioral thermoregulation via microhabitat selection supports the hypothesis grasshoppers are taking advantage of thermal opportunity to gain more energy.

My third chapter is in the field of discipline-based education research. Understanding the interplay between contrasting theories of motivation can help us make mechanistic predictions about the effects of classroom experiences on student outcomes. Building on past literature about classroom experience, I use structural equation models to explore the relationships between students' perception of classroom competition, sense of control, sense of belonging, and exam performance. Two contrasting theories of motivation, Control-Value Theory and Self-Determination Theory, posit sense of control and sense of belonging as keys to motivation and ultimately performance. I leverage a large survey dataset of students in introductory college STEM classes in the United States. I test contrasting mechanisms for the impact of competition on exam z-score: an indirect effect via sense of control and an indirect effect via sense of belonging. I find that perceived competition lowers students' sense of belonging, sense of control, and, via sense of control, exam performance. I also disaggregate these effects by race to investigate whether the impact of perceived competition varies for students with different racial identities. I refer to the Quantitative Critical Race Theory literature to guide our analysis and interpretation of these results. I find that the relationships between perceived competition, sense of control, and sense of belonging do not differ for students with different racial identities. I do, however, find differences in the mean levels of perceived competition, sense of belonging, and sense of control among students with different identities. These differences reflect systems of oppression in our society and our classrooms.

Table of contents

Chapter 1: How do climate and population differences affect grasshopper energy gain across an elevational gradient?

In preparation for publication

Abstract

Grasshoppers can feed at exceptionally hot temperatures among insects, which has implications for rates of energy gain and responses to climate change, particularly in cool, montane environments. Does temperature sensitivity vary along a montane elevation gradient and what are the energetic implications? We assess the temperature dependence of digestion rate (thermal performance curves, or TPCs) for populations of two grasshopper species – a thermal generalist (*Melanoplus sanguinipes*) and a montane specialist (*Melanoplus boulderensis*) – along an elevation gradient. We find that the temperature dependence of digestion is similar across elevations, but higher elevation populations tend to have higher maximum rates. We use the TPCs as the basis for an energy budget model balancing estimated rates of energy acquisition against estimated rates of energy use. We conduct a simulated reciprocal transplant, which indicates that the majority of differences in net energy gain across elevation are attributable to climate rather than population-specific traits (physiology and body size) for *M. boulderensis* and the converse for *M. sanguinipes*. In both species, the directions of the trends are the same: net energy gain is maximized at low-elevation sites and by high-elevation populations. We estimate that 65 years of warming (from 1950-1959 to 2015-2024) has increased net energy gain for grasshoppers at all sites, especially high-elevation sites, where the most warming has occurred. The digestion rates of our focal grasshopper species suggest selection to capitalize on limited thermal opportunity and that the potential to capitalize will increase with warming.

Introduction

Forecasting the biological impacts of climate change remains a challenge. Organismal responses are variable and idiosyncratic. Mechanistic approaches that account for organismal thermal biology may offer more robustness and accuracy than correlative (phenomenological) models (Briscoe et al., 2023) especially because mechanisms may extrapolate better into novel climates (Evans et al., 2015; Williams & Jackson, 2007).

Temperature strongly influences organismal performance, energy use, and fitness by constraining physiological rates and behavior (Kingsolver, 2009). Despite the central role of thermal biology, predicting how organisms will respond to current and future climate change remains a challenge (Srivastava et al., 2021). In ectotherms, spatial and temporal variation in thermal environments can alter energy acquisition and metabolic demands in ways that shape population dynamics and influence development, behavior, and other responses. Mechanistic approaches that link temperature-dependent performance to energetics provide a powerful framework for understanding how organisms respond along environmental gradients and to environmental change (Briscoe et al., 2023).

Understanding whether survival or fecundity predominates as a fitness constraint for a given population can help predict its response to climate change. Rare, extreme temperatures can create selective pressures via survival, while mean temperatures affect fecundity via energy use and acquisition once there is sufficient net energy gain for survival (Buckley et al., 2021A). Both temperature variability and means are forecasted to increase under climate change. Which component is more salient for an organism depends on that organism's thermal tolerance, its thermal sensitivity of performance, and the temperatures it experiences in its environment. Moreover, the constraining fitness component – survival or fecundity – can vary for a species spatially along an environmental gradient and temporally under climate change. At lower temperatures (i.e. the leading edge of the species range under climate change), fecundity may be the larger constraint, while at higher temperatures (the trailing range edge) survival may be more constraining. Measuring the temperature dependence of these individual fitness components and understanding which constraints are more salient in a given situation can improve forecasting of the organismal impacts of climate change (Buckley et al., 2021A).

The temperature dependence of energy acquisition and use provides one promising mechanistic approach to quantifying the fitness effects of climate change. The temperatures organisms experience over their lives cumulatively affect fecundity by constraining their activity time for acquiring and processing food and the discretionary energy they have left for growth and reproduction after accounting for nutritional gains and metabolic costs (Kingsolver and Buckley, 2017; Buckley et al., 2021A; Kearney and Porter, 2004). When food availability is not limited, the temperature dependence of digestion tends to limit

energy gain more than consumption or foraging time (Angilletta, 2001; Harrison & Fewell, 1995). The temperature dependence of digestion provides a basis for predicting population dynamics and distributions (Levy et al., 2017; Youngblood et al., 2023). Thermoregulation can importantly alter ectotherm energy budgets as they select body temperatures in order to maximize their energy gain or seek refuge from thermal extremes (Coggan et al., 2011; Scheffers et al., 2014; Stark et al., 2023). Energetic determinants of fecundity can be complemented by modeling the impacts of short-term, damaging temperature extremes on survival (Amarasekare & Savage, 2012; Buckley et al., 2021A; Levy et al., 2015).

Energy budgets can be estimated using thermal performance curves (TPCs), which describe how the rate of a fitness-conferring performance varies depending on the organism's body temperature. As body temperature rises from the lower to the upper thermal limit for performance (*Tmin* and *Tmax*, respectively), performance tends to increase gradually to peak performance (*Pmax*) at the thermal optimum (*Topt*) and then to decline rapidly (Huey & Kingsolver, 1989). Biochemical rates and physiological stress constrain the shape of TPCs (Logan & Cox, 2020), but TPC shape remains a key form of adaptation to environmental gradients (Buckley & Kingsolver, 2021; Kingsolver, 2009). The construction and application of TPCs involves numerous assumptions, but TPCs can be integrated over environmental conditions as a powerful, if coarse, tool to estimate performance and fitness responses to variable and changing environments (Sinclair et al., 2016).

Different physiological processes are often optimized at different body temperatures, shaping behavioral thermoregulation and the timing of activities in ectotherms (Huey & Kingsolver, 1989). In grasshoppers, digestion tends to peak at relatively high body temperatures, creating tradeoffs between maximizing energy gain during warm periods and incurring higher metabolic costs, which rise exponentially with temperature (Dillon et al., 2010). In cool, seasonal environments, rapid digestion during brief windows of favorable warmer temperatures may allow grasshoppers to capitalize on limited thermal opportunities, potentially influencing growth, reproduction, and population persistence (Nufio et al., 2025). The balance between temperature-dependent energy acquisition and metabolic expenditure determines

the net energy gain, which is expected to vary across elevational gradients and to shift under recent climate warming.

In this study, we use a mechanistic, energetics-based approach to examine how temperature-dependent digestion performance and metabolic costs shape net energy gain across an elevational gradient in grasshoppers. Grasshoppers play an important role in nutrient cycling and increased herbivory can accelerate nutrient cycling in some systems (Belovsky & Slade, 2000). Past work on the thermal physiology of grasshoppers has characterized temperature-dependent measures of feeding, digestion, assimilation efficiency, and metabolic expenditure, providing a mechanistic link between thermal environments and fitness-relevant outcomes (Harrison & Fewell, 1995; Belovsky, 1986, Buckley et al., 2014; Buckley & Nufio, 2014). We focus on two montane species in Boulder County, Colorado, that differ in their elevational distributions, thermal biology and dispersal limitation. *Melanoplus sanguinipes* is a mid-season thermal generalist found along the high plains to the subalpine, while *Melanoplus boulderensis* is a short winged early-season thermal specialist restricted to montane environments. We estimate digestion thermal performance curves (TPCs) for populations along the elevational gradient. We construct an energy budget using these TPCs, metabolic rates, and modeled body temperatures over the growing season in order to estimate net energy gain across the elevation gradient under historical (1950–1959) and contemporary (2015–2024) climate conditions. This approach allows us to evaluate how climate, population- and species- specific responses, and their interaction influence energy budgets across space and time. We assess how recent climate warming has altered energetic constraints. By linking thermal biology, energetics, and environmental change, our study aims to clarify the role of energy acquisition in shaping organismal responses to climate warming.

We predict that both climate and physiology will affect net energy gain. The warming from the historical time period to the contemporary time period and from high-elevation sites to low may well increase hoppers' net energy gain, particularly at high elevations that have experienced more warming. We build on a past study showing that hoppers' digestion is optimized at high temperatures by extending the test temperatures above 40°C, which is near the digestion *Topt* (Buckley & Nufio, 2014). Still, metabolic costs also increase at higher temperatures, so the overall effect of warming on net energy gain

depends on the relative magnitude of the increases in acquisition and costs. We predict that, especially in the more genetically differentiated *M. boulderensis* (Slatyer et al., 2020), populations at different sites along the elevational gradient will exhibit TPCs that align with their environment. For example, a high-elevation population may have a lower *Topt* than a low-elevation population since it would be beneficial to optimize digestion at a temperature that is more common in one's environment.

Methods

Study system

We studied grasshopper populations along an elevational gradient in Boulder County, Colorado, which has been extensively studied since the 1950s (Alexander & Hilliard, 1969; Nufio et al., 2010). We sampled populations from 5 sites at different elevations: 1,740m, 2,195m, 2,591m, 3,048m, and 3,515m (average population sample size: 48, range: 30-77). These elevations include the upper limit of *M. sanguinipes* and the lower limit of *M. boulderensis*, so *M. sanguinipes* only occurred at the lower 3 sites (1,740-2,591m), while *M. boulderensis* only occurred at the upper four sites (2,195-3,515m; Table S1). We collected *M. sanguinipes* individuals in August of 2023 and we collected *M. boulderensis* individuals in July of 2023 and 2024.

Thermal performance curves

To quantify the temperature dependence of digestion, we measured gut throughput as feces production at controlled body temperatures following Harrison and Fewell (1995). Adult grasshoppers were collected from each site and, within two days of capture, transported to Seattle by air (carry-on) or overnight shipment. Transported individuals were housed with others from the same site and sex in Ziploc bags with wheatgrass. Upon arrival, grasshoppers were weighed (Mettler Toledo XSR105DU) and housed individually in plastic tubs with wheatgrass. Individuals acclimated to the Percival (model 36VL) environmental chambers (27 °C, 14 h light / 21 °C, 10 h dark) for at least one day and then they were fasted overnight at 25 °C for at least 12 hours to empty the gut, with moisture provided via a damp paper towel.

Each grasshopper was assayed at three of six test temperatures (16, 23, 30, 37, 40, and 43 °C) over three consecutive days in a semi-randomized order. On each assay day, individuals were acclimated

to the test temperature for one hour before being transferred to tubs containing 1 g of hydrated wheatgrass, which exceeded consumption in all trials. We inserted the wheatgrass in floral tubes for hydration. After 8 hours, fecal material was collected, dried at 50–65 °C for at least 72 hours, and weighed to estimate mass-specific feces production rates. We chose an 8-hour duration to primarily reflect digestion rather than feeding rates (Harrison and Fewell 1995).

We used the R package brms to fit TPCs with Bayesian statistics, referring to the methods of Siemers et al. (2024)**.** Our dependent variable was dried feces mass / hopper mass / time (mg/g/hr). We selected the Lobry-Rosso-Flandrois (LRF) function based on assessing the fit of several TPC functions (Lobry et al., 1991):

$$rate = Pmax \times \frac{(T - Tmax) \times (T - Tmin)^2}{(Topt - Tmin) \times \{(Topt - Tmin) \times (T - Topt) - (Topt - Tmax) \times (Topt - Tmin - 2T)\}}$$

We estimated the equation parameters *Topt*, *Pmax*, *Tmin*, and *Tmax* with site as a fixed effect on all parameters, sex as a group-level fixed effect (in effect, only affecting *Pmax*, or the height of the curve), and individual as a group-level random effect. We accounted for sex since females are considerably larger (Nufio et al., 2025). We log-transformed the LRF function and fit it using a lognormal error distribution (so sex and individual effects were proportional rather than additive). For *M. boulderensis*, we collected data over two years and accounted for year as a fixed effect. Since we improved methodology in the second year by drying and weighting feces more promptly, we used the (higher) 2nd year fecal masses to fit *Pmax* but allowed both years of data to inform the shape of the TPC. We used moderately informative, biologically grounded priors, identical between populations and species so as not to bias the results (*Tmin* = normal(10°C, (5°C)^2), Topt = normal(38°C, (5°C)^2), *Tmax*-*Topt* = lognormal(log(12), (log(2))^2) [°C], Pmax = normal (4mg/g/hr, (3mg/g/hr)^2). We compared maximum a posteriori (MAP) estimates and 90% high-density credible intervals of the parameter posterior distributions for different populations. The complex LRF function, which includes multiplicative interactions among parameters, results in broad credible intervals. Parameter differences may thus indicate meaningful biological differences despite overlapping parameter credible intervals, so we use the distinct MAP estimates for each population in our energy budget.

Body temperatures

We modeled climate using ERA5 (Hersbach et al., n.d.) and ERA5-Land (Muñoz Sabater, n.d.) hourly data on single levels, processed with the mcera5 package in R (Klinges et al., 2022). We examined the growing seasons (June through August) for historical (1950-1959) and contemporary (2015-2024) time periods. We used ERA5-Land (0.1° x 0.1° grids) for hourly air temperature, dewpoint temperature, surface pressure, windspeed, total precipitation, and downward solar radiation at the surface. For total cloud cover, total direct solar radiation, mean downward longwave radiation flux, and mean net longwave radiation flux, all of which weren't available from ERA-5 Land, we used the griddata() function from the scipy.interpolate python library to interpolate ERA5 data (0.25° x 0.25° grids) to match the spatial resolution of ERA5-Land (Virtanen et al., 2020). We used inverse distance weighting of the four nearest grid cells (weight ~ 1/distance) to further downscale all climate variables to site latitude and longitude (Klinges et al., 2022). We used the micro_era5() function from the mcera5 package (Klinges et al., 2022), which interfaces with the microclimate model in the R package NicheMapR (Kearney & Porter, 2017), to estimate air temperature, wind speed, solar radiation, and soil surface temperature at the four combinations of 0% and 80% shading and organismal heights of 0.01m and 0.15m. We additionally estimated an intermediate microclimate condition (30% shading and 0.03m). For all shading cases, we used unshaded air temperature (since our sites are meadows with plenty of air mixing), but we allowed the soil temperatures to be influenced by the level of shade. Aside from the above settings, we used the default settings of micro_era5().

We used the function Tb_grasshopper() from the TrenchR package (Buckley et al., 2023) to estimate body temperatures via a biophysical model as a function of the climate variables above, as well as additional microclimate, organismal, and substrate variables. This function calculates operative temperature by modeling the heat exchange between the grasshopper and the surrounding environment via convection, conduction, and radiation (Buckley et al., 2014). Inputs to this function include the zenith angle of the sun, clearness index (0.7), grasshopper solar absorptivity (0.7; Anderson et al., 1979) and length (0.0225m), substrate reflectivity (0.3), and the proportion of organismal surface area in contact with this substrate (0). We generated 5 possible body temperatures corresponding to the 5 different shade and height settings used in NicheMapR.

Energy budget

We estimated annual net energy gain for each species and sex at a given site by summing hourly estimates over the growing season (June-August). We estimated hourly body temperature both without (assuming intermediate microclimate: 30% shading and 0.03m) and with thermoregulation. For the case of thermoregulation, we assumed grasshoppers selected the body temperature that is closest to the species' preferred body temperature (*M. boulderensis:* 32.8°C, *M. sanguinipes*: 30.6°C; Buckley et al., 2014) from the range of available body temperatures (defined by the 5 body temperature scenarios representing the range of shade and height).

We estimated net energy gain as energy input from digestion minus the energetic cost of metabolism. We used our thermal performance curve fits to estimate feces production for an individual of a given sex, from a given population, and at a given body temperature. To convert grams of feces to energy gain in units of kilojoules, we assumed an assimilation rate (proportion of energy absorbed) of 0.4 and used values for the energy content of feces (12.8kJ/g) and wheatgrass (14.8kJ/g) (Harrison & Fewell, 1995).

We estimated resting metabolism using a previous parameterization of the equation $B \propto M^x e^{-E/kT_b}$, where M is the grasshopper's mass (g, averaged across population and sex), x is the (unitless) scaling coefficient, *E* is the apparent activation energy of biochemical reactions (eV) and *k* is the Boltzmann constant ($k = 8{\cdot}62 \times 10^{-5}$ eV K^{-1}) (Buckley et al., 2014). This equation was linearized by taking the natural log and then previously separately fit for each grasshopper species (with a coefficient for dependence on elevation, Table S2). Grasshoppers spend the vast majority of their time at rest (unpublished observations), so we did not account for the additional energetic costs of active metabolism. To convert metabolism in units of milliliters of carbon dioxide produced to energy expenditure in units of kilojoules, we assumed that grasshoppers were metabolizing lipids. We used a respiratory quotient of 0.7 to convert from CO2 produced to O2 consumed, a ratio of 2L O2 consumed for every 1g of lipids catabolized, and an energy density of 39kJ per gram of lipid, to calculate the expenditure in kilojoules (Lighton, 2008).

Analysis

We used R for analysis and Python to access and merge ERA5 and ERA5-Land data. The first author used a large language model (Claude by Anthropic) to help write code.

We simulated a reciprocal transplant by combining the climate conditions of one site with the population traits of another to determine how much the populations' net energy gains are driven by each component. In other words, our simulations represented the net energy gain of each population when exposed to each climate. Our "observations" were the total energy gain for each population at each climate (where the species occurs) at each of the 10 years in the contemporary time period. Each observation averaged the estimated net energy gains across males and females. To understand the direction of the effects of climate, population, and their interaction, we used the R lmer function in the Lme4 package to fit a linear mixed-effects (LME) model for each species, with year as a random effect (Bates et al., 2015). We used AIC to select a model that included the interaction term (ΔAICc = 60 for *M. boulderensis* and 78 for *M. sanguinipes*). We conducted a type II ANOVA of the output of this LME model to assess the magnitude and direction of the effects of site climate and population (Fox et al., 2001). We restrict our interpretation to the direction and magnitude of effects (ignoring statistical significance) since estimates are based upon simulated data without propagation of error. We also used hierarchical partitioning of variance to compare the proportion of variance explained by site versus population for each species (glmm.hp R package, Lai et al., 2022).

We also compared historical and contemporary net energy gains for populations at their home sites (assuming that historical physiology and body size are the same as our contemporary measurements). We compared net energy gain over time using LME models including time period, site, and the interaction of the two as fixed effects and year as a random effect. We also compared estimated marginal means from the historical and contemporary time periods for each site to understand the specific effect on each population's net energy gain (Lenth and Piaskowski, 2025). Once again, statistical significance for these models should not be interpreted.

Results

Our focal populations and species optimize digestion at high and similar temperatures (Fig. 1). Higher elevation populations of both species tend to exhibit higher peaks (*Pmaxes*) and lower thermal

maxima for performance (*Tmax*, Fig. 1), but the confidence intervals of parameter estimates overlap (Table S3-S4). In *M. boulderensis*, males feed more per gram of their body mass than females, whereas we do not see a significant difference in *Pmax* between males and females in *M. sanguinipes* (Table S3-S4). Our subsequent analyses utilize average TPC fits and energy gain across males and females. Individual variation in feces production is sizable at high temperatures (Fig. 1; Table S3-S4; *Pmax* individual SD = 0.417 for *M*. *boulderensis* and 0.334 for *M*. *sanguinipes*). Surprisingly, the *Tmax* estimate for the montane *M. boulderensis* species (~50°C) is higher than that for the thermal generalist *M. sanguinipes* (~46°C). Persistent digestion capacity even at the highest temperatures assayed limited our ability to accurately characterize *Tmax*.

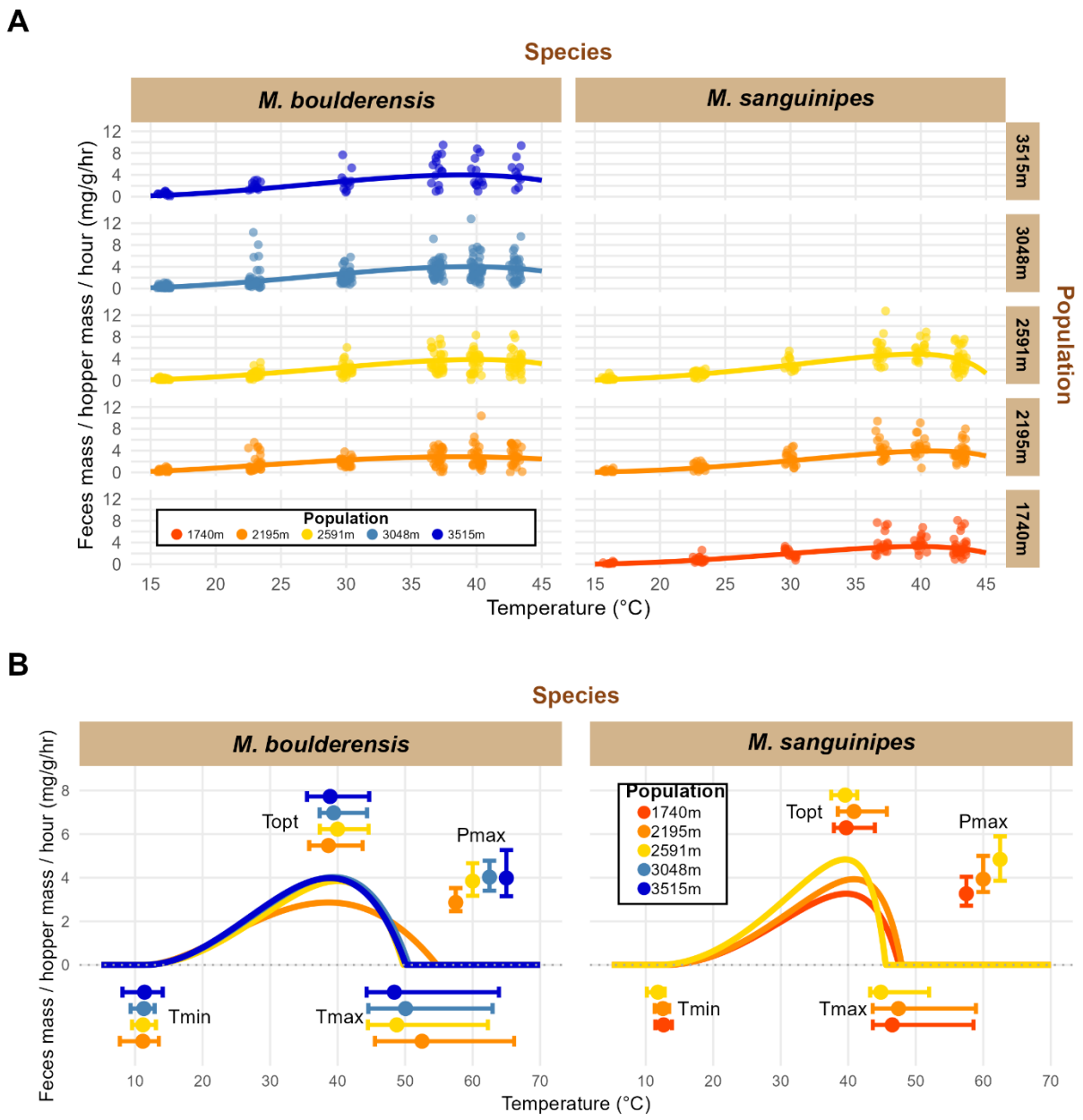

Fig. 1. *The thermal performance curves using the Lobry-Rosso-Flandrois equation showing (A) data for individuals and (B) parameter estimates (posterior means and 90% credible intervals). Colors indicate the populations at the different sites, and not all species are present at all sites.*

Digestion performance is optimized at quite high temperatures (~40°C), above the body temperatures grasshoppers regularly experience in the field (Fig. 2). Mean grasshopper body temperature estimates decrease with site elevation and increase from historical to contemporary time periods. The temperature-dependence of net energy gain is mainly driven by the digestion TPCs, with a small subtraction for resting metabolism, so energy gain is optimized at high temperatures. Populations at

low elevations (and contemporary time periods) more frequently encounter the high temperatures that maximize energy gain (Fig. 2).

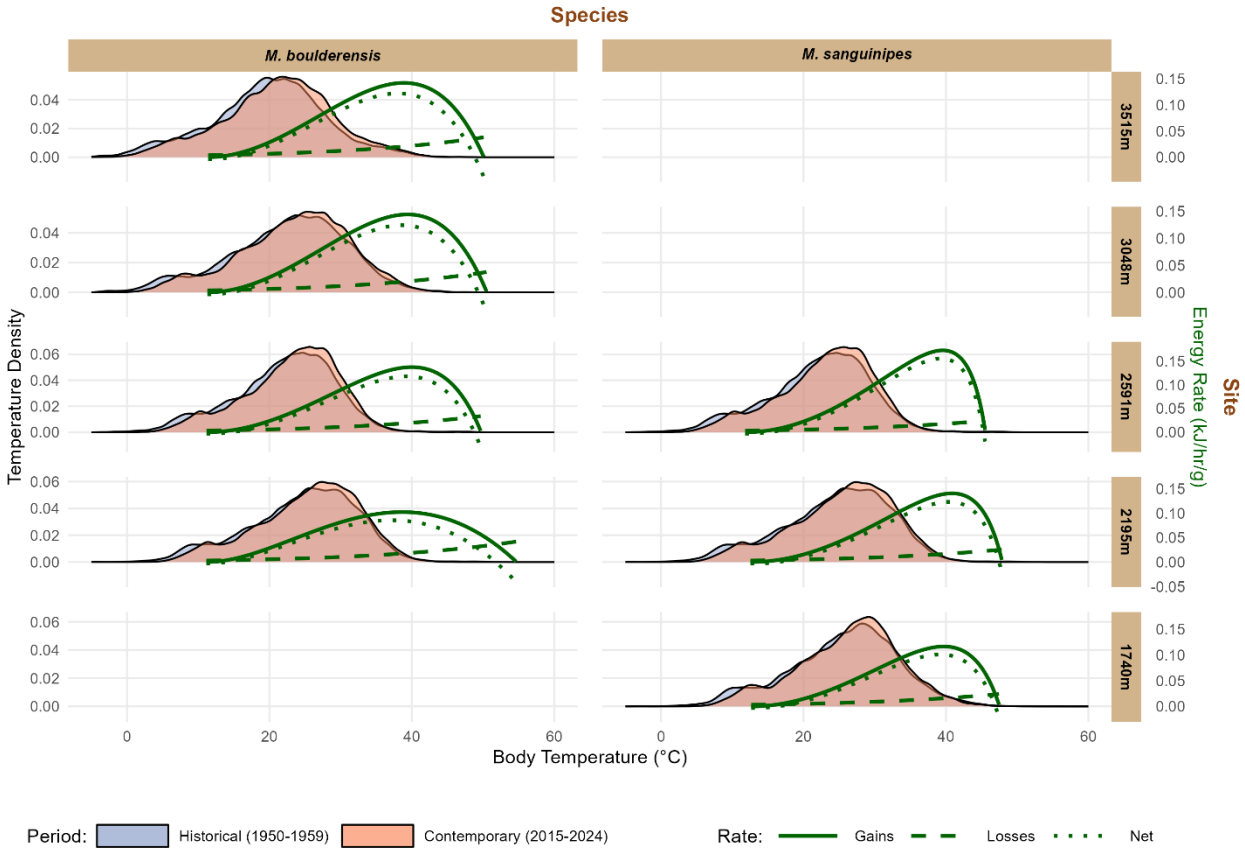

Fig. 2. *Daytime (sunrise to sunset) estimated body temperature distributions have shifted between the historical and contemporary time periods (blue and red). Overlaying digestion thermal performance curves (solid lines, Fig. 1), rates of energy use (dashed lines), and estimates of new energy gain (dotted lines) for each population indicate that energy gain is optimized at temperatures warmer than the body temperatures experienced.*

Our simulated reciprocal transplant shows that an individual's energy gain is driven by both physiological differences between the populations and by the climates these populations experience (Fig. 3, Table S5-56, $p<0.001$ for all statistical tests). The climate that maximizes energy yield for grasshoppers of all populations is the warmest one (the lowest elevation; *M. boulderensis*: $F_{3,135} = 2006.03$; *M. sanguinipes*: $F_{2,72} = 1956.79$). Within a given climate, however, TPC parameters for higher elevation populations maximize energy gain (*M. boulderensis*: $F_{3,135} = 544.63$; *M. sanguinipes*: $F_{2,72} = 3224.08$). This effect is particularly pronounced for *M. sanguinipes* and seems to be driven by differences in *Pmax*. Indeed, when we modify our energy budget to divide total energy gain through digestion by the area

under the TPC, much of the population effect disappears (Fig. S1). Differences in energy yields between populations are larger at warmer sites (*M. boulderensis*: $F_{9,135}$ = 10.80; *M. sanguinipes*: $F_{4,72}$ = 35.93). Trends remain similar when thermoregulation is modeled (Fig. S3). We use hierarchical partitioning of variance (without the interaction term) to understand the relative contribution of physiology and climate to energy balances for each species. For *M. boulderensis*, 79% of the variance is explained by climate, while the remaining 21% is explained by population physiology. For *M. sanguinipes*, 38% of the variance explained is due to climate, while the other 62% is due to population physiology.

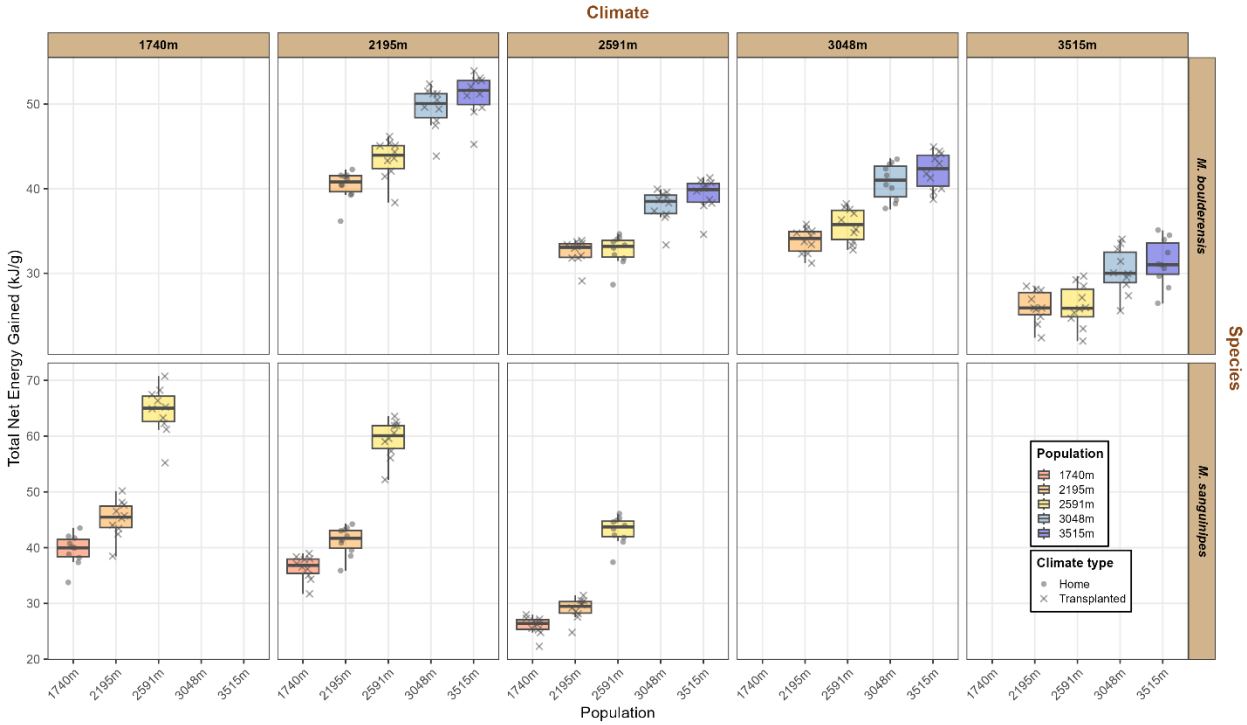

Fig. 3. *Estimated net energy gained for our simulated reciprocal transplant. We depict estimates for each contemporary year (2015-2024, points) along with 25, 50, and 75% quantiles and maxima and minima. Comparing populations in their home climates (dots as opposed to Xs), energy balances are similar across sites. This is due to opposing influences of population physiology (color, see comparisons within facet) and site climate (columns).*

Finally, we compare estimates of net energy gain for home populations in historical (1950-1959) and contemporary climates (2015-2024). Net energy gain has increased over 65 years of warming for all populations (Table S7; contrast ± SE *M. boulderensis*: 3.83 ± 0.825 kJ/g, $F_{1,18}$ = 21.52, p<0.001; *M. sanguinipes*: 3.89 ± 1.13 kJ/g, $F_{1,18}$ = 11.92, p=0.003), with the greatest increase occurring at higher elevation sites (Fig. 4, Table S8-S9, Fig. S2, S4) where warming has been most pronounced (Fig. 2).

Shifts in energy gain across time vary across sites for both species (Table S7-S9; *M. boulderensis*: $F_{3,54}$ = 2.96, p=0.040; *M. sanguinipes*: $F_{2,36}$ = 13.60, p<0.001).

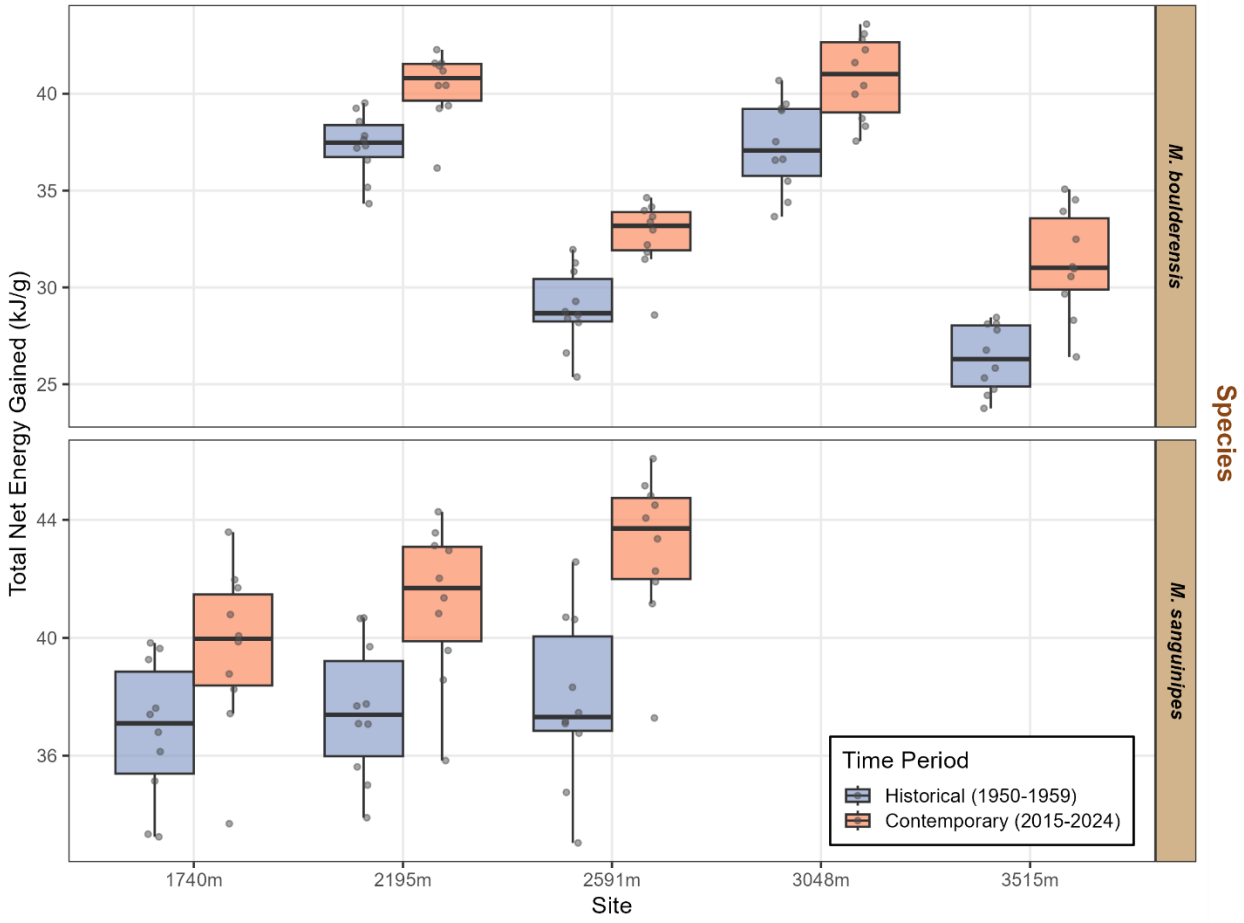

Fig. 4. *65 years of warming have increased estimated energy balances for all populations within their home sites. We depict estimates for each historical or contemporary year (dots) along with 25, 50, and 75% quantiles and maxima and minima.*

Discussion

Our study shows that digestion performance in both focal grasshopper species is optimized at relatively high body temperatures (~40°C), indicating that periods of warm conditions can strongly enhance energy acquisition (Fig. 1). Although such high body temperatures occur less frequently than moderate temperatures in these environments (Fig. 2), a high thermal optimum may allow grasshoppers to capitalize on brief episodes of thermal opportunity when they arise, consistent with theoretical expectations for ectotherms in seasonal environments (Angilletta et al., 2010). Population TPCs did not

differ significantly along the elevational gradient but we detected a trend in the magnitude of peak performance (*Pmax*). Contrary to our predictions, digestion thermal optima did not decline with elevation, providing little suggestion of local adaptation in thermal optima. *Pmax* differences among populations were more pronounced for *M. sanguinipes*. The finding was unexpected given the potential for local adaption in the dispersal-limited species, *M. boulderensis*, which exhibits more genetic differentiation (Slatyer et al., 2020).

Another unexpected result was the relatively high estimated upper thermal limit for digestion (*Tmax*) in the montane species *M. boulderensis* (~50°C) compared with the thermal generalist *M. sanguinipes* (~46 °C). This pattern contrasts with previous estimates of critical thermal maxima (*CTmax*) based on righting response, which tended to be slightly higher for *M. sanguinipes* (Buckley et al., 2014; Taylor, 2024). Our having few test temperatures above *Topt* limits the accuracy of our *Tmax* estimates. It is also possible that the high upper thermal limit for *M. boulderensis* digestion is real and adaptive – although mean and lower temperatures are lower at high elevation, grasshoppers there can reach higher thermal extremes due to absorption of higher levels solar radiation (Fig. 2; Buckley et al., 2013). A high *Tmax* may allow them to capitalize on this thermal opportunity. Preferred body temperature estimates in a thermal gradient were higher for *M. boulderensis* (Buckley et al., 2014), particularly at higher elevations, consistent with the montane species optimizing limited thermal opportunity.

The main trend in our TPCs is that peak performance (*Pmax*) is highest for high-elevation populations. This is consistent with past results for leaf area eaten for *M. boulderensis*, particularly for an additional 3739m population. However, the past study found the reverse elevation trend for *M. sanguinipes* (Buckley & Nufio, 2014). The discrepancy may be due to this study fitting the data to a different TPC equation, including higher temperatures, and examining feces production rather than leaf area eaten. Interestingly, the value of *Pmax* does not seem to trade off with thermal breadth, as in a specialist-generalist tradeoff; likewise, higher *Pmax* is not associated with higher thermal optima, as predicted by hotter-is-better (Angilletta, 2009).

Mass scaling could explain the observed *Pmax* trend. The performance metric in our TPCs is mass of feces / mass of grasshopper / time. Grasshoppers, like many insects in season-limited

environments, follow a reverse temperature size rule (Nufio et al., 2025) – so the populations at high elevations which have the largest area under the TPC are also the smallest. Perhaps digestion performance should not be expected to scale linearly with mass: the metabolic scaling exponent for our focal grasshoppers is ~0.9 (Buckley et al., 2014). However, assimilation tends to increase disproportionally with size through insect development (Maino & Kearney, 2015).

Another explanation is that selection is strongest on high-elevation populations: heat is a limited resource for these populations and efficient conversion of heat into energy through digestion is highly beneficial (Buckley et al., 2021A; Hodkinson, 2005). Assimilation rates tend to decline with increases in the duration of permissive temperatures (Levy et al., 2017). Thus, high-elevation populations with limited durations of thermally permissive conditions have the most to gain from increased capacity for energy assimilation.

Our energy budget estimates indicate that warm climates and high-elevation TPCs maximize net energy gain. Our simulated reciprocal transplant reveals that the effects of climate and population physiology are working in opposite directions. Taking both into account, the estimated net energy gains for populations along an elevation gradient are more or less stable (Fig. 3). This pattern is consistent with counter-gradient variation, in which environmental and physiological effects offset one another across spatial gradients (Conover et al., 2009; Pettersen, 2020; Yamahira et al., 2007). Notably, when we no longer scale by population body size (mass), *M. sanguinipes* exhibits counter-gradient variation to a lesser extent and *M. boulderensis* exhibits co-gradient variation, with environmental and physiological components both pointing in the same direction (Fig. S5-S6). Our hierarchical partitioning of variance indicates that the explanatory power of climate vs. physiology depends upon the species: for *M. boulderensis*, climate is far more influential, while for *M. sanguinipes*, physiology is somewhat more influential. It is important to note that the partitioning is approximate since it is based on variation in energy estimates among years without accounting for error in parameter estimates (Dietze et al., 2024).

Our comparison of historical and contemporary energy budgets suggests that warming has been beneficial for all populations' energy balances, especially at high elevations where the most warming has occurred (Fig. 4). Our analysis indicates that the potential for increased energy assimilation predominates

in determining energetic responses to climate change (Levy et al., 2017), but increased metabolic costs and food limitations can be limiting factors for energy gain in other systems (Huey & Kingsolver, 2019). The capacity of grasshoppers to assimilate at high temperatures results in the potential for climate warming to increase net energy gain even in warm environments, increasing the chance of outbreaks that are detrimental to vegetation (Youngblood et al., 2023). However, water limitations can reduce thermal tolerance and limit the potential to capitalize on warming (Youngblood et al., 2025). Nutrient density can sometimes limit opportunities for energy gain, but responses to nutrient density vary widely among grasshoppers (Cease, 2024).

Conclusion

Although this study finds that warming provides thermal opportunity for increased net energy gain, past assessments of the effect of climate change on the same populations suggest mixed impacts of climate change on these grasshoppers. Resurveys show that from 1959-1960 to 2006-2008, abundances of these populations have either decreased or, in the case of high-elevation *M. boulderensis*, remained the same (Buckley et al., 2014; Nufio et al., 2010). Analyzing all abundance data (1958-1960 and 2006-2015), Buckley et al. (2021B) found that total abundance declines or stays the same for our populations in years with a higher number of growing degree days. Past estimates of these populations' thermal sensitivity also indicated that climate change has increased digestion performance, but indicated decrease in locomotion performance (which is optimized at ~25°C) with warming (Buckley & Nufio, 2014). The increased energy gain with warming may be accompanied by thermal stress or survival decreases that prevent abundance increases with warming (Buckley et al., 2021A). Factors such as seasonal timing and elevation may determine whether climate change poses an opportunity or a stress. In our focal grasshopper system, early season species and higher elevation populations have generally increased body size over recent decades while later season and lower elevation species have declined in size (Nufio et al. 2025).

Energetics is a promising lens for understanding organismal impacts of climate change. However, additional biological mechanisms are also important to consider, such as survival in the face of thermal extremes and shifting biotic interactions (Urban et al., 2016). Behavior is also an important mediator and buffer. All of these avenues should be pursued and ideally integrated. Understanding how different

biological mechanisms interact and which are the most salient in a given system or population may be crucial for predicting organismal impacts of climate change.

Acknowledgements

I would like to thank Monica Sheffer, César Nufio, Caroline Williams, Sean Schoville, and Lauren Buckley for their substantial contributions to this project.

I would like to thank Anna Brasket, Jared Haar, Lucie Reizian, Marcos Alvarez, Varun Krishnakumar, Alexander Juan, and Hannah Woods for help with the digestion thermal performance curve measurements. I would like to thank Taylor Hatcher, Spike Pike, Madeleine Strait, Simran Bawa, Michael Troutman, Berry Brosi, Elli Theobald, Emily Carrington, and Andrew Berdahl for feedback on the project and/or manuscript.

Data and Code Accessibility Statement

Code and data can be accessed at https://github.com/juliamsmith/EB_Ch1.git. We omitted large climate and microclimate data files but provide the code for downloading and generating the data.

Bibliography

Alexander, G., & Hilliard, J. R. (1969). Altitudinal and Seasonal Distribution of Orthoptera in the Rocky Mountains of Northern Colorado. *Ecological Monographs*, *39*(4), 385–432. https://doi.org/10.2307/1942354

Amarasekare, P., & Savage, V. (2012). A Framework for Elucidating the Temperature Dependence of Fitness. *The American Naturalist*, *179*(2), 178–191. https://doi.org/10.1086/663677

Angilletta, M. J. (2001). Thermal and Physiological Constraints on Energy Assimilation in a Widespread Lizard (sceloporus Undulatus). *Ecology*, *82*(11), 3044–3056. https://doi.org/10.1890/0012-9658(2001)082%255B3044:TAPCOE%255D2.0.CO;2

Angilletta, M. J. (2009). *Thermal adaptation: A theoretical and empirical synthesis*. https://books.google.com/books?hl=en&lr=&id=5RRREAAAQBAJ&oi=fnd&pg=PR7&dq=info:0o39D7IFBv0J:scholar.google.com&ots=jM_CrwThjf&sig=sC6I_QL1Cr0PhbpbfXC-X0tugI8

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Belovsky, G. E., & Slade, J. B. (2000). Insect herbivory accelerates nutrient cycling and increases plant production. *Proceedings of the National Academy of Sciences*, *97*(26), 14412–14417. https://doi.org/10.1073/pnas.250483797

Briscoe, N. J., Morris, S. D., Mathewson, P. D., Buckley, L. B., Jusup, M., Levy, O., Maclean, I. M. D., Pincebourde, S., Riddell, E. A., Roberts, J. A., Schouten, R., Sears, M. W., & Kearney, M. R. (2023). Mechanistic forecasts of species responses to climate change: The promise of biophysical ecology. *Global Change Biology*, *29*(6), 1451–1470. https://doi.org/10.1111/gcb.16557

Buckley, L. B., Graham, S. I., & Nufio, C. R. (2021). Grasshopper species' seasonal timing underlies shifts in phenological overlap in response to climate gradients, variability and change. *Journal of Animal Ecology*, *90*(5), 1252–1263. https://doi.org/10.1111/1365-2656.13451

Buckley, L. B., & Kingsolver, J. G. (2021). Evolution of Thermal Sensitivity in Changing and Variable Climates. *Annual Review of Ecology, Evolution, and Systematics*, *52*(Volume 52, 2021), 563–586. https://doi.org/10.1146/annurev-ecolsys-011521-102856

Buckley, L. B., Miller, E. F., & Kingsolver, J. G. (2013). Ectotherm Thermal Stress and Specialization Across Altitude and Latitude. *Integrative and Comparative Biology*, *53*(4), 571–581. https://doi.org/10.1093/icb/ict026

Buckley, L. B., & Nufio, C. R. (2014). Elevational clines in the temperature dependence of insect performance and implications for ecological responses to climate change. *Conservation Physiology*, *2*(1), cou035. https://doi.org/10.1093/conphys/cou035

Buckley, L. B., Nufio, C. R., & Kingsolver, J. G. (2014). Phenotypic clines, energy balances and ecological responses to climate change. *Journal of Animal Ecology*, *83*(1), 41–50. https://doi.org/10.1111/1365-2656.12083

Buckley, L. B., Ortiz, B. A. B., Caruso, I., John, A., Levy, O., Meyer, A. V., Riddell, E. A., Sakairi, Y., & Simonis, J. L. (2023). TrenchR: An R package for modular and accessible microclimate and biophysical ecology. *PLOS Climate*, *2*(8), e0000139. https://doi.org/10.1371/journal.pclm.0000139

Buckley, L. B., Schoville, S. D., & Williams, C. M. (2021). Shifts in the relative fitness contributions of fecundity and survival in variable and changing environments. *Journal of Experimental Biology*, *224*(Suppl_1), jeb228031. https://doi.org/10.1242/jeb.228031

Cease, A. J. (2024). How Nutrients Mediate the Impacts of Global Change on Locust Outbreaks. *Annual Review of Entomology*, *69*(Volume 69, 2024), 527–550. https://doi.org/10.1146/annurev-ento-120220-110415

Coggan, N., Clissold, F. J., & Simpson, S. J. (2011). Locusts use dynamic thermoregulatory behaviour to optimize nutritional outcomes. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2745–2752. https://doi.org/10.1098/rspb.2010.2675

Conover, D. O., Duffy, T. A., & Hice, L. A. (2009). The Covariance between Genetic and Environmental Influences across Ecological Gradients. *Annals of the New York Academy of Sciences*, *1168*(1), 100–129. https://doi.org/10.1111/j.1749-6632.2009.04575.x

Dietze, M., White, E. P., Abeyta, A., Boettiger, C., Bueno Watts, N., Carey, C. C., Chaplin-Kramer, R., Emanuel, R. E., Ernest, S. K. M., Figueiredo, R. J., Gerst, M. D., Johnson, L. R., Kenney, M. A., McLachlan, J. S., Paschalidis, I. C., Peters, J. A., Rollinson, C. R., Simonis, J., Sullivan-Wiley, K., … Zwart, J. (2024). Near-term ecological forecasting for climate change action. *Nature Climate Change*, *14*(12), 1236–1244. https://doi.org/10.1038/s41558-024-02182-0

Evans, T. G., Diamond, S. E., & Kelly, M. W. (2015). Mechanistic species distribution modelling as a link between physiology and conservation. *Conservation Physiology*, *3*(1), cov056. https://doi.org/10.1093/conphys/cov056

Fox, J., Weisberg, S., & Price, B. (2001). *car: Companion to Applied Regression* (p. 3.1-3) [Dataset]. https://doi.org/10.32614/CRAN.package.car

Harrison, J. F., & Fewell, J. H. (1995). Thermal Effects on Feeding Behavior and Net Energy Intake in a Grasshopper Experiencing Large Diurnal Fluctuations in Body Temperature. *Physiological Zoology*, *68*(3), 453–473. https://doi.org/10.1086/physzool.68.3.30163779

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J.-N. (n.d.). *ERA5 hourly*

data on single levels from 1940 to present. Retrieved November 16, 2025, from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview

Hodkinson, I. D. (2005). Terrestrial insects along elevation gradients: Species and community responses to altitude. *Biological Reviews*, *80*(3), 489–513. https://doi.org/10.1017/S1464793105006767

Huey, R. B., & Kingsolver, J. G. (1989). Evolution of thermal sensitivity of ectotherm performance. *Trends in Ecology & Evolution*, *4*(5), 131–135. https://doi.org/10.1016/0169-5347(89)90211-5

Huey, R. B., & Kingsolver, J. G. (2019). Climate Warming, Resource Availability, and the Metabolic Meltdown of Ectotherms. *The American Naturalist*, *194*(6), E140–E150. https://doi.org/10.1086/705679

Kearney, M. R., & Porter, W. P. (2017). NicheMapR - an R package for biophysical modelling: The microclimate model. *Ecography*, *40*(5), 664–674. https://doi.org/10.1111/ecog.02360

Kingsolver, J. G. (2009). The Well-Temperatured Biologist. *The American Naturalist*, *174*(6), 755–768. https://doi.org/10.1086/648310

Klinges, D. H., Duffy, J. P., Kearney, M. R., & Maclean, I. M. D. (2022). mcera5: Driving microclimate models with ERA5 global gridded climate data. *Methods in Ecology and Evolution*, *13*(7), 1402–1411. https://doi.org/10.1111/2041-210X.13877

Lai, J., Zou, Y., Zhang, S., Zhang, X., & Mao, L. (2022). glmm.hp: An R package for computing individual effect of predictors in generalized linear mixed models. *Journal of Plant Ecology*, *15*(6), 1302–1307. https://doi.org/10.1093/jpe/rtac096

Levy, O., Borchert, J. D., Rusch, T. W., Buckley, L. B., & Angilletta Jr., M. J. (2017). Diminishing returns limit energetic costs of climate change. *Ecology*, *98*(5), 1217–1228. https://doi.org/10.1002/ecy.1803

Levy, O., Buckley, L. B., Keitt, T. H., Smith, C. D., Boateng, K. O., Kumar, D. S., & Angilletta, M. J. (2015). Resolving the life cycle alters expected impacts of climate change. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1813), 20150837. https://doi.org/10.1098/rspb.2015.0837

Lobry, J. R., Rosso, L., & Flandrois, J.-P. (1991). A FORTRAN subroutine for the determination of parameter confidence limits in non-linear Models. *Binary*, *3*, 86–93.

Logan, M. L., & Cox, C. L. (2020). Genetic Constraints, Transcriptome Plasticity, and the Evolutionary Response to Climate Change. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.538226

Maino, J. L., & Kearney, M. R. (2015). Testing mechanistic models of growth in insects. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1819), 20151973. https://doi.org/10.1098/rspb.2015.1973

Muñoz Sabater, J. (n.d.). *ERA5-Land hourly data from 1950 to present*. Retrieved November 16, 2025, from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview

Nufio, C. R., McGuire, C. R., Bowers, M. D., & Guralnick, R. P. (2010). Grasshopper Community Response to Climatic Change: Variation Along an Elevational Gradient. *PLOS ONE*, *5*(9), e12977. https://doi.org/10.1371/journal.pone.0012977

Nufio, C. R., Sheffer, M. M., Smith, J. M., Troutman, M. T., Bawa, S. J., Taylor, E. D., Schoville, S. D., Williams, C. M., & Buckley, L. B. (2025). Insect size responses to climate change vary across elevations according to seasonal timing. *PLOS Biology*, *23*(1), e3002805. https://doi.org/10.1371/journal.pbio.3002805

Pettersen, A. K. (2020). Countergradient Variation in Reptiles: Thermal Sensitivity of Developmental and Metabolic Rates Across Locally Adapted Populations. *Frontiers in Physiology*, *11*. https://doi.org/10.3389/fphys.2020.00547

Scheffers, B. R., Edwards, D. P., Diesmos, A., Williams, S. E., & Evans, T. A. (2014). Microhabitats reduce animal's exposure to climate extremes. *Global Change Biology*, *20*(2), 495–503. https://doi.org/10.1111/gcb.12439

Siemers, I., von Schmalensee, L., Ziesemer, S., Gotthard, K., & Lehmann, P. (2024). Predictable local adaptation in butterfly photoperiodism but not thermal performance along a latitudinal cline. *Ecosphere*, *15*(5), e4842. https://doi.org/10.1002/ecs2.4842

Sinclair, B. J., Marshall, K. E., Sewell, M. A., Levesque, D. L., Willett, C. S., Slotsbo, S., Dong, Y., Harley, C. D. G., Marshall, D. J., Helmuth, B. S., & Huey, R. B. (2016). Can we predict ectotherm responses to climate change using thermal performance curves and body temperatures? *Ecology Letters*, *19*(11), 1372–1385. https://doi.org/10.1111/ele.12686

Slatyer, R. A., Schoville, S. D., Nufio, C. R., & Buckley, L. B. (2020). Do different rates of gene flow underlie variation in phenotypic and phenological clines in a montane grasshopper community? *Ecology and Evolution*, *10*(2), 980–997. https://doi.org/10.1002/ece3.5961

Srivastava, D. S., Coristine, L., Angert, A. L., Bontrager, M., Amundrud, S. L., Williams, J. L., Yeung, A. C. Y., de Zwaan, D. R., Thompson, P. L., Aitken, S. N., Sunday, J. M., O'Connor, M. I., Whitton, J., Brown, N. E. M., MacLeod, C. D., Parfrey, L. W., Bernhardt, J. R., Carrillo, J., Harley, C. D. G., … Donner, S. D. (2021). Wildcards in climate change biology. *Ecological Monographs*, *91*(4), e01471. https://doi.org/10.1002/ecm.1471

Stark, G., Ma, L., Zeng, Z.-G., Du, W.-G., & Levy, O. (2023). Cool shade and not-so-cool shade: How habitat loss may accelerate thermal stress under current and future climate. *Global Change Biology*, *29*(22), 6201–6216. https://doi.org/10.1111/gcb.16802

Taylor, E. (2024). *Thermal Adaptation and Plasticity in Melanoplus Grasshoppers: Elevated Temperatures and Heat-Responsive Genes* [Thesis]. https://minds.wisconsin.edu/handle/1793/85808

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Pe'er, G., Singer, A., Bridle, J. R., Crozier, L. G., De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., … Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, *353*(6304), aad8466. https://doi.org/10.1126/science.aad8466

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Williams, J. W., & Jackson, S. T. (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, *5*(9), 475–482. https://doi.org/10.1890/070037

Yamahira, K., Kawajiri, M., Takeshi, K., & Irie, T. (2007). Inter- and Intrapopulation Variation in Thermal Reaction Norms for Growth Rate: Evolution of Latitudinal Compensation in Ectotherms with a Genetic Constraint. *Evolution*, *61*(7), 1577–1589.

Youngblood, J. P., Cease, A. J., Talal, S., Copa, F., Medina, H. E., Rojas, J. E., Trumper, E. V., Angilletta Jr., M. J., & Harrison, J. F. (2023). Climate change expected to improve digestive rate and trigger range expansion in outbreaking locusts. *Ecological Monographs*, *93*(1), e1550. https://doi.org/10.1002/ecm.1550

Youngblood, J. P., Overgaard, J., Ørsted, M., Angilletta Jr., M. J., & VandenBrooks, J. M. (2025). Dehydration worsens heat tolerance of locusts and amplifies predicted impacts of climate change. *Functional Ecology*, *39*(5), 1194–1207. https://doi.org/10.1111/1365-2435.70038

Supplement

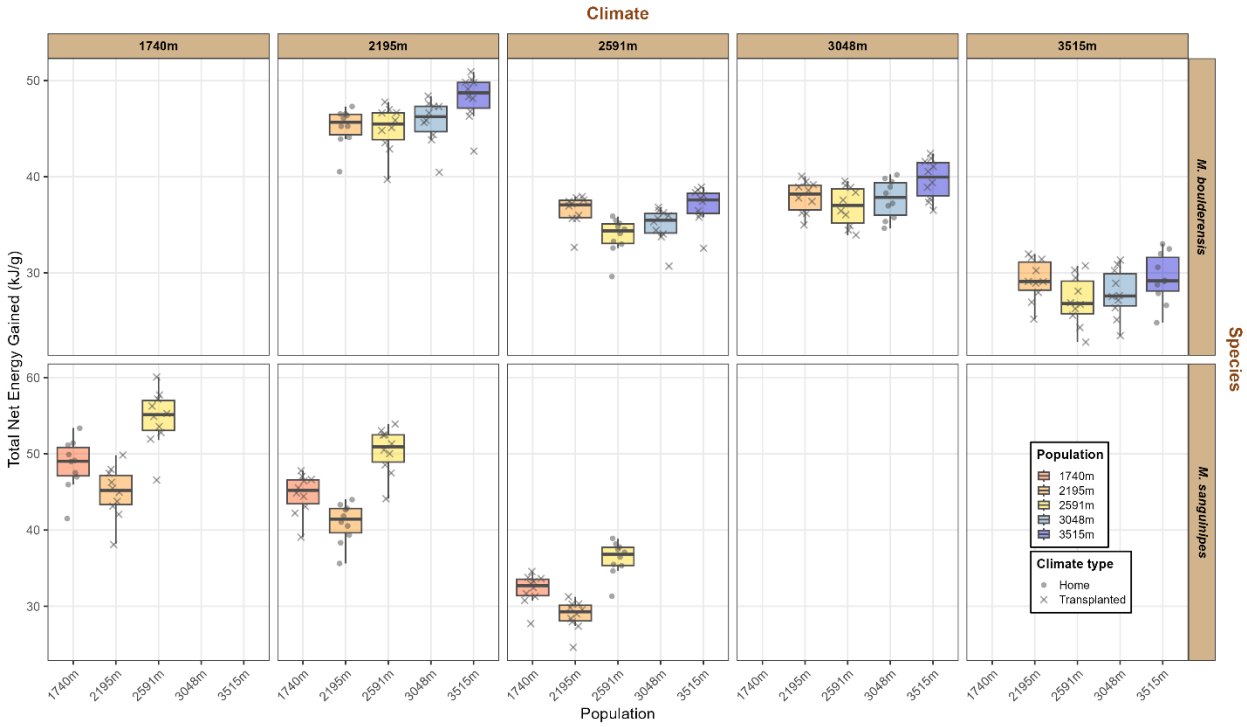

Fig. S1. *Estimated net energy gained for our simulated reciprocal transplant, corrected for area under the thermal performance curve. (Akin to Fig. 3 in the text.) We depict estimates for each contemporary year (points) along with 25, 50, and 75% quantiles and maxima and minima. Comparing populations in their home climates (dots as opposed to Xs), energy balances tend to decline in colder sites. After accounting for differences in Pmax, physiology (see comparisons within facet) has little effect on energy balance.*

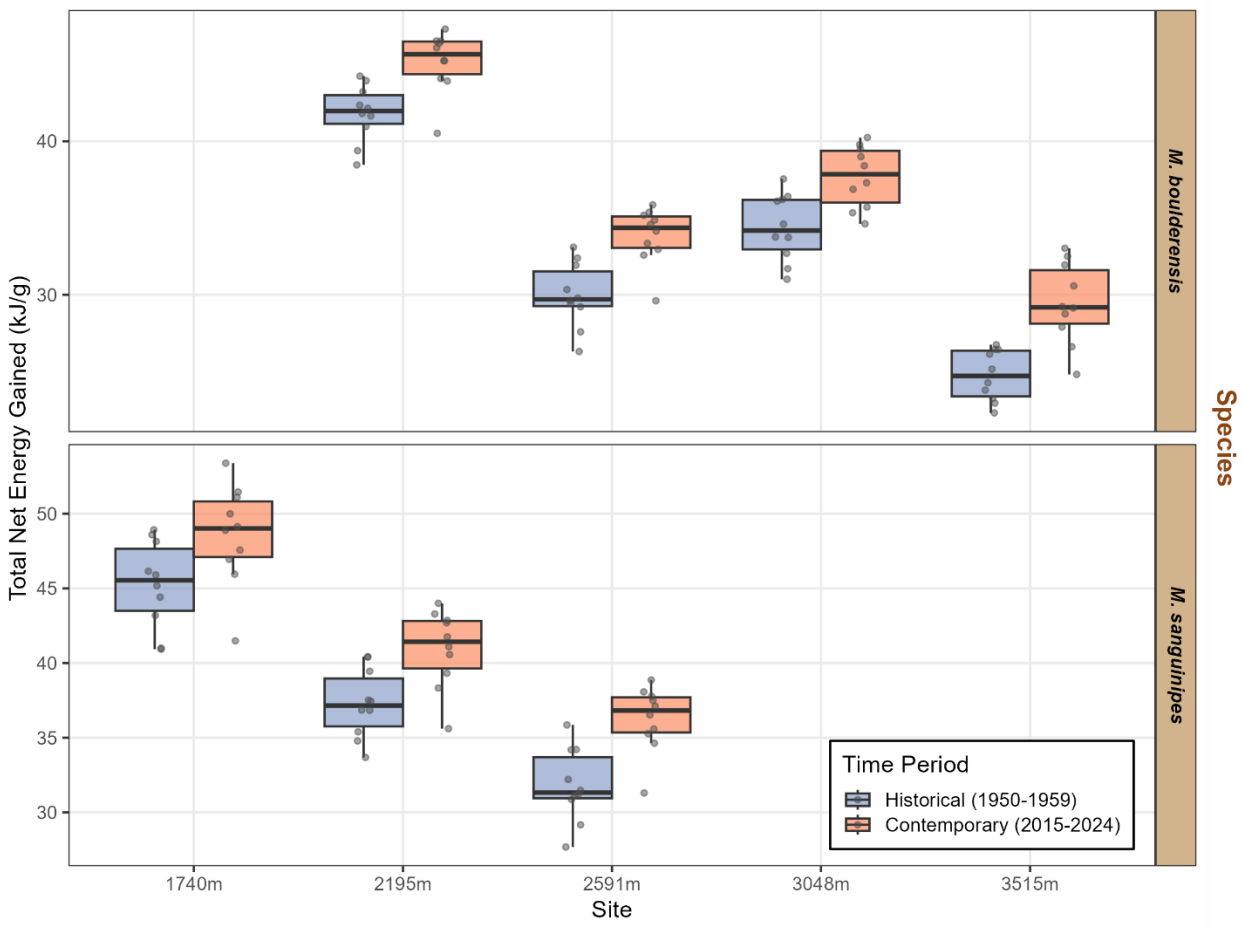

Fig. S2. *65 years of warming have increased estimated energy balances for all populations within their home sites (like Fig. 4 in text but corrected for area under the TPC). We depict estimates for each historical or contemporary year (dots) along with 25, 50, and 75% quantiles and maxima and minima. After accounting for differences in Pmax, we see that higher-elevation populations tend to gain less energy on net.*

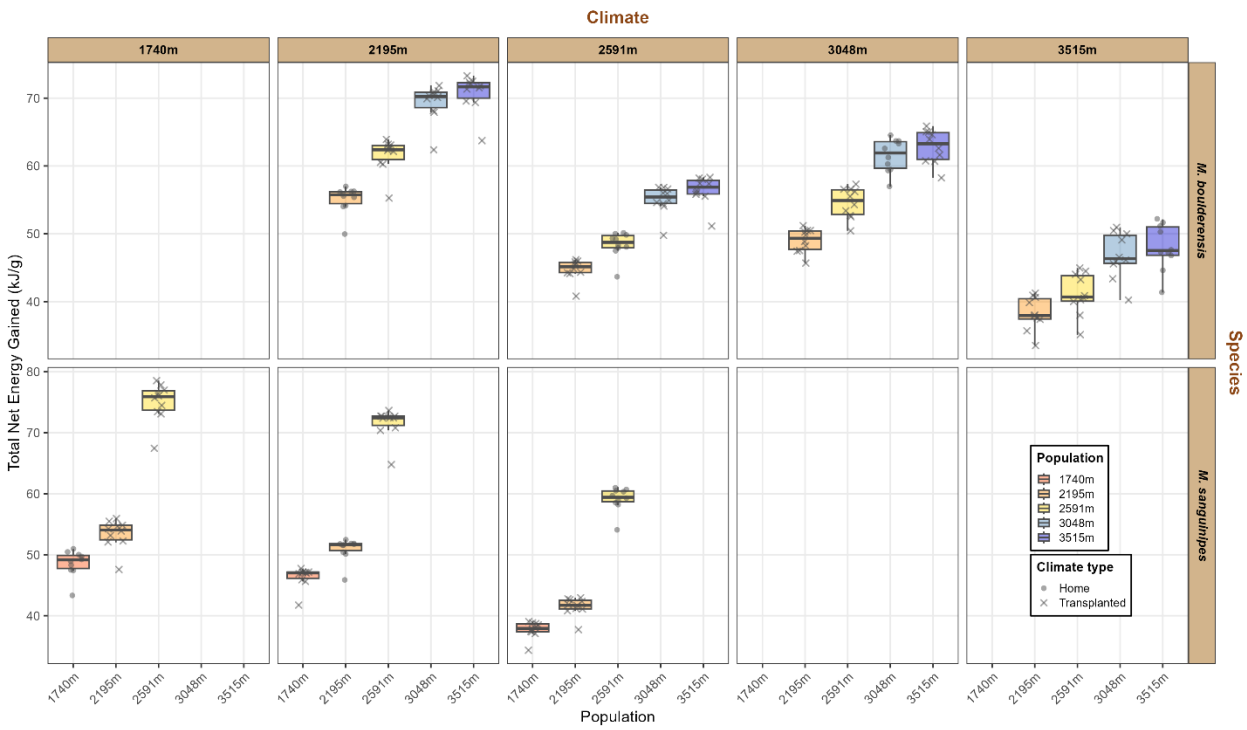

Fig. S3. *Estimated net energy gained for our simulated reciprocal transplant allowing for behavioral thermoregulation toward preferred body temperature. (Akin to Fig. 3 in the text.) We depict estimates for each contemporary year (points) along with 25, 50, and 75% quantiles and maxima and minima. Comparing populations in their home climates (dots as opposed to Xs), energy balances are similar across sites. This is due to opposing influences of population physiology (color, see comparisons within facet) and site climate (columns).*

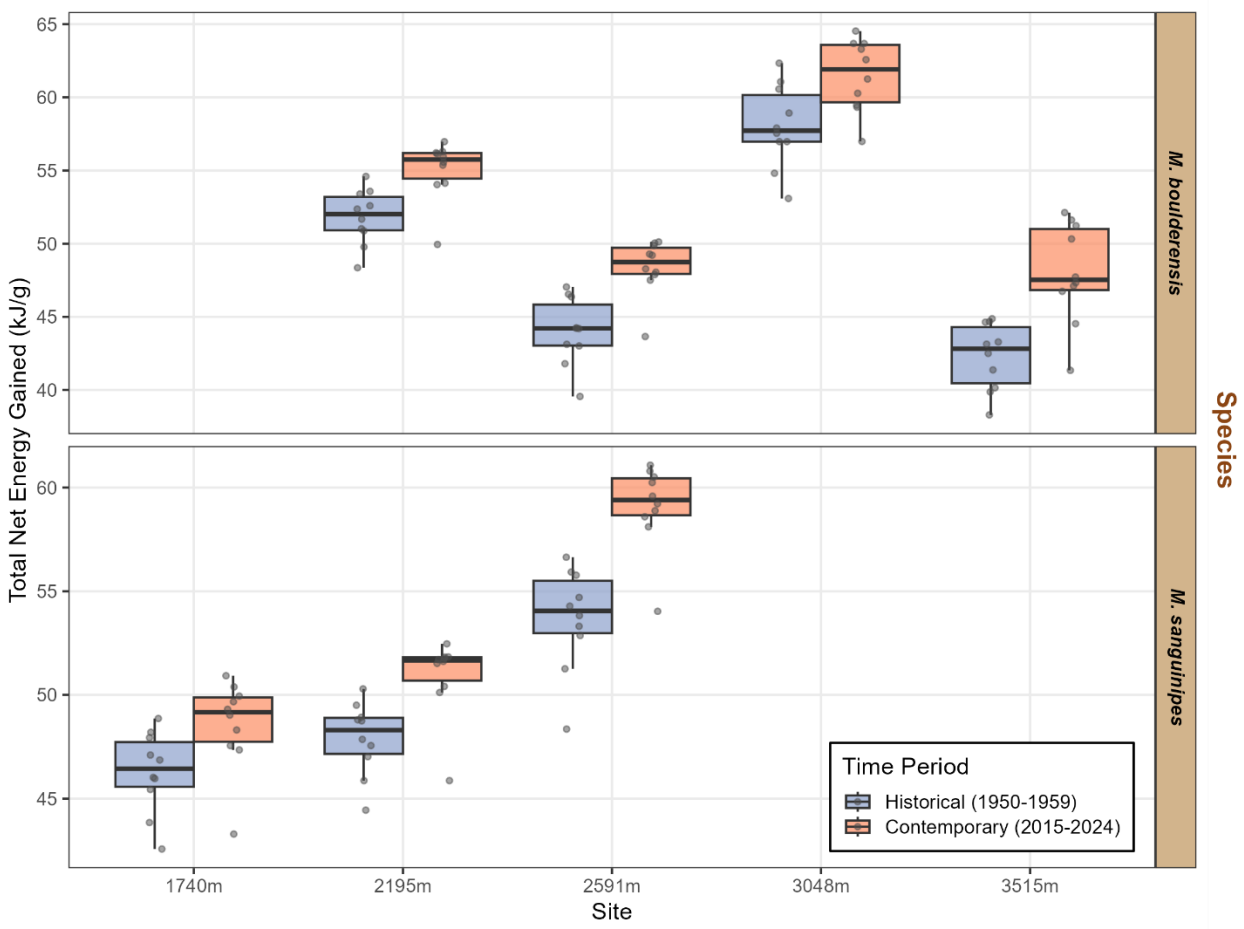

Fig. S4. *65 years of warming have increased estimated energy balances for all populations within their home sites (like Fig. 4 in text but allowing behavioral thermoregulation). We depict estimates for each historical or contemporary year (dots) along with 25, 50, and 75% quantiles and maxima and minima.*

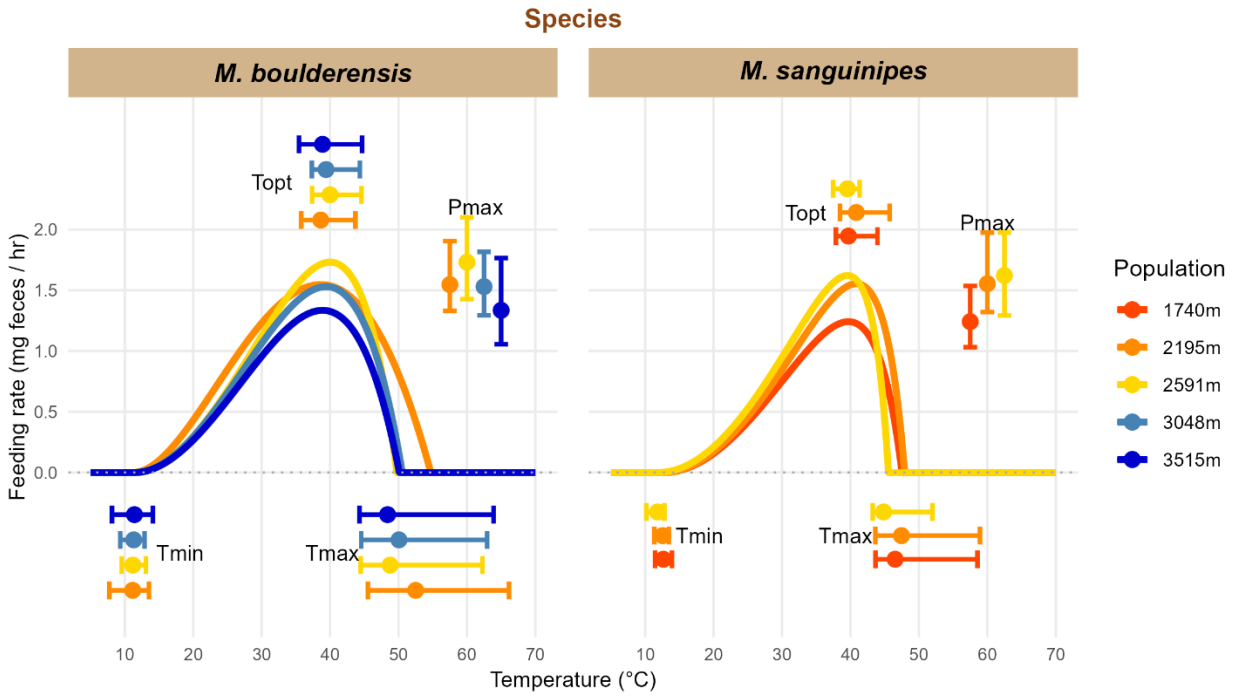

Fig. S5. *The thermal performance curves using the Lobry-Rosso-Flandrois equation showing parameter estimates (posterior means and 90% credible intervals) without mass scaling (Akin to Fig. 1 in the text but multiplied by the average population masses). Colors indicate the populations at the different sites, and not all species are present at all sites.*

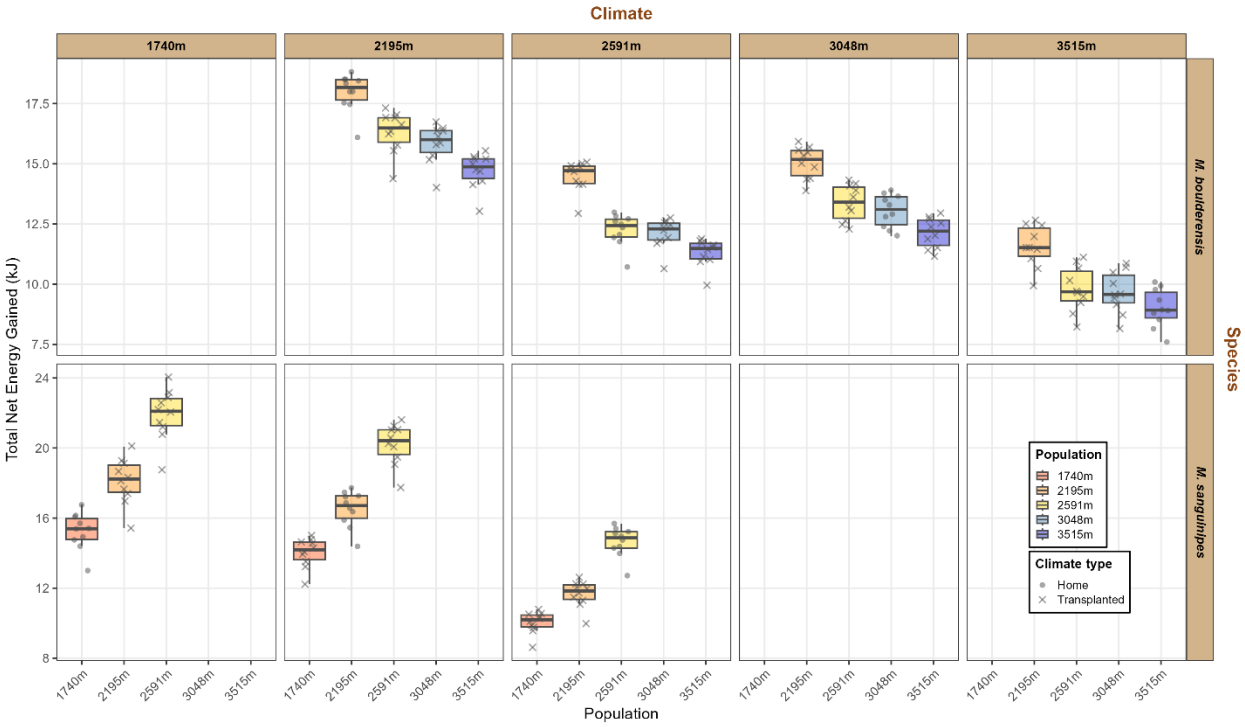

Fig. S6. *Estimated net energy gained for our simulated reciprocal transplant without mass scaling. (Akin to Fig. 3 in the text.) We depict estimates for each contemporary year (points) along with 25, 50, and 75%*

quantiles and maxima and minima. Comparing populations in their home climates (dots as opposed to Xs), energy balances are similar across sites for M. sanguinipes and decline at higher elevations for M. boulderensis. The trend in population physiology (color, see comparisons within facet) is reduced in M. sanguinipes and reversed in M. boulderensis when we do not scale by mass. In the latter case, we now see co-gradient variation, in which population physiology and site climate (columns) trend in the same direction.

Table S1. *The sites along an elevational gradient in Colorado where we collected grasshoppers to measure their digestion thermal performance curves.*

Elevation	Coordinates	Species	Number of individuals (males / females)
1,740m	40.00N, 105.28W	*M. sanguinipes*	19 / 17
2,195m	40.01N, 105.37W	*M. sanguinipes*	18 / 19
		M. boulderensis	27 / 32
2,591m	40.02N, 105.43W	*M. sanguinipes*	20 / 15
		M. boulderensis	32 / 32
3,048m	40.03N, 105.55W	*M. boulderensis*	40 / 37
3,515m	40.06N, 105.62W	*M. boulderensis*	10 / 20

Table S2. *Past data from (Buckley et al., 2014) on temperature-dependence of resting metabolic rate and on preferred body temperature. log(rmr) = b0 + b1*log_mass + b2*1/(k*Tb) + b3*elevation, where k= 8·62 × 10^{-5} eV K^{-1} and Tb is in Kelvin.*

Species	b0	b1	b2	b3	PBT
M. boulderensis	17.1	0.98	-0.48	0.000124	32.83
M. sanguinipes	16.5	0.85	-0.47	0.000108	30.63

Table S3. *TPC parameter estimates for M. boulderensis with 90% credible intervals.*

Population	T_{min} (°C)	T_{opt} (°C)	T_{max} (°C)	P_{max} (mg/g/hr)
2195m	11.14 (7.74, 13.51)	38.64 (35.78, 43.70)	52.50 (45.52, 66.15)	2.865 (2.464, 3.527)
2591m	11.17 (9.54, 13.08)	39.99 (37.37, 44.59)	48.76 (44.48, 62.26)	3.848 (3.172, 4.666)
3048m	11.30 (9.33, 12.88)	39.41 (37.33, 44.35)	50.06 (44.55, 62.95)	4.026 (3.405, 4.780)
3515m	11.40 (8.12, 14.09)	38.89 (35.46, 44.67)	48.40 (44.30, 63.90)	3.985 (3.153, 5.270)
Fixed Effects				
Sex effect				0.116 (-0.028, 0.255)
Year effect				-0.583 (-0.772, -0.449)
Random Effects				
Individual SD				0.417 (0.314, 0.496)

Table S4. *TPC parameter estimates for M. sanguinipes with 90% credible intervals.*

Population	T_{min} (°C)	T_{opt} (°C)	T_{max} (°C)	P_{max} (mg/g/hr)
1740m	12.65 (11.43, 13.90)	39.70 (37.87, 43.94)	46.52 (43.65, 58.56)	3.268 (2.715, 4.042)
2195m	12.54 (11.33, 13.46)	40.85 (38.47, 45.72)	47.45 (43.63, 58.92)	3.935 (3.345, 5.000)
2591m	11.80 (10.16, 12.79)	39.55 (37.48, 41.34)	44.84 (43.25, 51.99)	4.840 (3.858, 5.897)
Fixed Effects				
Sex effect				-0.415 (-0.587, -0.241)
Random Effects				
Individual SD				0.334 (0.221, 0.454)

Table S5. *Output of Type II ANOVA on our simulated reciprocal transplant regression.*

	M. boulderensis				***M. sanguinipes***			
	num df	den df	F	p	num df	den df	F	p
Climate	3	135	2006.03	< 0.001	2	72	1956.79	< 0.001
Population	3	135	544.63	< 0.001	2	72	3224.08	< 0.001
Population:Climate	9	135	10.80	< 0.001	4	72	35.93	< 0.001

Table S6. *Coefficients from the same simulated reciprocal transplant regression model. The dependent variable is total energy gained adjusted for body size (kJ/g).*

	M. boulderensis				*M. sanguinipes*			
	Coefficient	**SE**	**t**	**p**	**Coefficient**	**SE**	**t**	**p**
(Intercept)	40.368	0.686	58.845	< 0.001	41.210	0.908	45.376	< 0.001
clim_1740m				—	3.971	0.493	8.052	< 0.001
pop_1740m				—	-4.797	0.493	-9.727	< 0.001
clim_2591m	-7.825	0.460	-16.999	< 0.001	-12.161	0.493	-24.660	< 0.001
pop_2591m	3.119	0.460	6.774	< 0.001	18.254	0.493	37.016	< 0.001
clim_3048m	-6.547	0.460	-14.223	< 0.001				—
pop_3048m	9.162	0.460	19.903	< 0.001				—
clim_3515m	-14.341	0.460	-31.154	< 0.001				—
pop_3515m	10.703	0.460	23.250	< 0.001				—
pop_1740m:clim_1740m				—	-0.771	0.697	-1.105	0.273
pop_1740m:clim_2591m				—	1.758	0.697	2.520	0.014
pop_2591m:clim_1740m				—	1.034	0.697	1.482	0.143
pop_2591m:clim_2591m	-2.980	0.651	-4.577	< 0.001	-4.245	0.697	-6.087	< 0.001
pop_2591m:clim_3048m	-1.268	0.651	-1.948	0.053				—
pop_2591m:clim_3515m	-2.979	0.651	-4.576	< 0.001				—
pop_3048m:clim_2591m	-3.795	0.651	-5.830	< 0.001				—
pop_3515m:clim_2591m	-3.976	0.651	-6.107	< 0.001				—
pop_3048m:clim_3048m	-2.149	0.651	-3.300	0.001				—
pop_3048m:clim_3515m	-4.894	0.651	-7.517	< 0.001				—
pop_3515m:clim_3048m	-2.386	0.651	-3.665	< 0.001				—
pop_3515m:clim_3515m	-5.432	0.651	-8.343	< 0.001				—

Table S7. *Output of type II ANOVA on our time period regression. "Population" accounts for the effects of both physiological and climate differences, and "Time period" refers to historical (1950-1959) or contemporary (2015-2024). The dependent variable is total energy gained adjusted for body size (kJ/g).*

	M. boulderensis				*M. sanguinipes*			
	num df	den df	F	p	num df	den df	F	p
Population	3	54	505.57	< 0.001	2	36	44.57	< 0.001
Time period	1	18	21.52	< 0.001	1	18	11.92	0.003
Time period:Population	3	54	2.96	0.040	2	36	13.60	< 0.001

Table S8. *Coefficients from the same time period regression model. The dependent variable is total energy gained adjusted for body size (kJ/g).*

	M. boulderensis				*M. sanguinipes*			
	Coefficient	SE	t	p	Coefficient	SE	t	p
(Intercept)	40.368	0.652	61.907	< 0.001	41.210	0.820	50.267	< 0.001
historical	-3.030	0.922	-3.286	0.003	-3.694	1.159	-3.186	0.005
1740m				—	-1.597	0.334	-4.787	< 0.001
2591m	-7.686	0.475	-16.174	< 0.001	1.848	0.334	5.540	< 0.001
3048m	0.466	0.475	0.981	0.331				—
3515m	-9.070	0.475	-19.086	< 0.001				—
historical:1740m				—	0.922	0.472	1.953	0.059
historical:2591m	-0.730	0.672	-1.086	0.282	-1.515	0.472	-3.210	0.003
historical:3048m	-0.527	0.672	-0.784	0.436				—
historical:3515m	-1.934	0.672	-2.878	0.006				—

Table S9. *Estimated contrasts (contemporary – historical) by site in our time period regression. Estimates represent change in total energy gained adjusted for body size (kJ/g).*

	M. boulderensis				*M. sanguinipes*			
Site	**Estimate**	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
1740m				—	2.772	1.159	2.391	0.0267
2195m	3.030	0.922	3.286	0.0028	3.694	1.159	3.186	0.0046
2591m	3.760	0.922	4.078	< 0.001	5.208	1.159	4.492	< 0.001
3048m	3.557	0.922	3.858	< 0.001				—
3515m	4.964	0.922	5.383	< 0.001				—

Chapter 2: Implications of behavioral thermoregulation and activity timing for energy acquisition in montane grasshoppers

In preparation for publication

Abstract

Behavioral thermoregulation, via microhabitat selection and temporal activity patterns, allows ectotherms some control in the temperatures they are experiencing and consequently their ability to perform essential activities and promote energy acquisition. Here we characterize behavioral thermoregulation and its implications for energy acquisition in two grasshopper species along a montane elevational gradient with thermal optima for food processing that often exceed environmental temperatures. We hypothesized that lower thermal optima for locomotion may result in diel temporal partitioning such that these grasshoppers fill their guts in the morning, then bask in the sunlight in the afternoon to increase their temperature and optimize their digestive throughput. We find that grasshoppers in the field thermoregulate to elevate body temperatures when below their thermal optima. Although most of a grasshopper's time is spent stationary, feeding and movement activity are observed to increase with increasing operative temperatures. We do not find clear diel patterns in gut fullness. We find that time of day dictates thermal preference in one species, with individuals seeking hotter temperatures in a thermal gradient during midday than in the morning. Overall, we find that our focal grasshoppers are able to elevate their body temperatures such that the high thermal optima for food processing are likely to promote energy acquisition. We find mixed support for temporal partitioning of activities promoting energy acquisition.

Introduction

Behavior is an underappreciated modifier of how organisms experience their environment (Sih et al., 2010). Environment and behavior form a feedback loop – an organism's environment affects its behavior, and organisms can change their environment through behavioral thermoregulation. Environmental choices are made both spatially, through microhabitat selection, and temporally, through the timing of activities (Woods et al., 2015). In ectotherms, these environmental choices determine body temperature (T_b) and thus response to climate variability and change (Kearney et al., 2009; Levy et al., 2017; Sears et al., 2019). Behavior can also co-evolve with non-behavioral traits, such as coloration or thermal performance curves (Angilletta et al., 2006).

Microhabitat selection can allow ectotherms to maintain a body temperature within their preferred range even in the face of diel temperature fluctuations (Angilletta, 2009; Whitman, 1987). This is important since ectotherms' performance of fitness-enhancing activities varies greatly with temperature (Chown & Terblanche, 2006). Behavioral thermoregulation is important to ectotherms from an energetic perspective. Warmer body temperatures increase metabolic expenditures (Dillon et al., 2010), and behavioral thermoregulation is one of the first lines of defense (Kearney et al., 2009). Conversely, warmer body temperatures can pose an opportunity to increase rates of feeding and energy gains (Buckley & Kingsolver, 2012; Youngblood et al., 2023).

Digestion tends to be the rate limiting process of energy gain for ectotherms (Levy et al., 2017). The muscle and related performances that enable foraging and mechanical processing of food are viable over a broader range of temperatures than the biochemical processing of food via digestion and assimilation (Angilletta et al., 2002; Huey, 1982; Stevenson et al., 1985). Many ectotherms, particularly herbivores that consume tough plant matter, can initiate foraging activity at cool temperatures early in the day and fill their gut much faster than they can process food (Harrison & Fewell, 1995). Additionally, short-term feeding (filling the gut) and digestion (emptying the gut) can vary in their thermal optima (*Topts*), often with digestion rate optimized at a higher temperature (Harrison & Fewell, 1995; Kingsolver & Woods, 1997). Rates of biochemical reactions increase exponentially with temperature based on the number of molecules with kinetic energies exceeding the required activation energy (Gillooly et al., 2001). Thus, enzymes adapted to warmer temperatures increase overall rates of reactions (Angilletta et al., 2010; Kingsolver, 2009).

One energetically advantageous strategy for ectotherms is to optimize food processing at high temperatures, which can be achieved even in cool environments by using solar radiation to behaviorally thermoregulate (Brewster et al., 2021; Sinervo & Adolph, 1994). Ectotherms can then temporally partition activities to actively forage in cool morning conditions before behaviorally thermoregulating to warm conditions to optimize food processing later in the day. Studies have documented diurnal temporal partitioning in a variety of contexts, but demonstrations of temporal partitioning to maximize food acquisition and processing are limited (Kronfeld-Schor & Dayan, 2003; O'Neill & Rolston, 2007).

The body temperatures ectotherms select can depend on their food status (Hardison & Eliason, 2024). When food processing is optimized at high temperatures and food is abundant, seeking a high body temperature can maximize growth; however, when food is scarce, lower body temperatures reduce metabolic costs and often allow for more efficient absorption of nutrients (Coggan et al., 2011; Huey & Kingsolver, 2019; Miller et al., 2009). For example, Treidel et al. (2022) found that fed crickets selected warmer temperatures than starved crickets. Locusts deprived of certain nutrients sought the temperature at which those nutrients are most efficiently absorbed (Clissold et al., 2013). Consequently, the trend of seeking warmer temperatures after feeding is common in ectotherms (Angilletta, 2009). However, in some systems other considerations – such as development, predator evasion, or immunological defense – dominate thermal preference (Schuler et al., 2011).

Grasshoppers have exceptionally high *Topts* for digestion among insects (Buckley & Nufio, 2014; Harrison & Fewell, 1995; Chapter 1). Field observations show grasshoppers can effectively maintain body temperatures above ~30°C during sunny days even in cold environments (Anderson et al., 1979; Harrison & Fewell, 1995; Whitman, 1987; Nufio CR personal observations). Whitman (1987) found that in the desert *Taeniopoda eques* was able to maintain consistent body temperatures near its preferred body temperature (32.6°C) during daytime by occupying different postures and microclimates. Harrison and Fewell, (1995) found that, in the lab, *Melanoplus bivitatus* were able to gain more energy at high temperatures (35°C) than lower temperatures (25°C and 15°C), and that this was primarily driven by the temperature dependence of digestion rates. In the field, grasshoppers began moving once their body temperatures exceeded 21°C and feeding once T_bs exceeded 25°C (although they fed at lower temperatures in the lab when they had been deprived of food). They found no particular relationship found between T_b and activity above those thresholds (Harrison & Fewell, 1995).

We build on this research to evaluate thermoregulation, temporal partitioning, and their energetic consequences along a montane elevation gradient for two grasshopper species that differ in their elevational distributions, thermal biology and dispersal limitation. *Melanoplus sanguinipes* is a mid-season thermal generalist found along the high plains to the subalpine, while *Melanoplus boulderensis* is an early-season thermal specialist restricted to montane environments. The short wings of *M. boulderensis*

restrict dispersal, which has contributed to greater genetic differentiation—and the potential for local adaptation —across the elevation gradient (Slatyer et al., 2020). Both species exhibit high *Topts* for digestion (~40°C) (Chapter 1), which exceed average air temperatures and the *Topts* for locomotion (~25°C) (Buckley & Nufio, 2014).

The high *Topts* could be advantageous over lower *Topts* given that digestion depends on biochemical reactions whose rates increase exponentially with temperature (Gillooly et al., 2001). We examine the implications of the observed high *Topt* for energy gain at an example site by adapting our energy budget from Chapter 1. The peak of the distribution of estimated available body temperatures (operative temperatures Te, see methods) is below the observed digestion *Topt* and near the *Topt* for locomotion (25°C, Fig. 1A). We compare the observed thermal performance curve (TPC) to two alternative TPCs with optima of 25°C (Fig. 1A). One alternative TPC corresponds to the “hotter-is-better” or the thermodynamic constraint hypothesis, which suggests hot shifted thermal performance curves are advantageous because they allow for greater overall performance (a greater area under the TPC) due to the biochemical advantages of being warmer (Angilletta et al., 2010). We scaled down the peak performance for the TPC with a *Topt* of 25°C using the Boltzmann-Arrhenius relationship, with an activation energy of 0.6eV (Asbury & Angilletta Jr., 2010; Gillooly et al., 2001). Alternatively, the biochemical adaptation hypothesis posits that, through molecular and cellular changes, species can adapt to counter the thermodynamic constraint. Our other alternative TPC thus maintains a similar peak performance for the *Topt* of 25°C.

Our estimates suggest that the lower *Topt* would elevate seasonal (June-August) energy gain in the absence of a thermodynamic advantage of the higher *Topt* (Fig. 1B). We estimated that behavioral thermoregulation to approaching *Topt* (40°C) can help the grasshoppers reach these higher temperatures more frequently and consequently increase their energy gain (Fig. 1 C and D). Thus, behavioral thermoregulation may help grasshoppers realize potential thermodynamic advantages of a high *Topt*.

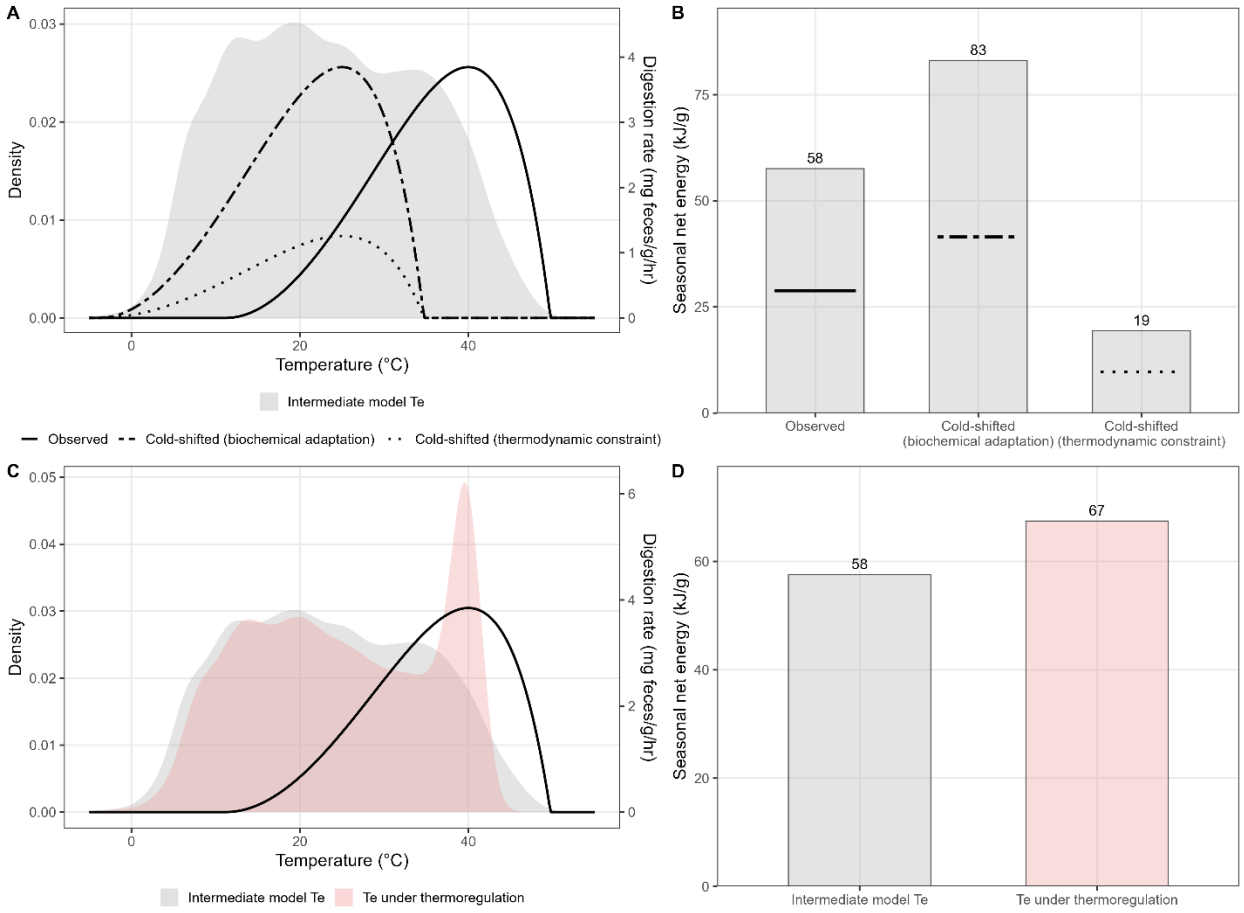

Fig. 1. *Energy gain varies with digestion TPC and thermoregulation. (A) We depict how the seasonal (June-August 2022) distribution of daylight operative temperatures at our 2591m site modelled using our biophysical model corresponds to TPCs for digestion performance* Chapter 1)*. The solid curve represents the observed digestion TPC for M. boulderensis from the 2591m site. The other two curves represent hypothetical cold-shifted TPCs under the biochemical adaptation and thermodynamic constraint hypotheses. (B) We estimate the seasonal energy gained when using each of the thermal performance curves in our energy budget, leaving other variables such as metabolic rate and mass unchanged. (C) Simulating thermoregulation toward the TPC Topt (D) increases estimated seasonal energy gain.*

We integrate field observations and controlled experiments to assess the hypothesis that our focal grasshoppers are using behavioral thermoregulation and temporal partitioning to maximize energy acquisition in their cool, montane environments. We first (1) assess whether grasshoppers are behaviorally thermoregulating to temperatures near their *Topt* for digestion. We then (2) assess grasshopper field behaviors to evaluate temporal partitioning of activity and the temperature dependence of behavior. We (3) ask whether gut fullness follows a diurnal pattern consistent with temporal partitioning.

Finally, we (4) evaluate selected body temperatures in a thermal gradient to ask whether temperature preferences depend upon time of day and/or gut fullness.

We anticipate that behavioral thermoregulation will enable grasshoppers to achieve body temperatures approximating their *Topt* for digestion. We expect the grasshoppers to exhibit the most activity at intermediate temperatures in the morning to fill their guts before remaining in warm temperatures in the afternoon conducive to digestion. We thus anticipate that gut fullness will peak by midday and then decline. We expect grasshoppers to prefer warmer temperatures in the afternoon and with full guts consistent with temporal partitioning of activity to optimize digestion.

Methods

Study system, climate data, and biophysical model

We studied grasshopper populations along an elevational gradient in Boulder County, Colorado, which has been extensively studied since the 1950s (Alexander & Hilliard, 1969; Nufio et al., 2010). We sampled populations from 4 sites at different elevations: 1740m, 2195m, 2591m, and 3048m. These elevations span the upper limit of *M. sanguinipes* and the lower limit of *M. boulderensis*, so *M. sanguinipes* only occurred at the lower 3 sites (1740-2591m), while *M. boulderensis* only occurred at the upper 3 sites (2195-3048m).

We used a biophysical model to estimate grasshopper body temperatures (operative temperatures T_e; Buckley et al., 2023) as a function of the following climate variables: wind speed, solar radiation, soil (surface) temperature, and air temperature. We deployed weather stations measuring these variables at each site. In 2022 we used Pace XR5 data loggers (Pace Scientific, Mooresville, NC, USA) logging every 5 minutes with Pace PT907 thermistors, a Pace SRS-100 Silicon Photodiode, and an anemometer. In 2023, we used Micromet stations (Dotmote Labs, Seattle, Washington, USA) logging every 10 minutes. Shaded air temperature and wind speed were measured at 1m above the ground. However, there were gaps and issues in our coverage of these variables. We used external climate data from nearby sites to fill in these gaps, fitting linear models between our data and external data for air temperature and solar radiation (external data from: Barnard & Schiff, 2026; Jager & Andreas, 1996; Morse, 2024; R^2 values for these fits were 0.77-0.94). We ultimately did not use our measured wind speed

or soil temperature data and instead made simplifying assumptions or modeled these variables. For wind speed, we assumed a constant wind speed of .5m/s at 1m height. We used TrenchR's soil_temperature() to estimate soil temperature based on air temperature, solar radiation, wind speed, air pressure, and some soil properties (solar absorptivity of 0.7, water content of 0.2, particle density of 1620; Buckley et al., 2023).

We used the function Tb_grasshopper() from the TrenchR package (Buckley et al., 2023) to estimate body temperatures via a biophysical model as a function of the climate variables above, as well as additional microclimate, organismal, and substrate variables. This function calculates operative temperature by modeling the heat exchange between the grasshopper and the surrounding environment via convection, conduction, and radiation (Buckley et al., 2014). Inputs to this function include the zenith angle of the sun, clearness index (0.7), grasshopper solar absorptivity (0.7; Anderson et al., 1979) and length (0.0225m), substrate reflectivity (0.3), and the proportion of organismal surface area in contact with this substrate (0). We chose to model a grasshopper 0.03m above the ground and experiencing 50% shading in order to estimate an intermediate body temperature.

Analysis

We fit linear models for field body temperatures and gut content, and linear mixed models for lab preferred body temperature, using the lm() and lmer() functions in R (Bates et al., 2015). We assessed the significance of each predictor using the Anova() function from the car package (Fox et al., 2001) with Type II tests, obtaining F statistics. For field behavioral observations we fit generalized additive models with binomial family using bam() from the mgcv package (Wood, 2000). Significance of parametric terms was assessed via Wald χ^2 tests, and significance of smooth terms via approximate χ^2 tests, both obtained from mgcv (Wood, 2013). In all cases, we fit separate models for our two species. We always treated site elevation as an ordered categorical variable. Our model selection was informed by AICc and our research questions. Sample sizes are summarized in the methods section, and more in-depth coverage can be found in Table S1.

Field body temperatures

In 2022 and 2023, we collected field body temperature data using an Extech type K thermocouple. A researcher would pin the head of the thermocouple between the hind leg and the abdomen to obtain a body temperature reading and release the hopper within ~10 seconds of capture. The sex and species of the grasshopper were noted along with the degree of shading where the hopper was first noticed. We performed a linear regression to assess the dependence of observed body temperature on modelled body temperature, site, and observed degree of shading. We measured 150 *M. sanguinipes* at the 1740m and 2591 sites and 287 *M. boulderensis* at the 2195m and 3014m sites. Measurements were collected over 22 days and spanned ~9AM-4PM.

Field behavioral observations

In 2022 and 2023, we leveraged an ongoing reciprocal transplant experiment to conduct instantaneous field behavioral observations within 16-inch diameter cylindrical metal cages constructed from tomato cages and wire mesh. 8-12 grasshoppers were reared within each of 3-4 cages per site. We marked adults with non-toxic paint to allow identification. Behavioral observations occurred opportunistically, with an effort to span the daylight hours of 7AM to 4PM. We had a total of 1927 observations at the four sites, occurring over 48 days in July-August (with observations at one or two sites per day). Observation periods at a single site varied but were ~1.5hrs. A single researcher cycled across each cage (sometimes completing several cycles in a day), noting the identity, activity, position, and shading of each hopper one at a time. Possible activities arose from a period of preliminary observations and included: stationary, feeding, moving, mating, and grooming. Most grasshoppers were stationary followed by moving and feeding, so we simplified the categorization by removing observations of other activities. We then collapsed moving and feeding into an “active” category, while stationary was considered “inactive.” We fit a binomial-family generalized additive model in which activity depended operative temperature (smooth k=8), site, year, and sex, with cage, individual, and observation bout (date-site combination) as random effects.

Field gut content

In 2023, we collected adult grasshoppers from each study population to assess their gut content at three times of day (10AM, 1:30PM, and 5PM), capturing ~10 males and 10 females at each site and time. Collections took place over 23 days, and usually collections for different times occurred on different days at a given site. We stored collected grasshoppers individually with no food in small plastic tubs or in Ziploc bags at room temperature (~21°C) for at least 24 hours (and no more than 36 hours). Grasshoppers were then weighed, stored in a new container with food, and released at their site of origin. The feces in the bags or containers were dried in a drying oven (65°C for at least 48hrs) and then weighed (Mettler Toledo XSR105DU). We used a linear regression to assess the dependency of hopper-mass-adjusted feces mass (mg feces / g hopper) on sex, site, and time of day (treated as an ordered factor).

Lab preferred body temperatures (PBTs)

In 2023, we reared *M. sanguinipes* from two populations (1740m and 2195m) in Percival model 36VL incubators under ramping 14-34°C temperature conditions with a 14 h light / 10 h dark cycle. Grasshoppers were reared in individual plastic cups and fed romaine lettuce and wheat bran. As part of another experiment, some individuals were exposed to low (+5°C) or high (+10°C) intensity heatwave conditions for three days at 3rd, 4th, and 5th instars. Once the individuals had reached adulthood, each grasshopper's preferred body temperature (PBT) was assessed under four conditions, whose order was semi-randomized. Specifically, we varied the time of day they were tested (8-10AM vs. 2-4PM, denoted as AM and PM) and whether the hoppers were deprived of food for the ~16 hours before they were tested or fed normally (ad lib), for four total combinations. In 2024, we assessed the PBT of field-caught adult *M. boulderensis* from three sites, and we varied only the timing of the PBT test, while all individuals were fed normally. Before testing, hoppers were exposed to various constant temperatures for the purpose of assessing the thermal dependence of feeding performance. Then they recovered for 1-3 days in incubators (27°C, 14 h light / 21°C, 10 h dark) before PBT trials. We measured 342 *M. sanguinipes* and 153 *M. boulderensis* individuals.

To assess PBT, we built a thermal gradient on an (0.125″ × 24″ × 48″) aluminum sheet. One end was in contact with an ice bath, while the other end was in contact with a heating band (QWORK Drum Heater Band 1200W 120V), creating a 10-55°C thermal gradient. The sheet was covered in a thin layer of sand and 11 5-cm-wide lanes were created by lengthwise corrugated plastic dividers. To keep grasshoppers inside and to still allow accurate thermal imaging, the top of the thermal gradient was covered in plastic wrap, heated by a blow dryer to pull the wrap taut. A FLIR camera (T640) was suspended over the thermal gradient to take infrared images. Tape measures were run lengthwise across the sides of the gradient, to allow visual assessment of a grasshoppers' distance from the hot end of the gradient. This was necessary because IR and regular images captured by the FLIR camera did not always discernably capture the grasshoppers' locations.

For each PBT trial, one grasshopper per lane was placed in the thermal gradient on the cold end, where a small strip of plastic wrap could be quickly opened and closed. Grasshoppers explored the thermal gradient for 30 minutes, and then, every 10 minutes for 50 minutes (following Forsman, 2000; Springate & Thomas, 2005), a researcher took a thermal image of the gradient using the FLIR app and visually estimated the distance of each hopper from the hot end of the thermal gradient (to the 0.5cm). Most -- but not all -- of the grasshoppers were relatively stationary after the 30-minute acclimation period.

These data were then processed in R using the ThermImage package (Tattersall, 2017). We wrote an R script to locate the important features of the thermal gradient, such as the four edges and each of the lanes. We translated distance from the hot end (cm) into pixels and captured a corresponding temperature from the thermal image at the center of the lane. For each grasshopper, we took the median of the six temperatures collected over the 50-minute period to get a single PBT. For our 2023 *M. sanguinipes* data, we fit a mixed effects model of median PBT's dependence on sex, site, heatwave treatment (three ordered categories: control, low-intensity, and high-intensity), feeding state, time of day and the interaction between feeding state and time of day, with individual identity as a random effect. For our 2024 *M. boulderensis* data we ran a similar model but with no treatment, and no terms involving feeding state, reflecting our simpler study design.

Results

Field body temperatures

There was a clear relationship between modelled operative temperature (T_e) and observed body temperature (T_b) (*M. sanguinipes*: $F_{1,145}$=161.7, p<0.001; *M. boulderensis*: $F_{1,282}$=240.3, p < 0.001; Table 1, Table S2). At low operative environmental temperatures (modelled in intermediate shade conditions), grasshoppers were able to behaviorally thermoregulate to elevated body temperatures, while at high T_es, hoppers depressed their T_bs (Fig. 2). For both species, the degree of observed shading had a statistically significant effect, with more shading leading to lower T_bs (Table 1). In *M. sanguinipes*, site also had a significant impact on T_bs, with the lower elevation site having higher observed T_bs (Table 1). An interaction term between T_e and site elevation was not supported (AICc was slightly smaller for the model with no interaction term for both species).

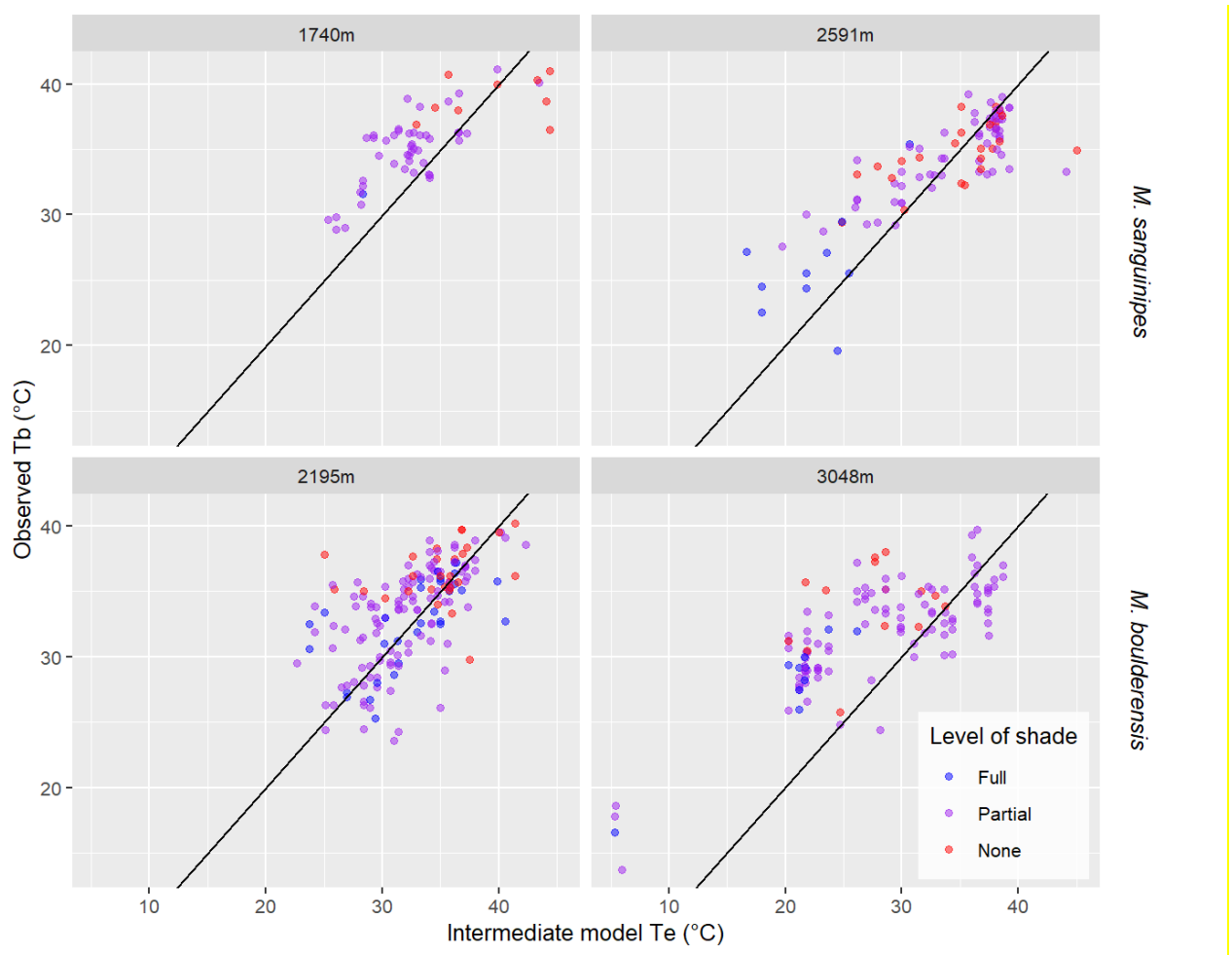

Fig. 2. *The relationship between observed body temperatures (T_bs) and modelled intermediate operative temperatures (T_es) across sites. M. sanguinipes were measured at 1740m and 2591m, while M. boulderensis were measured at 2195m and 3048m. The line denotes a 1:1 relationship of observed T_bs to modelled T_es.*

Table 1. *Anova table for how observed field body temperature (T_b) is predicted by intermediate modelled operative temperature (T_e), level of shading, and site elevation. * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	num df	**den df**	**F**	**p**	**num df**	**den df**	**F**	**p**
T_e	1	145	161.7	2.34e-25***	1	282	240.3	1.28e-39***
Shade level	2	145	7.9	5.46e-04***	2	282	11.8	1.23e-05***
Elevation	1	145	17.3	5.54e-05***	1	282	3.7	5.57e-02

Field behavioral observations

Grasshoppers were active on average only ~15-20% of the time. Our hypothesis touches on both time of day and temperature, so, in addition to our other covariates and random effects, we compared the AICc of three generalized additive models predicting activity (binary): operative temperature, time of day, and an additive model accounting for both (k=8, where k is the maximum number of degrees of freedom of smooth terms, in this case T_e and time of day). Operative temperature and the additive model were indistinguishable (ΔAICc <2), while the time-of-day-only model had a higher AICc (ΔAICc > 5). Here we present the results of the simpler, T_e-only model (Fig. 3A) but note that the additive model shows a similar albeit steeper relationship with T_e and a negative linear relationship with time of day (Fig. 3B; Table S3-S4). We find a significant effect of T_e on activity in both species: for *M. sanguinipes* it is a logistic increase (χ_1 = 5.684, p=0.0171, edf=1; Table 2), while for *M. boulderensis* it has more curvature but is still monotonically increasing ($\chi_{3.6}$ = 17.376, p=0.0016, edf=2.76; Table 2). There is, however, a lot of uncertainty and sparse data at very high T_es above 40°C, and some of the *M. sanguinipes* observations at these temperatures suggest a drop in activity (Fig. 3A). For *M. sanguinipes*, we also found that males were more active than females (χ_1 =7.056, p=0.0079; Table 2, Table S5).

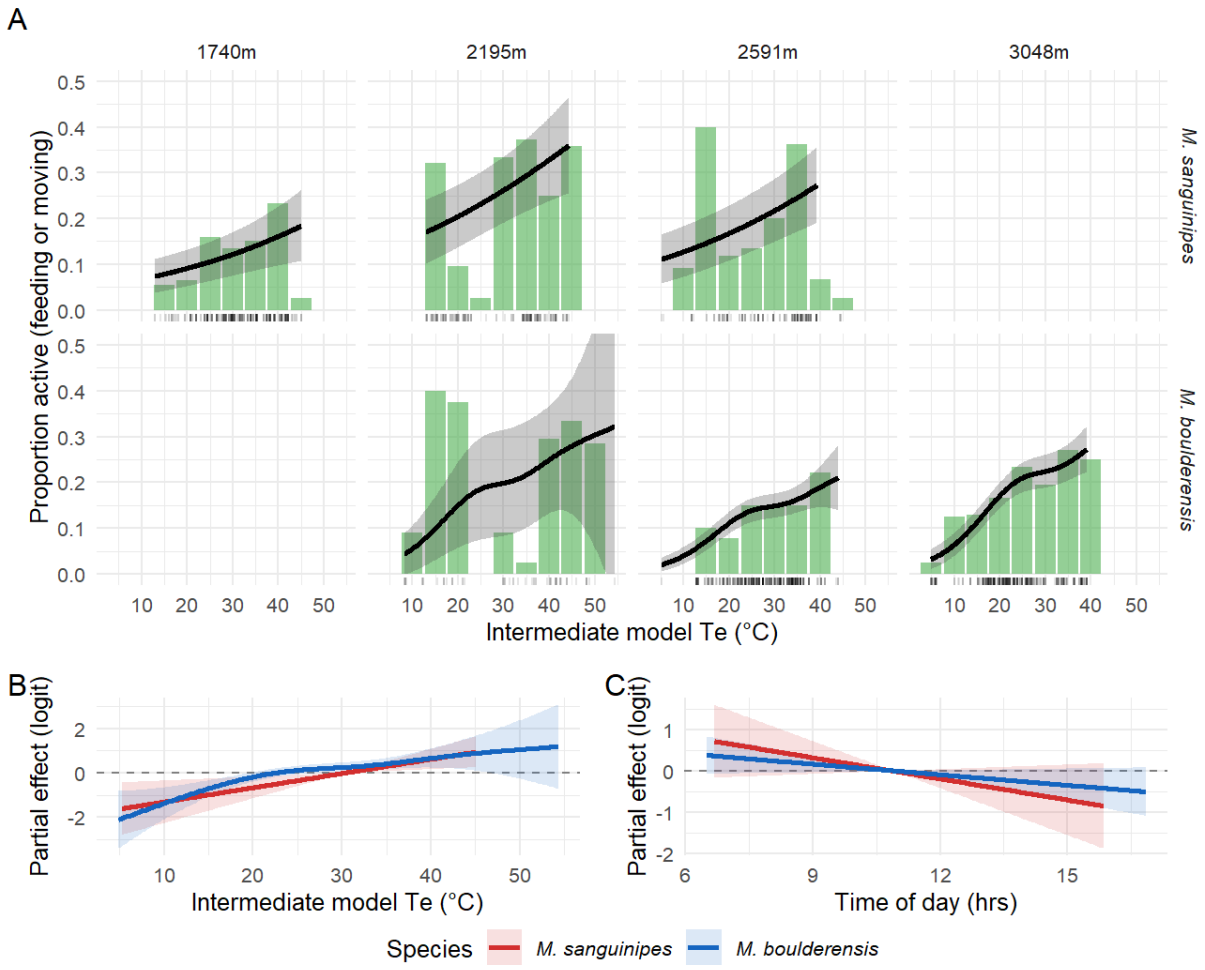

Fig. 3. *Activity varies with operative temperature (T_e). (A) The predictions of our primary generalized additive model are presented alongside a histogram of our observations. T_e intervals with fewer than 5 observations are not displayed in the histogram, while intervals with 5 or more observations but none that were active are visualized as a proportion of 0.025. (B and C) The partial effects and 95% confidence intervals of both T_e (B) and time of day (C) estimated from a secondary generalized additive model.*

Table 2. *Generalized additive models for each species regressing activity (binary: active or inactive) on operative temperature (T_e), sex, and site elevation. Degrees of freedom are estimated (edf) for smooth terms, and an adjusted edf (Ref.df) is used to calculate the p value for the χ^2 test (Wood, 2013). * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	df / edf	**Ref.df**	**χ^2**	**p**	**df / edf**	**Ref.df**	**χ^2**	**p**
Parametric terms								
Elevation	2		3.051	2.18e-01	2.00		5.302	7.06e-02
Year	1		0.020	8.87e-01	1.00		2.009	1.56e-01
Sex	1		7.056	7.90e-03**	1.00		0.135	7.14e-01
Smooth terms								
$s(T_e)$	1	1	5.684	1.71e-02*	2.78	3.6	17.376	1.59e-03**

Field gut content

Gut fullness did not vary significantly with the time of day for either species (Fig. 4, Table 3, Table S6). A model including the interaction of time of day and site was not supported by AIC (ΔAIC < 2 for *M. sanguinipes* and AIC > 5 in favor of simpler model for *M. boulderensis*). In both species, males produced less feces per gram of body mass than females (*M. sanguinipes:* $F_{1,182}$=32.3, p=5.24e-08; *M. boulderensis:* $F_{1,161}$=26.4, p=8.06e-07; Table 3). In *M. sanguinipes* there was a significant U-shaped relationship between gut fullness and site elevation ($F_{2,182}$=5.3, p=0.00578, Table 3).

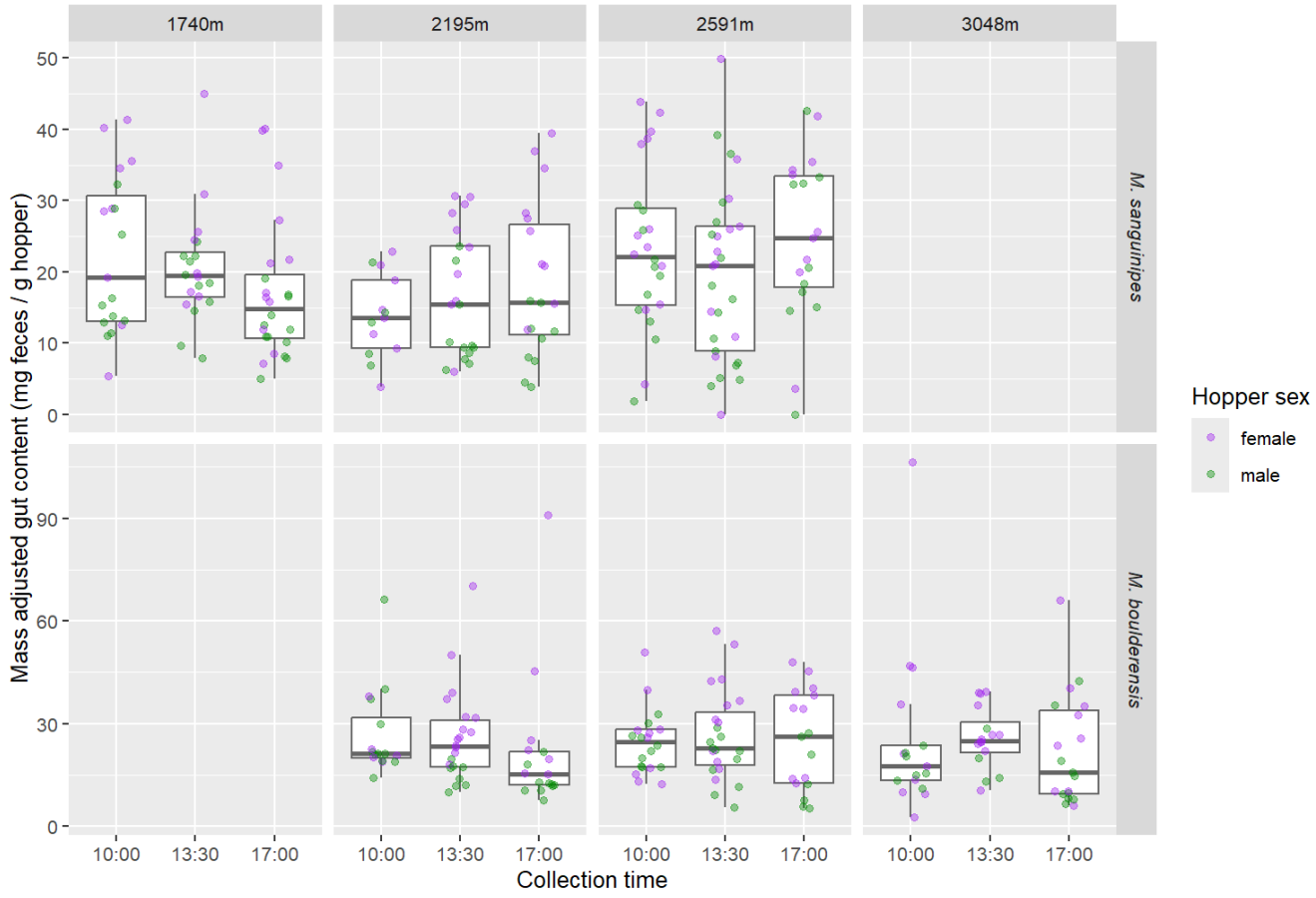

Fig. 4. *Gut fullness (mass of the feces adjusted for hopper mass, y axis) does not vary significantly with time of day across species and site elevations.*

Table 3. *Anova table for how mass-adjusted gut content (mg dried feces / g hopper) is predicted by elevation, collection time, and sex. * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	num df	**den df**	**F**	**p**	**num df**	**den df**	**F**	**p**
Elevation	2	182	5.3	5.78e-03**	2	161	0.3	7.37e-01
Collection time	2	182	0.1	8.65e-01	2	161	0.5	5.85e-01
Sex	1	182	32.3	5.24e-08***	1	161	26.4	8.06e-07***

Lab preferred body temperatures

Individuals of *M. sanguinipes* preferred significantly warmer temperatures when measured in the afternoon than in the morning ($F_{1,252.5}$=24.2, p < 0.001; Table 4, Table S7, Fig. 5A). In a marginal trend, unfed individuals selected lower temperatures ($F_{1,254.4}$=2.2; p=0.143), although, this was mostly counteracted by a (non-significant) tendency for unfed individuals to select warmer afternoon temperatures ($F_{1,255.5}$=0.8, p=0.371; Table 3, Table S7). Heatwave treatment did not have a significant effect on PBT (Table 4). We did not vary *M. boulderensis* feeding state (due to lower sample size and findings in *M. sanguinipes*) and we did not find a significant trend in PBT with time of day (Table 4, Fig. 5B). We used median selected temperatures for individuals, and the average standard deviation was 2.67°C.

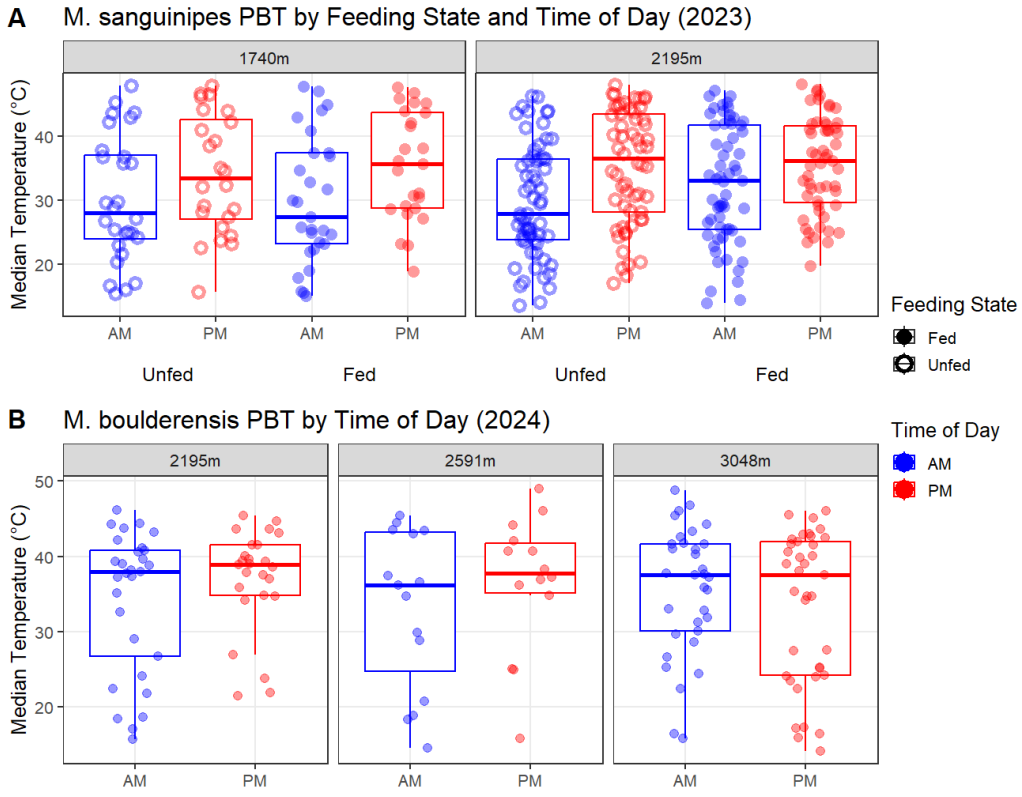

Fig. 5. *Preferred body temperature (PBT) trials. Columns represent site of origin, colors represent time of day of the trial, points represent the median selected body temperature of an individual during a trial (out of 6 time points taken every 10 minutes after a 30-minute period of acclimation to the thermal gradient). (A) M. sanguinipes PBTs of lab-reared adults. Both time of day and feeding conditions leading up to the trial were varied (x-axis / color and fill). Our linear mixed effects model shows that hoppers have higher PBTs in the afternoon (PM) than in the morning (AM). (B) M. boulderensis PBTs of field-caught adults. Our linear mixed effects model found no significant difference between morning and afternoon PBTs.*

Table 4. *Anova table of linear mixed effects models for each species predicting preferred body temperature. Treatment indicates the heat wave treatment. Differences in independent variables between species are based on differences in the experimental design in the different years. Denominator degrees of freedom are estimated using the Satterthwaite approximation. * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	***M. sanguinipes* (2023)**				***M. boulderensis* (2024)**			
	num df	**den df**	**F**	**p**	**num df**	**den df**	**F**	**p**
Sex	1	86.6	0.0	9.07e-01	1	66.9	0.3	5.98e-01
Elevation	1	89.2	0.9	3.52e-01	2	68.4	0.2	7.85e-01
Time of day	1	252.5	24.2	1.53e-06***	1	87.0	0.5	4.73e-01
Feeding state	1	254.4	2.2	1.43e-01				
Feeding state:Time of day	1	255.5	0.8	3.71e-01				
Treatment	2	87.6	1.6	2.07e-01				

Discussion

Our energy budget analysis suggests that energetic advantages of the observed high *Topt* for digestion are contingent on biochemical rate increases at warm temperatures (hotter-is-better, Fig. 1). Thermal optima are similar across species and populations along the elevation gradient (Chapter 1), perhaps reflecting constraints on biochemical adaptation (Somero, 2010). We found that *M. boulderensis* selects body temperatures approximating the *Topt* for digestion and *M. sanguinpes* selects slightly lower temperatures (Fig. 5). The energy budget suggests that raising body temperatures via thermoregulation can modestly increase energy gain along our focal elevation gradient (Fig. 1). Accordingly, we found that grasshoppers were able to elevate their body temperatures somewhat at low operative temperatures but not maintain constant body temperatures (Fig. 2).

Consistent with a limited capacity for thermoregulation, we hypothesized temporal partitioning of activity whereby grasshoppers would use their capacity for feeding at lower temperatures to fill their gut in the morning and then bask in the afternoon when warmer temperatures allow thermoregulating to temperatures near their *Topt* for digestion. We found some support for temporal partitioning of activity: *M.*

sanguinipes selected warmer body temperatures in the afternoon (Fig. 5). We found that activity increased with body temperature, but observed grasshoppers spent most of their day at body temperatures well below their thermal optima for digestion (Fig. 3). We did observe a partial effect of activity declining through the day, consistent with temporal partitioning. The overall low level of observed activity (15-20% of available time) suggest that temporal partitioning of activity may not be needed to enable foraging. Indeed, grasshopper guts appeared to reach a constant level of fullness by our first measurement period (10AM, Fig. 4). Overall, our findings indicate that our focal grasshoppers seek opportunities for energy gain whenever conditions are permissive in their cool, season-limited montane environment.

Our range of preferred field body temperatures were in line with other studies (Anderson et al., 1979; Kemp, 1986; O'Neill and Rolston, 2007). Body temperatures varied considerably as a function of environmental conditions. Body temperatures were elevated at low temperatures and depressed at very high temperatures, somewhat buffering the effect of diel environmental variation (Fig. 2). Grasshoppers are able to continue locomotion (Fig. 3A; Buckley & Nufio, 2014) and digestion (Chapter 1) at hot temperatures close to 50°C. Although we observed few temperatures that high, grasshoppers appeared able to effectively thermoregulate to avoid stressfully hot temperatures. Expanding the times of day and operative temperatures (T_es) sampled would be desirable for confirming our findings.

Our behavioral observations and our gut content findings contradict our expectation that grasshoppers would fill their guts in the morning, then bask and digest in the midday (Fig. 3, Fig. 4). Gut content does not seem to follow a diel cycle. This is not necessarily surprising since days at our Colorado sites are quite variable in terms of temperature and sunniness. A grasshopper will not necessarily end the day with an empty gut, and, since during the cold night minimal digestion occurs, the hopper would not start the next day with an empty gut either. It is more likely that feeding happens intermittently whenever digestion has made room in the gut and temperature allows for activity. Indeed, we found that feeding and movement – two behaviors linked to foraging – occurred more frequently at higher body temperatures. Harrison and Fewell (1995) described a temperature threshold for feeding and movement above which the occurrence of the activity did not depend on body temperature. In our generalized additive model,

there was not a clear threshold and the slope continued upwards even above Harrison and Fewell's threshold of 25°C. We did not find the operative temperature at which activity dropped, but that may be due to small sample sizes at very high operative temperatures. When accounting for operative temperature and time of day, we found that activity declines as it gets later, but this expected relationship was outweighed by activity increasing with temperature. In both the behavioral observations and the gut content data, we were limited by opportunistic sampling – some of the variation in our outcomes is the result of variation in the days data collection occurred.

Finally, we found that in *M. sanguinipes*, the preferred body temperature was higher in midday than it was in the morning, consistent with our expectations for diurnal partitioning of activity (Fig. 5). Diel variation in PBT has been found before in ectotherms (Angilletta et al., 1999; Kaneko et al., 2012; Treidel et al., 2022). We did not find the same trend in *M. boulderensis*, but incubator conditions were different. The finding in *M. sanguinipes* could be due to acclimatization or it could be due to a circadian temperature preference rhythm as is found in *Drosophila* (Kaneko et al., 2012). In *M. sanguinipes*, there was a nearly significant trend of lower preferred body temperature in unfed individuals. This is what we expected to see and may warrant further investigation. High variability in preferred body temperatures of individuals may result from methodological limitation. Individuals were placed in the gradient on the cold side (~10°C), and a minority of individuals did not move from there. The warm side was ~55°C and the heat quickly dissipated from there, so the temperature may not have been warm enough to dissuade the grasshoppers from trying to escape at the ends. Still, an overall difference in thermal preference by time of day emerges in *M. sanguinipes*.

Conclusion

Behavior is an important consideration in a warming world. We find that thermoregulation enables grasshoppers to maintain body temperatures that increase energy acquisition and to avoid stressfully hot temperatures (Fig. 1). Given the low and variable available body temperatures, our finding suggests that our focal grasshoppers may stand to increase their energy acquisition while avoiding thermally stressful conditions with climate change (Chapter 1; Youngblood et al., 2023). As generalist consumers of vegetation, our grasshoppers are unlikely to face food shortages and thus energetic challenges (i.e., metabolic meltdown, Huey and Kingsolver 2019). But declines in food quality with climate change (Cease,

2024) have the potential to detriment grasshoppers (Welti et al., 2020). Of course, even as rising mean temperatures could be energetically beneficial and have the potential to drive outbreaks (Youngblood et al., 2023), higher extremes could pose a threat for survival or limit activity time (Buckley et al., 2021; Dillon & Woods, 2016). Although we don't observe many operative temperatures above 40°C, it is not clear that activity time is limited at high temperatures from our behavioral observations (Fig. 3). However, drier conditions accompanying climate change can depress thermal tolerance and induce thermal stress (Youngblood et al., 2025). If suitable refugia are present, behavioral thermoregulation – via seeking shade and avoiding the hot soil – can help to buffer the effects of increasing temperature extremes under climate change (Scheffers et al., 2014; Whitman, 1987). Body size changes in our focal grasshopper system over recent decades are consistent with climate change alternatively resulting in thermal opportunity or stress: early season and higher elevation species have generally increased body size while later season and lower elevation species have declines in size (Nufio et al. 2025).

Acknowledgements

I would like to thank Anna Brasket, Leo Rupp, Monica Sheffer, César Nufio, Caroline Williams, Sean Schoville, and Lauren Buckley for their substantial contributions to this project.

I would like to thank Berry Brosi, Elli Theobald, Emily Carrington, and Andrew Berdahl for feedback on the project and/or manuscript.

Data and Code Accessibility Statement

Code and data can be accessed at https://github.com/juliamsmith/Ch2_Behavior.git.

Bibliography

Alexander, G., & Hilliard, J. R. (1969). Altitudinal and Seasonal Distribution of Orthoptera in the Rocky Mountains of Northern Colorado. *Ecological Monographs*, *39*(4), 385–432. https://doi.org/10.2307/1942354

Anderson, R. V., Tracy, C. R., & Abramsky, Z. (1979). Habitat Selection in Two Species of Short-Horned Grasshoppers. The Role of Thermal and Hydric Stresses. *Oecologia*, *38*(3), 359–374.

Angilletta, M. J. (2009). *Thermal adaptation: A theoretical and empirical synthesis*. https://books.google.com/books?hl=en&lr=&id=5RRREAAAQBAJ&oi=fnd&pg=PR7&dq=info:0o39D7IFBv0J:scholar.google.com&ots=jM_CrwThjf&sig=sC6I_QL1Cr0PhbpbfXC-X0tugI8

Angilletta, M. J., Bennett, A. F., Guderley, H., Navas, C. A., Seebacher, F., & Wilson, R. S. (2006). Coadaptation: A Unifying Principle in Evolutionary Thermal Biology. *Physiological and Biochemical Zoology*, *79*(2), 282–294. https://doi.org/10.1086/499990

Angilletta, M. J., Hill, T., & Robson, M. A. (2002). Is physiological performance optimized by thermoregulatory behavior?: A case study of the eastern fence lizard, *Sceloporus undulatus*. *Journal of Thermal Biology*, *27*(3), 199–204. https://doi.org/10.1016/S0306-4565(01)00084-5

Angilletta, M. J., Huey, R. B., & Frazier, M. R. (2010). Thermodynamic Effects on Organismal Performance: Is Hotter Better? *Physiological and Biochemical Zoology*, *83*(2), 197–206. https://doi.org/10.1086/648567

Angilletta, M. J., Montgomery, L. G., & Werner, Y. L. (1999). Temperature Preference in Geckos: Diel Variation in Juveniles and Adults. *Herpetologica*, *55*(2), 212–222.

Asbury, D. A., & Angilletta Jr., M. J. (2010). Thermodynamic Effects on the Evolution of Performance Curves. *The American Naturalist*, *176*(2), E40–E49. https://doi.org/10.1086/653659

Barnard, H. R., & Schiff, M. (2026). *DWCZ - CO - Gordon Gulch 1, South & North Facing Meteorological Towers—2020-Ongoing (DWCZ-GG1-SF-NF-Met)* [Dataset]. https://www.hydroshare.org/resource/d14f77371bb348b3b8d90bdaefe7b432/

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Brewster, C. L., Gifford, M., Ortega, J., & Beaupre, S. J. (2021). Analyzing Time-Energy Constraints to Understand the Links between Environmental Change and Local Extinctions in Terrestrial Ectotherms. *The American Naturalist*, *198*(6), 719–733. https://doi.org/10.1086/716725

Buckley, L. B., & Kingsolver, J. G. (2012). The demographic impacts of shifts in climate means and extremes on alpine butterflies. *Functional Ecology*, *26*(4), 969–977. https://doi.org/10.1111/j.1365-2435.2012.01969.x

Buckley, L. B., & Nufio, C. R. (2014). Elevational clines in the temperature dependence of insect performance and implications for ecological responses to climate change. *Conservation Physiology*, *2*(1), cou035. https://doi.org/10.1093/conphys/cou035

Buckley, L. B., Nufio, C. R., & Kingsolver, J. G. (2014). Phenotypic clines, energy balances and ecological responses to climate change. *Journal of Animal Ecology*, *83*(1), 41–50. https://doi.org/10.1111/1365-2656.12083

Buckley, L. B., Ortiz, B. A. B., Caruso, I., John, A., Levy, O., Meyer, A. V., Riddell, E. A., Sakairi, Y., & Simonis, J. L. (2023). TrenchR: An R package for modular and accessible microclimate and biophysical ecology. *PLOS Climate*, *2*(8), e0000139. https://doi.org/10.1371/journal.pclm.0000139

Buckley, L. B., Schoville, S. D., & Williams, C. M. (2021). Shifts in the relative fitness contributions of fecundity and survival in variable and changing environments. *Journal of Experimental Biology*, *224*(Suppl_1), jeb228031. https://doi.org/10.1242/jeb.228031

Cease, A. J. (2024). How Nutrients Mediate the Impacts of Global Change on Locust Outbreaks. *Annual Review of Entomology*, *69*(Volume 69, 2024), 527–550. https://doi.org/10.1146/annurev-ento-120220-110415

Chown, S. L., & Terblanche, J. S. (2006). Physiological Diversity in Insects: Ecological and Evolutionary Contexts. In *Advances in Insect Physiology* (Vol. 33, pp. 50–152). Academic Press. https://doi.org/10.1016/S0065-2806(06)33002-0

Clissold, F. J., Coggan, N., & Simpson, S. J. (2013). Insect herbivores can choose microclimates to achieve nutritional homeostasis. *Journal of Experimental Biology*, jeb.078782. https://doi.org/10.1242/jeb.078782

Coggan, N., Clissold, F. J., & Simpson, S. J. (2011). Locusts use dynamic thermoregulatory behaviour to optimize nutritional outcomes. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1719), 2745–2752. https://doi.org/10.1098/rspb.2010.2675

Dillon, M. E., Wang, G., & Huey, R. B. (2010). Global metabolic impacts of recent climate warming. *Nature*, *467*(7316), 704–706. https://doi.org/10.1038/nature09407

Dillon, M. E., & Woods, H. A. (2016). Introduction to the Symposium: Beyond the Mean: Biological Impacts of Changing Patterns of Temperature Variation. *Integrative and Comparative Biology*, *56*(1), 11–13. https://doi.org/10.1093/icb/icw020

Forsman, A. (2000). Some like it hot: Intra-Population Variation in behavioral Thermoregulation in Color-Polymorphic pygmy Grasshoppers. *Evolutionary Ecology*, *14*(1), 25–38. https://doi.org/10.1023/A:1011024320725

Fox, J., Weisberg, S., & Price, B. (2001). *car: Companion to Applied Regression* (p. 3.1-3) [Dataset]. https://doi.org/10.32614/CRAN.package.car

Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M., & Charnov, E. L. (2001). Effects of Size and Temperature on Metabolic Rate. *Science*, *293*(5538), 2248–2251. https://doi.org/10.1126/science.1061967

Hardison, E. A., & Eliason, E. J. (2024). Diet effects on ectotherm thermal performance. *Biological Reviews*, *99*(4), 1537–1555. https://doi.org/10.1111/brv.13081

Harrison, J. F., & Fewell, J. H. (1995). Thermal Effects on Feeding Behavior and Net Energy Intake in a Grasshopper Experiencing Large Diurnal Fluctuations in Body Temperature. *Physiological Zoology*, *68*(3), 453–473. https://doi.org/10.1086/physzool.68.3.30163779

Huey, R. B. (1982). Temperature, physiology, and the ecology of reptiles. In C. Gans & F. H. Pough (Eds.), *Biology of the Reptilia* (Vol. 12, pp. 25–91). Academic Press.

Huey, R. B., & Kingsolver, J. G. (2019). Climate Warming, Resource Availability, and the Metabolic Meltdown of Ectotherms. *The American Naturalist*, *194*(6), E140–E150. https://doi.org/10.1086/705679

Jager, D., & Andreas, A. (1996). *NREL National Wind Technology Center (NWTC): M2 Tower; Boulder, Colorado (Data)* [Dataset]. Not Available. https://doi.org/10.5439/1052222

Kaneko, H., Head, L. M., Ling, J., Tang, X., Liu, Y., Hardin, P. E., Emery, P., & Hamada, F. N. (2012). Circadian Rhythm of Temperature Preference and Its Neural Control in Drosophila. *Current Biology*, *22*(19), 1851–1857. https://doi.org/10.1016/j.cub.2012.08.006

Kearney, M., Shine, R., & Porter, W. P. (2009). The potential for behavioral thermoregulation to buffer “cold-blooded” animals against climate warming. *Proceedings of the National Academy of Sciences*, *106*(10), 3835–3840. https://doi.org/10.1073/pnas.0808913106

Kemp, W. P. (1986). Thermoregulation in Three Rangeland Grasshopper Species. *The Canadian Entomologist*, *118*(4), 335–343. https://doi.org/10.4039/Ent118335-4

Kingsolver, J. G. (2009). The Well-Temperatured Biologist. *The American Naturalist*, *174*(6), 755–768. https://doi.org/10.1086/648310

Kingsolver, J. G., & Woods, H. A. (1997). Thermal Sensitivity of Growth and Feeding in Manduca sexta Caterpillars. *Physiological Zoology*, *70*(6), 631–638. https://doi.org/10.1086/515872

Kronfeld-Schor, N., & Dayan, T. (2003). Partitioning of Time as an Ecological Resource. *Annual Review of Ecology, Evolution, and Systematics*, *34*(Volume 34, 2003), 153–181. https://doi.org/10.1146/annurev.ecolsys.34.011802.132435

Levy, O., Borchert, J. D., Rusch, T. W., Buckley, L. B., & Angilletta Jr., M. J. (2017). Diminishing returns limit energetic costs of climate change. *Ecology*, *98*(5), 1217–1228. https://doi.org/10.1002/ecy.1803

Miller, G. A., Clissold, F. J., Mayntz, D., & Simpson, S. J. (2009). Speed over efficiency: Locusts select body temperatures that favour growth rate over efficient nutrient utilization. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1673), 3581–3589. https://doi.org/10.1098/rspb.2009.1030

Morse, J. F. (2024). *Climate data for C1 data loggers (CR23X and CR1000), 2000—Ongoing, daily.* [Dataset]. Environmental Data Initiative. https://doi.org/10.6073/PASTA/C29B9A57257E2BB5EA110E6E0564C65B

Nufio, C. R., McGuire, C. R., Bowers, M. D., & Guralnick, R. P. (2010). Grasshopper Community Response to Climatic Change: Variation Along an Elevational Gradient. *PLOS ONE*, *5*(9), e12977. https://doi.org/10.1371/journal.pone.0012977

O'Neill, K. M., & Rolston, M. G. (2007). Short-term dynamics of behavioral thermoregulation by adults of the grasshopper Melanoplus sanguinipes. *Journal of Insect Science*, *7*(1), 27. https://doi.org/10.1673/031.007.2701

Scheffers, B. R., Edwards, D. P., Diesmos, A., Williams, S. E., & Evans, T. A. (2014). Microhabitats reduce animal's exposure to climate extremes. *Global Change Biology*, *20*(2), 495–503. https://doi.org/10.1111/gcb.12439

Schuler, M. S., Sears, M. W., & Angilletta, M. J. (2011). Food consumption does not affect the preferred body temperature of Yarrow's spiny lizard (Sceloporus jarrovi). *Journal of Thermal Biology*, *36*(2), 112–115. https://doi.org/10.1016/j.jtherbio.2010.12.002

Sears, M. W., Riddell, E. A., Rusch, T. W., & Angilletta, M. J., Jr. (2019). The World Still Is Not Flat: Lessons Learned from Organismal Interactions with Environmental Heterogeneity in Terrestrial Environments. *Integrative and Comparative Biology*, *59*(4), 1049–1058. https://doi.org/10.1093/icb/icz130

Sih, A., Stamps, J., Yang, L. H., McElreath, R., & Ramenofsky, M. (2010). Behavior as a Key Component of Integrative Biology in a Human-altered World. *Integrative and Comparative Biology*, *50*(6), 934–944. https://doi.org/10.1093/icb/icq148

Sinervo, B., & Adolph, S. C. (1994). Growth Plasticity and Thermal Opportunity in Sceloporus Lizards. *Ecology*, *75*(3), 776–790. https://doi.org/10.2307/1941734

Slatyer, R. A., Schoville, S. D., Nufio, C. R., & Buckley, L. B. (2020). Do different rates of gene flow underlie variation in phenotypic and phenological clines in a montane grasshopper community? *Ecology and Evolution*, *10*(2), 980–997. https://doi.org/10.1002/ece3.5961

Somero, G. N. (2010). The physiology of climate change: How potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers.' *Journal of Experimental Biology*, *213*(6), 912–920. https://doi.org/10.1242/jeb.037473

Springate, S., & Thomas, M. B. (2005). Thermal biology of the meadow grasshopper, Chorthippus parallelus, and the implications for resistance to disease. *Ecological Entomology*, *30*(6), 724–732. https://doi.org/10.1111/j.0307-6946.2005.00743.x

Stevenson, R. D., Peterson, C. R., & Tsuji, J. S. (1985). The Thermal Dependence of Locomotion, Tongue Flicking, Digestion, and Oxygen Consumption in the Wandering Garter Snake. *Physiological Zoology*, *58*(1), 46–57.

Tattersall, G. J. (2017). *Thermimage: Thermal Image Analysis* [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.1069705

Treidel, L. A., Huebner, C., Roberts, K. T., & Williams, C. M. (2022). Life history strategy dictates thermal preferences across the diel cycle and in response to starvation in variable field crickets, *Gryllus lineaticeps*. *Current Research in Insect Science*, *2*, 100038. https://doi.org/10.1016/j.cris.2022.100038

Whitman, D. W. (1987). Thermoregulation and daily activity patterns in a black desert grasshopper, *Taeniopoda eques*. *Animal Behaviour*, *35*(6), 1814–1826. https://doi.org/10.1016/S0003-3472(87)80074-X

Wood, S. N. (2000). *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation* (p. 1.9-4) [Dataset]. https://doi.org/10.32614/CRAN.package.mgcv

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika*, *100*(1), 221–228. https://doi.org/10.1093/biomet/ass048

Woods, H. A., Dillon, M. E., & Pincebourde, S. (2015). The roles of microclimatic diversity and of behavior in mediating the responses of ectotherms to climate change. *Journal of Thermal Biology, What Sets the Limit? How Thermal Limits, Performance and Preference in Ectotherms Are Influenced by Water or Energy Balance*, *54*, 86–97. https://doi.org/10.1016/j.jtherbio.2014.10.002

Youngblood, J. P., Cease, A. J., Talal, S., Copa, F., Medina, H. E., Rojas, J. E., Trumper, E. V., Angilletta Jr., M. J., & Harrison, J. F. (2023). Climate change expected to improve digestive rate and trigger range expansion in outbreaking locusts. *Ecological Monographs*, *93*(1), e1550. https://doi.org/10.1002/ecm.1550

Youngblood, J. P., Overgaard, J., Ørsted, M., Angilletta Jr., M. J., & VandenBrooks, J. M. (2025). Dehydration worsens heat tolerance of locusts and amplifies predicted impacts of climate change. *Functional Ecology*, *39*(5), 1194–1207. https://doi.org/10.1111/1365-2435.70038

Supplement

Table S1. *Sample sizes of individuals from each population (species-site combination) used for each experiment. MS = Melanoplus sanguinipes; MB = Melanoplus boulderensis. TOD = time of day.*

		1740m	2195m	2591m	3048m
Field T_bs	MS	56		94	
	MB		167		120
Behavioral observations	MS	376	239	204	
	MB		71	516	521
Gut content	MS	63 (19-24/TOD)	53 (13-21/TOD)	72 (19-29/TOD)	
	MB		55 (16-22/TOD)	60 (17-23/TOD)	52 (16-19/TOD)
PBT	MS	106	236		
	MB		54	29	70

Table S2. Coefficients of our regression relating observed body temperatures (T_bs) to operative temperatures (T_es), shade level, and site elevation. ** denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	Estimate	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
(Intercept)	16.85	1.02	16.48	4.74e-35***	17.76	0.97	18.35	4.90e-50***
T_e	0.45	0.04	12.72	2.34e-25***	0.47	0.03	15.50	1.28e-39***
Partial shade	2.93	0.75	3.91	1.42e-04***	0.50	0.49	1.02	3.11e-01
No shade	3.19	0.85	3.77	2.39e-04***	2.66	0.63	4.26	2.83e-05***
Elevation (linear)	-1.03	0.25	-4.15	5.54e-05***	0.51	0.27	1.92	5.57e-02

Table S3. *Alternative generalized additive model for behavioral observations including a smooth parameter for time of day (main model is presented in Table 2). Activity (binary: active or inactive) is regressed on operative temperature (T_e), time of day, sex, and site elevation. Degrees of freedom are estimated (edf) for smooth terms, and an adjusted edf (Ref.df) is used to calculate the p value for the χ^2 test (Wood, 2013). * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	df / edf	**Ref.df**	**χ^2**	**p**	**df / edf**	**Ref.df**	**χ^2**	**p**
Parametric terms								
Elevation	2		2.706	2.58e-01	2.00		7.496	2.36e-02*
Year	1		0.011	9.15e-01	1.00		0.469	4.93e-01
Sex	1		6.996	8.17e-03**	1.00		0.139	7.10e-01
Smooth terms								
s(T_e)	1	1	7.157	7.47e-03**	2.76	3.5	19.352	6.74e-04***
s(Time of day)	1	1	2.569	1.09e-01	1.00	1.0	2.793	9.47e-02

Table S4. *Coefficients for the parametric terms of the alternative model for behavioral observations. Activity (binary: active or inactive) is regressed on operative temperature (T_e), time of day, sex, and site elevation. * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	Estimate	**SE**	**z**	**p**	**Estimate**	**SE**	**z**	**p**
Intercept	-1.632	0.254	-6.430	1.28e-10***	-1.696	0.175	-9.699	3.05e-22***
Elevation (linear)	0.343	0.311	1.102	2.70e-01	0.267	0.278	0.959	3.37e-01
Elevation (quadratic)	-0.430	0.344	-1.250	2.11e-01	0.306	0.211	1.452	1.47e-01
2023	-0.047	0.440	-0.107	9.15e-01	-0.139	0.203	-0.685	4.93e-01
Sex (male)	0.517	0.195	2.645	8.17e-03**	0.069	0.185	0.372	7.10e-01

Table S5. *Coefficients for the parametric terms of the main model for behavioral observations (see Table 2). Activity (binary: active or inactive) is regressed on operative temperature (T_e), time of day, sex, and site elevation. * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	Estimate	**SE**	**z**	**p**	**Estimate**	**SE**	**z**	**p**
Intercept	-1.595	0.252	-6.321	2.60e-10***	-1.581	0.158	-9.980	1.87e-23***
Elevation (linear)	0.383	0.312	1.227	2.20e-01	0.184	0.274	0.672	5.02e-01
Elevation (quadratic)	-0.440	0.343	-1.282	2.00e-01	0.247	0.206	1.199	2.31e-01
2023	-0.063	0.442	-0.142	8.87e-01	-0.264	0.187	-1.417	1.56e-01
Sex (male)	0.517	0.195	2.656	7.90e-03**	0.068	0.185	0.367	7.14e-01

Table S6. *Coefficients for our model of how mass-adjusted gut content (mg dried feces / g hopper) is predicted by elevation, collection time, and sex (female is the reference level). * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	M. sanguinipes				*M. boulderensis*			
	Estimate	**SE**	**t**	**p**	**Estimate**	**SE**	**t**	**p**
(Intercept)	23.62	1.01	23.47	5.99e-57***	30.22	1.52	19.92	2.72e-45***
Elevation (linear)	1.69	1.19	1.42	1.57e-01	-1.39	1.95	-0.71	4.78e-01
Elevation (quadratic)	3.71	1.29	2.88	4.44e-03**	-0.63	1.86	-0.34	7.34e-01
Collection time (linear)	-0.13	1.27	-0.10	9.17e-01	-2.01	1.94	-1.03	3.03e-01
Collection time (quadratic)	0.64	1.20	0.53	5.95e-01	-0.17	1.88	-0.09	9.26e-01
Sex (male)	-8.05	1.42	-5.68	5.24e-08***	-11.43	2.23	-5.13	8.06e-07***

Table S7. *Coefficients table for two linear mixed effects models (one for each species) predicting preferred body temperature. Differences in independent variables between species are based on differences in the experimental design in the different years. Reference levels are female (sex), AM (time of day), and fed (feeding state). * denotes p<0.05, ** means p <0.01, and *** means p<0.001.*

	***M. sanguinipes* (2023)**				***M. boulderensis* (2024)**			
	Coeff.	**SE**	**t**	**p**	**Coeff.**	**SE**	**t**	**p**
(Intercept)	31.70	1.14	27.77	1.85e-73***	33.98	1.47	23.06	2.89e-40***
Sex (male)	0.14	1.20	0.12	9.07e-01	0.97	1.83	0.53	5.98e-01
Elevation (linear)	0.85	0.90	0.94	3.52e-01	-0.95	1.37	-0.69	4.90e-01
Elevation (quadratic)					-0.11	1.81	-0.06	9.53e-01
PM	3.56	1.26	2.83	5.05e-03**	0.92	1.27	0.72	4.72e-01
Unfed	-2.05	1.22	-1.68	9.34e-02				
Unfed:PM	1.59	1.77	0.90	3.71e-01				
Treatment (linear)	1.18	1.07	1.11	2.71e-01				
Treatment (quadratic)	1.70	1.06	1.60	1.13e-01				

Chapter 3: Perceived competition, mediated by sense of control, is associated with lower STEM student success regardless of race

Submitted for publication

Abstract

Background

Competition is one of the leading reasons undergraduates cite for leaving STEM. Competition is characterized by instructor and student behavior, course policies, and outside-of-class factors. Past studies on competition find that it has several negative effects on students' experience in the classroom, although its effect on student performance is less clear. Using a structural equation model analysis of survey and exam score data from 6,463 students in 46 introductory STEM classrooms at 23 institutions, we ask: 1) Does perceived competition affect students' performance directly or only indirectly, mediated through sense of belonging and/or sense of control? 2) In what ways do the relationships between perceived competition, sense of belonging, sense of control, and performance vary for students with different racial identities? 3) Do students with different racial identities perceive different levels of competition, belonging, and control? In selecting our mediators, we refer to Control-Value Theory and Self-Determination Theory. Our analysis and interpretation of results disaggregated by students' racial identities are guided by Quantitative Critical Race Theory. We consider six racial categories, balancing sample size considerations with our desire to reflect students' identities and noting that race is a social construct and racial identity is complex. We disaggregate this data with the goal of identifying and addressing inequities resulting from racism.

Results

We found several negative effects of competition. Specifically, the negative effect of perceived competition on exam z-score was mediated via students' sense of control, not sense of belonging, and there was no direct effect of competition on exam z-score. Notably, perceived competition was associated with a lower sense of belonging, but belonging was not correlated with exam z-score. These relationships did not significantly differ between students with different racial identities. However, we did find some differences in mean levels of competition, belonging, and control among students with different identities, reflecting inequities for Black and Asian students.

Conclusions

Overall, our findings indicate that perceived classroom competition is associated with lower performance for students on average, and that these relationships don't vary for students with different racial identities. Competition's negative effect on sense of belonging and sense of control are also notable in and of themselves. Although competition is multi-faceted and affected by many factors, instructors should work to reduce competition in their classrooms. Where it is not possible to reduce competition explicitly, restoring students' sense of control may help mitigate the negative effect of competition on performance.

Introduction

Overview

Competition is often cited by undergraduates as a reason for leaving their STEM major (Seymour and Hunter, 2019; Strenta et al. 1994). Introductory STEM classes – the locus of our study – serve as students' first direct interaction with these competitive environments, and these introductory courses are the main site of STEM major attrition (Olson and Riordan, 2012). In this study, we use structural equation modeling to investigate the effects of perceived competition on students' sense of belonging, sense of control, and, ultimately, performance in their introductory STEM course. We also disaggregate our model by race to investigate whether the impact of perceived competition varies for students with different racial identities. We refer to the Quantitative Critical Race Theory (QuantCrit) literature to guide our analysis and interpretation of results (Gillborn et al., 2018). We note in particular that numbers are not neutral, numbers require interpretation, and differences we see in the classroom are due not to "race" itself but due to racism (Gillborn et al., 2018).

What is competition and how does it manifest in STEM classrooms?

While there is no consensus on a single definition of the term "competition" in education research, the literature points to several aspects of the classroom experience that contribute to perceived competition. Tatapudy et al. (in rev.) interviewed students in an introductory Biology class on their perceptions of competition in their STEM courses and synthesized the themes that emerged. Student interviews revealed four types of factors that made a class feel competitive: instructor factors, student factors, course factors, and out-of-class factors (Tatapudy et al., in rev). One example of an instructor

factor is when the instructor frames their courses as a trial that only the best students will overcome (Gasiewski et al., 2012; Seymour and Hunter, 2019; Wallwey, 2022). Student factors occur when students bring a competitive orientation to the class (Weis, 2016; Seymour and Hunter, 2019), for example by frequently comparing grades. An example of a course factor is norm-referenced grading (i.e., grading on a curve) (Hughes et al., 2014; Fines, 1997). Out-of-class factors also shape the in-classroom experience: if students know they are competing with one another for internships or medical school admission, that in turn can frame the class itself as a competition (Nguyen and Lewis, 2020; Gasiewski et al., 2012). These four factors often co-exist and interact with one another (Tatapudy et al., in rev.). Our survey questions focus mainly on the student and instructor factors, emphasizing student or instruction behavior of comparing students (more details in the methods section).

Because competition is undertheorized in the literature and is a complex, multifaceted phenomenon, it can be helpful to consider with what it has been contrasted. For example, in the achievement goal literature, competition is contrasted with a focus on learning the material. According to achievement goal theory, there are two broad types of goals a student can adopt: a *mastery goal* is an aspiration to develop competence and understanding of the material, while a *performance goal* is about demonstrating one's ability by outperforming others (Meece et al., 2006; Bardach et al., 2020). In a meta-analysis, Bardach et al. (2020) found that students internalize the *goal structure* of the classroom: when students perceive a *performance goal structure* (a competitive classroom environment), they are more likely to adopt performance goals themselves. But adopting mastery goals is generally found to increase students' achievement-related behavior and self-efficacy (Meece et al., 2006). In the social interdependency literature, competition is contrasted with cooperation (Johnson et al. 1981). Here, competition is defined as one student's success negatively depending upon their peers' comparative failure. When cooperation is adopted – when students' success depends on their groupmates also succeeding – students have a higher effort to achieve, stronger interpersonal relationships in the classroom, and better psychological health (Johnson and Johnson, 2008).

In addition to these discrete contributing factors, competition is tightly intertwined with the culture of STEM in higher education. STEM is privileged within our society, and completing a STEM degree can lead to higher paying and more prestigious jobs (Imad et al., 2023; Burke et al., 2022). Compounding

that, science is presented as a meritocracy: objective, individualist, and elite (Morton et al., 2023). This prevalent narrative of science promotes competition and ignores the Eurocentric biases that make STEM – and the pathways to success in STEM – very much not "objective" (Morton et al. 2023; Imad et al., 2023). Additionally, many people, STEM professors and students included, believe that STEM capabilities are innate (Dweck, 2006; Bowman, 2023). For example, some individuals believe either you are a "math person" or you are not (Gonzalez-DeHass et al., 2024). These factors combine to create an unwelcoming environment, where STEM students feel more like they are being sorted rather than supported (Seymour and Hunter, 2019). However, STEM culture does not have to be this way: instead of a competition, STEM could be framed as a collective enterprise for the good of humanity, where everyone's perspectives are needed and valued (Imad et al., 2023).

Overall, competition is multi-faceted, and it operates at different levels, oftentimes simultaneously. What exactly competition is and isn't is still being characterized in the literature (Tatapudy et al., in rev.; Tatapudy et al., 2025), but because competition is routinely reported to be harmful to students (e.g., Canning et al., 2020; Morton et al., 2023; Imad et al., 2023, Tatapudy et al.. 2025), it is important to investigate its impacts, even without a perfect definition.

What are the effects of competition?

There is evidence that classroom competition has several negative effects on students, particularly on their motivation-related affect. Canning et al. (2020) found that perceived classroom competition in introductory STEM classrooms increased imposter feelings among students (especially first-generation students), and, as a result, indirectly decreased student engagement, attendance, and course grades and increased dropout intentions. Posselt and Lipson (2016) surveyed over 40,000 undergraduates across many institutions and fields of study and found that perceived competition was associated with higher levels of anxiety and depression, especially among queer, first-generation, Black, and Latinx students. Finally, Tatapudy and colleagues (2025) interviewed dozens of students in introductory STEM classes and found that competitive classes lowered students' sense of belonging. Cumulatively, the negative effects of competition lead to a retention problem: undergraduates who leave a STEM major cite the competitive environment as one of their top reasons for leaving (Seymour and Hunter, 2019).

While linked to many achievement-related behaviors and affects, the effect of competition on student performance itself is not entirely clear. In their meta-analysis, Murayama and Elliot (2012) found that perceived competition had no net effect on performance. Interestingly, though, they found support for an inconsistent mediator model in which competition induces both a beneficial performance-approach goal (eagerness to outcompete others) and a detrimental performance-avoidance goal (fear of underperforming compared to others).

Indeed, the literature on the topic of the impact of competition on course outcomes is mixed, often depending on how perceived classroom competition or performance goal structures are measured and the alternative to which they are compared. When competition is contrasted with cooperation the findings are fairly clear: students in cooperatively structured classrooms have a higher performance than both competitive and individualistic classrooms, with no difference between the latter two conditions (Johnson et al. 1981; Roseth et al., 2008; Qin et al., 1995). Findings on the effects of performance goals, mastery goals, and combinations thereof on achievement are more variable (Smeding et al. 2013, Bergin 1995; Bardach et al., 2020; Elliot and Church, 1997; Pintrich 2000; Senko, 2019). And in a review, Elliot (2020) notes that given negative and positive mediators of varying magnitude, the effect of competition on performance is quite context dependent and can vary with sociocultural identity.

Through what mechanisms might competition influence performance?

We investigate two potential mediators (intermediate variables) through which competition might impact student performance: sense of control and sense of belonging. We selected these two mediators because the theoretical and empirical literature position each of them as a fundamental human need, both within the classroom and beyond (Pekrun, 2006; Ryan & Deci, 2000).

Sense of control

Students' perceptions of competition may be associated with a lower sense of control in STEM classrooms. Perceived control refers to the belief that one's actions can influence outcomes, and sense of control in a classroom often emerges from cues in the instructional environment (Pekrun, 2006). Past studies suggest that if students see their success as contingent not only on their own effort but also on outperforming peers, their sense of control weakens (Frenzel et al., 2007; Choi et al., 2025). Relatedly, a decreased sense of control can compromise students' self-efficacy, or a student's belief in their ability to

succeed (Bandura, 1997; Chan and Lam, 2008). Similarly, lack of control can increase academic anxiety, which in turn lowers engagement and performance (Respondek et al., 2017).

According to Control-Value Theory, feeling a sense of control over one's ability to achieve is an important driver of student motivation and ultimately success (Pekrun, 2006). Valuing achievement is also posited as important for motivation, but Choi et al. (2025) found that competition only affected sense of control – not sense of value. Even when students value doing well in a class, if they feel unable to influence that outcome, their motivation to succeed may decline (Pekrun, 2006). This influence of perceived control on academic success has been the subject of several studies, with lower control predicting lower grades and higher dropout intention (You et al., 2011; Stupnisky et al., 2007; Respondek et al., 2017; Perry et al., 2001). Still, sense of control has not yet been tested as a mediator for the effect of competition on academic performance (although Choi et al. (2025) found an effect on student engagement).

Sense of belonging

Competition may also influence students' performance via their sense of belonging. According to Strayhorn (2018), sense of belonging refers to the extent to which students feel accepted, valued, and part of the academic community. However, sense of belonging can be conceptualized in a variety of ways. In their coding scheme, Rainey et al. (2018) divided sense of belonging into four categories: interpersonal relationships, science identity, personal interest, and (perceived) competence, highlighting that there are several components of sense of belonging. Generally speaking, competitive environments involve comparisons of students to one another, whether these comparisons are made by instructors, students, or course structures like norm-referenced grading schemes (Tatapudy et al. in rev.; Seymour and Hunter, 2017). The implication of these comparisons is that a student's status within the classroom is precarious and must be continually "proven" by outcompeting other students. This dynamic could threaten students' sense of belonging, especially if students are already feeling uncertain about whether they belong.

According to the Self-Determination Theory of motivation, belongingness, or relatedness, is a basic need that is an antecedent to motivation, with implications for performance (Ryan & Deci, 2000; Urhahne & Wijnia, 2023). Indeed, studies suggest that a lower sense of belonging may lead to lower

academic performance and reduced persistence in STEM (Walton & Cohen, 2011; Good et al., 2012; Rainey et al., 2018; Canning et al. 2020). Sense of belonging has not yet been tested as a mediator for the effect of competition on academic performance (although Canning et al. 2020 test “imposter feelings”, which are related).

Competition and race

Competition in STEM classrooms is rooted in a Eurocentric conceptualization of education and of science as 1) individualistic, 2) objective, and 3) elite (Morton et al. 2023). It is therefore unsurprising that competition might especially alienate racially minoritized students, especially in the context of introductory college STEM courses. First, the emphasis on individual success and advancement may feel unappealing to racially minoritized students, who often enter STEM fields with collaborative and prosocial values (McGee and Bentley, 2017; Jackson et al., 2017; Garibay, 2015; Morton et al., 2019). Second, “objectivity” posits science as immune to bias and cultural influence, justifying the competition as a “meritocracy,” despite historical and ongoing evidence to the contrary. Additionally, “objectivity” elevates European scientists and their ways of knowing as a presumed objective default (Imad et al., 2023). Third, the framing of STEM as elite and STEM capability as innate implies that introductory STEM courses serve a gatekeeping function (Hughes et al., 2014). Together, these three aspects of STEM culture make classrooms an unwelcoming environment for non-White students. These courses are not only a competition, but an unfair one at that: a sink-or-swim environment where prior inequities are amplified (Tatapudy et al., 2025), Eurocentric points of view are treated as objectively correct, and students of color are aware that instructors and peers may have preconceived notions about their ability to succeed in STEM (McGee, 2016; Steele and Aronson, 1995).

Especially in competitive classrooms, where performance and comparisons are centered, non-White undergraduates are forced to contend with racial stereotypes about their capacity for academic achievement. McGee (2016) found that Black and Latinx mathematics and engineering students at Predominantly White Institutions (PWIs) felt constant pressure to prove their intelligence and to assimilate to the dominant culture in order to counter stereotypes that questioned their capabilities. Even so-called “positive” stereotypes are dehumanizing and harmful: Asian and Asian American students are often perceived through the lens of the "model minority myth,” and this stereotype applies pressure on Asian

and Asian American STEM students to be perfect and ignores these students' individual passions, struggles, and needs (Museus and Park, 2015; McGee et al., 2017). Competition can be difficult for all students, but stereotypes raise the stakes for students of color.

Past research suggests that competition may have a disproportionate impact on racially minoritized students' experiences in the classroom, potentially exacerbating racial inequities in student outcomes. For example, several studies have documented competition as an aspect of the environment that racially minoritized students name when describing STEM classes (McGee, 2016; Fematt et al., 2024; Randolph, 2017). Other studies suggest that competition disproportionately harms racially minoritized students' mental health (Posselt and Lipson, 2016) and sense of belonging (Randolph, 2017; McGee, 2016). Since a student's classroom experience is known to contribute to their ultimate academic success (Eddy et al., 2015; Thoman et al., 2014), inequities in classroom experiences could lead to inequities in performance. If the same level of perceived competition has more negative impacts on racially minoritized students than on racially majoritized students, competition could be a driver of inequities in STEM courses. In the literature, there is currently no documented quantitative link between perceived competition and racial inequities in academic performance.

Sense of belonging may also vary among students with different identities. Due to negative stereotypes and historic and ongoing discrimination, the meaning and significance of belonging may be different for students with minoritized identities in STEM, who experience higher levels of stereotype threat and belonging uncertainty, than majoritized students (Murphy and Zirkel, 2015). Sense of belonging is inversely correlated to imposter feelings in academic settings (Peteet et al., 2015), and Canning and colleagues (2020) found that first-generation students experienced more impostor-like feelings in STEM environments they appraised as more competitive, compared to continuing-generation students. In our prior work, students from minoritized backgrounds (racially minoritized and/or first-generation students) described feelings of low belonging when reflecting on the impact of being in highly competitive environments (Tatapudy et al., in rev.).

Noting the Eurocentric and stereotyped nature of competition in American higher education, we test the hypothesis that competition is disproportionately harmful for students with racialized identities in

STEM. To be clear, any differences between students' experiences in the classroom are not due to race itself but are due to pervasive racism in education (Gillborn et al., 2018).

Goal for this study

In this study we quantitatively examine how students' perceptions of a competitive classroom environment relate to academic outcomes. We propose two possible mediators for this effect – sense of belonging and sense of control – and additionally test for a direct effect of perceived competition on performance. While there is past evidence for the intermediate relationships – for example, perceived competition affecting sense of belonging or sense of control affecting performance – we put these together in a mediation analysis that tests, even contrasts, sense of belonging and sense of control as impactful for student performance. We additionally test if the relationships and mean values of perceived competition, sense of control, and sense of belonging are different for students with various racial identities. In sum, we quantitatively link perceived competition, motivational factors, and racial inequities in academic performance. Our large sample size, structural equation modelling approach, comparison of two important potential mediators of performance, and disaggregation by students' racial identities contribute to the growing literature on competition.

Specifically, we address the following questions:

RQ1. Does perceived competition affect students' performance directly or only indirectly, mediated through sense of belonging and/or sense of control?

RQ2. In what ways do the relationships between perceived competition, sense of belonging, sense of control, and performance vary for students with different racial identities?

RQ3. Do students with different racial identities perceive different levels of competition, belonging, and control?

Since the history of racism and exclusion is inextricably linked with the history of education in the United States and since there is (largely qualitative) evidence that competition may have disproportionately detrimental effects on students who are minoritized in STEM (Randolph, 2017; McGee, 2016; Tatapudy et al., 2025), in our quantitative study we do not assume that the interactions between perceived competition, sense of belonging, sense of control, and performance are the same for students

with different racial identities. Instead, we disaggregate our data by students' racial identities and test for different relationships between these variables to get a clearer picture of how competition might affect different groups of students, rather than implicitly modeling the racial majority.

We employ Quantitative Critical Race Theory (QuantCrit) in our analyses and interpretation (Gillborn et al., 2018). QuantCrit stresses that quantitative studies are not objective – choices are made in the process of data collection, construction, analyses, interpretation, and presentation – and these choices are informed by the researchers' perspectives and biases. In looking for or detecting differences among racial groups vis-a-vis perceived competition, sense of belonging, sense of control, and performance, we are identifying inequities in hopes of changing the institutions that generate and perpetuate them. In other words, differences we see in the classroom are due not to "race" itself but due to racism (Castillo and Gillborn, 2023). Quantitative data are often privileged over qualitative work in research and in our society (Gillborn et al., 2018), but this is incorrect, as both have their strengths. In interpreting our findings, we turn to past qualitative work, which can better express the nuances of students' experiences. In addition, future qualitative work may shed even more light on the dynamics of perceived competition, sense of control, sense of belonging, and performance, and how they interact with racialized experiences in the classroom.

Methods

Data collection

The data were collected as part of a larger project (National Science Foundation Award numbers 2012792 and 2420369), and we are using a subset of the data here. As part of this larger project, we recruited instructors who were teaching introductory STEM classes at a variety of four-year undergraduate institutions (after data cleaning, we had 46 classes in this dataset; all but two were taught by unique instructors). In the first phase of recruitment, we reviewed studies written by introductory STEM course instructors implementing active learning strategies in their classrooms (Theobald et al., 2020). We emailed these instructors to invite them to participate in our study. If they were unable to participate, we asked these instructors to recommend colleagues who also implement active learning in their introductory STEM classrooms. Using this approach, we recruited 13 instructors at institutions across the United States. After this initial 13 instructors, we sought additional participants in three ways: 1) we asked prior

participants to suggest colleagues whom they thought might be interested in contributing, 2) using the IPEDs database, we divided the United States into eight regions (Southwest, Southeast, Rocky Mountains, Plains, New England, Mid East, Great Lakes, and the Far West), and sought even sampling by region: using institutional websites, we contacted directors of Centers for Teaching and Learning and department heads, asking them both for recommendations of faculty they think would be interested in contributing and asking them to forward our recruitment message to potential participants. We then reached out directly to instructors who were recommended. 3) We additionally recruited via faculty listservs and with individual invitations to instructors within our networks who teach introductory STEM classes. We primarily recruited instructors from 4-year public, research-intensive and regional institutions.

After obtaining required institutional permissions (see below and in the supplement), we invited students in these instructors' introductory STEM classes to complete a mid-course survey on their experience in the class and to share their grades at the end of the course. After data cleaning, our sample included data from 46 classes (25 Biology, 12 Chemistry, 4 Physics, 3 Mathematics, and 1 Computer Science and 1 Earth Science; median class enrollment: 241, range: 20-807) at 23 institutions, with responses from 6,463 students. For each class, the students took a single survey midway through the term asking, among other things, about their perception of competition in the classroom, their sense of control, their sense of belonging, and their demographic information. The survey was administered online and took around 15 minutes to complete. Students were given an online consent form detailing the study's purpose, outlining the potential risks and benefits of participating, and stressing that study participation was optional. In most classes, students were incentivized to take the survey with extra credit or course points, which were granted upon completion, whether or not the student opted to share their responses with the study. We did not use the responses of students under 18 years old in this research. Response rates after data cleaning ranged from 21% to 95% (median: 56%) per classroom. At the end of the term, we collected each student's average exam score from the professor. To ensure anonymity, student survey responses and grades were de-identified before analysis. See below and the supplement for more information about the data and questionnaire.

Human Ethics and Consent to Participate

The University of Washington Human Subjects Division determined this study to be exempt under protocol number MOD00017031. Additional permissions can be found in the supplementary material.

Survey design

At the end of our survey, we asked students to describe their racial and ethnic identity by allowing them to check all of the provided labels which applied to them, as well as offering an optional fill-in-the-blank field. The options we provided were: a) American Indian or Alaska Native, b) Asian, c) Black or African American, d) Hispanic, Latino, or Spanish, e) Middle Eastern or North African, f) Native Hawaiian or other Pacific Islander, g) White, h) Prefer not to respond, and i) Prefer to self-identify. We elaborate on our approach to the construction of racial categories in the next section.

Our survey questions about perceived competition, sense of control, and sense of belonging were adapted from previously published and/or widely used instruments (Arnold et al., 2009; Stornes and Bru, 2011; Canning et al., 2020; Perry et al., 2001; PERTs, n.d.). Students responded to questions on a 6-point Likert scale from strongly disagree to strongly agree (with no neutral option in the middle). There were 8 questions each about perceived competition and sense of control and 4 questions about sense of belonging. A full list of the questions and the response distributions can be found in the supplement (Table S1, Fig. S1).

The questions about perceived competition focused on instructor and student behavior and highlight comparisons in the classroom. We adapted two of our questions from Canning et al. (2020), one question from Stornes and Bru (2011), four questions from Arnold et al. (2009) (these questions were written to be administered in a workplace, but we reworded them slightly for the application in classrooms), and we wrote one of our own (“The grading in this class is competitive”). Examples of questions include: “The professor seems to pit students against each other in a competitive manner in this class” and “Students in this class frequently compare their performance to each other.”

While sense of belonging can be conceptualized in a variety of ways, our questions focused on social belonging. We used a previously established scale of social belonging in higher education (PERTs, n.d.), which is fairly broad and open to students’ interpretation, with statements such as “I feel accepted in this class” and “I feel I belong in this class.”

Our questions for sense of control focused on whether students perceived their performance to be a result of their efforts. Our questions came from Perry et al.'s (2001) perceived academic control scale. Examples of questions include: "The more effort I put into this class, the better I do," "I have a great deal of control over my academic performance in this class," and "No matter what I do, I can't seem to do well in this course" (reverse-coded).

Data exploration and preparation

Creating analytic categories for racial identity is nontrivial, as race is socially constructed, and using either too few or too many categories can be counterproductive to equity research (Zuberi, 2001; Gillborn et al., 2018). Using too few categories – for example, just two groups: "racially minoritized students" and "racially majoritized students" – can obscure important inequities by homogenizing very different racial experiences which may lead to different outcomes. Using too many categories can also be counterproductive when small group sample sizes mean that researchers lack the statistical power to identify inequities. It was important to us that we honored students' identities in our analyses. We also balanced this with sample size requirements of structural equation models. We wanted to represent distinct categories for Asian, Black, Hispanic, and White students. Beyond that, we wanted to represent additional identities where sample size permitted. Structural equation models require large sample sizes, and in the case of multi-group structural equation models, Wang and Wang (2019) recommend that each group have a sample size above 200, so this was our cutoff. With the QuantCrit literature in mind, we note that our decisions are subjective and imperfect, and our approach carries implications for the interpretation of our findings.

After examining our data, we created six categories for race: Asian, Black, Hispanic, White, non-PEER (PEER = persons excluded because of ethnicity and race (Asai, 2020); since Asian and White categories already exist, non-PEER here means only multiracial students who identify as Asian and White), and PEER (multiracial students who are not solely Asian and White, or students with a racial identity not captured in another category, such as "Middle Eastern or North African" (n=125), "American Indian or Alaska Native" (n=27), or "Native Hawaiian or other Pacific Islander" (n=13)). The categories we chose and the labels we applied to them were not perfect, and they often lump together a wide range of racialized experiences. For example, within the category of "Asian" there are many ethnicities with their

own specific histories, cultures, and experiences (Poon et al., 2016). “Asian” here encompasses both international students and Asian American students, who may have vastly different college experiences. Additionally, our “PEER” category lumps together multiracial students and monoracial students whose minoritized racial identities were very rare in our dataset. For students who selected a fill-in-the-blank option in addition to or instead of the options provided by our survey, we manually examined and categorized their responses. Our sample sizes of respondents in each analytic category can be viewed in the first column of Table 1.

We examined the distributions of the responses to our survey questions on competition, belonging, and control. We saw floor effects for perceived competition (for many questions “strongly disagree” was the most common response), and milder ceiling effects for sense of belonging and sense of control, so we chose to treat the data as ordinal rather than to approximate it as continuous and normally distributed.

To prepare our data for analysis, we 1) made the directionality of items within our constructs more uniform, 2) accounted for inter-classroom variation in grades, 3) examined data missingness, and 4) addressed sparse cells. Towards these ends, we executed the following procedures: 1) We flipped four reverse-coded Likert questions about control (we would later account for the method effect of reverse coding in our model). 2) We calculated the z-score of a student’s average exam grade within the context of that student’s classroom by subtracting the mean class score from that student’s score and dividing this difference by the standard deviation of scores in the class. Of note, we calculated the mean and standard deviation of exam scores from just our survey respondents, so they may not be entirely reflective of the whole class. We dropped 3 classes whose response rates were <20%, leaving 46 classes in our sample. 3) Our data contained some rows with one or more missing responses. We employed listwise deletion for any missing responses, removing 13% of participants. The majority of these missing observations were because survey respondents were missing exam scores (a total of 11% of our survey respondents did not have an associated exam score). For the perceived competition, sense of belonging, and sense of control questions, rates of missingness were less than 3% for each item. We chose listwise deletion in these cases because imputing outcome variables (in our case, exam scores) may be inadvisable (von Hippel, 2007) and because missingness due to our survey questions was minimal (Graham, 2009). After listwise

deletion, we had a sample of 6,463 students. 2% of our respondents did not share their racial identity (n=117); we included these students in our model that did not disaggregate by race and we omitted these students from the models which did include race. 4) There were several instances where specific responses to particular items were sparsely represented (less than 5 occurrences) in one or more racial categories. In these cases, and on an item-by-item basis, we collapsed this response with the adjacent response for all student surveys (regardless of race) to facilitate multi-group CFA (confirmatory factor analysis) and SEM (structural equation modeling) (DiStefano et al., 2021; details in Table S1, Fig. S1). For example, if only three Black students answered the second survey question on belonging with "Strongly disagree" (1), we would combine the “strongly disagree” category with the adjacent category – "somewhat disagree" (2) – for all students.

Overview of the link between research questions and modeling methods

Below is an overview of how we answered each of our research questions. More details on the methods can be found in later sections. All of the following analyses are performed in Mplus Version 9 and code (though, to preserve participant confidentiality, not the data) are shared in our “Availability of data and materials” statement. The first author used a large language model (Claude by Anthropic) to help write code.

RQ1. Does perceived competition affect students’ performance directly or only indirectly, mediated through sense of belonging and/or sense of control?

1. Performed a single-group CFA to confirm the fit of the measurement model.
2. Performed a single-group SEM to answer RQ1.

RQ2. In what ways do the relationships between perceived competition, sense of belonging, sense of control, and performance vary for students with different racial identities?

1. Performed a series of single- and multi- group CFAs to confirm scalar measurement invariance.
2. Performed two multi-group SEMs, one where path coefficients were allowed to vary among racial groups, and one where they were constrained to be equal among racial groups. Compared these models to answer RQ2.

RQ3. Do students with different racial identities perceive different levels of competition, belonging, and control?

1. Performed a multi-group scalar CFA with means. Set the population mean to be 0 for each latent variable and compared the mean for each racial identity to the population mean.

Confirmatory factor analysis

We used confirmatory factor analysis (CFA) to a) assess the fit of our perceived competition, sense of control, and sense of belonging model (in service of RQ1), b) assess measurement invariance of these constructs across racial categories (in service of RQ2 and RQ3), and c) assess the means of the constructs for racial categories (RQ3).

Measurement model

To assess the fit of our measurement model before performing a single-group SEM (not accounting for race), we used a three-factor CFA with the robust weighted least squares (WLSMV) estimator with the theta parameterization. We clustered by classroom (to account for the non-independence of students within class) and used correlated residuals among the subset of control questions which were originally reverse-coded and among the competition questions which asked about student-driven (as opposed to instructor-driven) competition. We chose the WLSMV estimator because it seemed best to treat our Likert data as ordinal (4-6 categories after collapsing, with ceiling and floor effects present, see Table S1 and Fig. S1) and because WLSMV handles low to moderate sample sizes better than other estimators for ordinal data (Brown, 2015).

To evaluate the overall fit of the measurement model (and all CFAs and SEMs in this paper), we used Hu and Bentler's (1999) cutoffs: CFI > 0.95, SRMR < 0.08, and RMSEA < 0.06. We also examined the omega (ω) for each factor, calculated using standardized loadings, λ_i: $\omega = (\Sigma\lambda_i)^2 / [(\Sigma\lambda_i)^2 + \Sigma(1-\lambda_i^2)]$, to assess the reliability of each of our constructs (note that this represents the reliability of the underlying latent variable, Flora 2020; Green and Yang, 2009). We used a cutoff of $\omega > 0.70$ to indicate good factor reliability (Kalkbrenner, 2024). Fit was good (CFI = 0.969, RMSEA = 0.037, SRMR = 0.050; ω = 0.92**,** 0.96**,** and 0.88 for perceived competition, sense belonging, and sense control, respectively; see Table S2 for loadings).

Measurement invariance

To address RQ2 and RQ3, we first had to establish measurement invariance between students with different racial identities (Rocabado et al., 2020; Vandenberg and Lance, 2000). To assess measurement invariance, we specified our model similarly to before, but we used a multi-group CFA. We compared the configural model (in which the same items load onto the same constructs) directly to the scalar model (additionally loadings, intercepts, and thresholds are constrained), as is the default approach to ordinal data with the WLSMV estimator (Muthén and Asparouhov, 2002). Before testing for configural invariance, we did single-group CFAs with each subpopulation (i.e., racial identity) to ensure adequate fit.

We achieved scalar invariance. To evaluate the fit of the single-group CFAs for each racial identity and of the configural model, we used Hu and Bentler's (1999) cutoffs. For comparison of configural and scalar invariance (in service of RQ2 and RQ3), we used the cutoffs used in Rutkowski and Svetina, 2017: ΔRMSEA $\leq$ 0.01 in conjunction with significant $\Delta\chi 2$ and ΔCFI $\geq$-0.004 (see Svetina et al., 2019 for overview of various cutoffs). For our single-group CFAs, all single-group models had good fit except that the fit for Asian students did not quite clear the RMSEA cutoff (RMSEA = 0.063, Table 1), but in light of the other fit indices we deemed this marginal fit to be acceptable and we proceeded to test configural invariance. The configural invariance model had acceptable CFI, RMSEA, and SRMR, so we deemed this fit to be suitable to continue to test scalar invariance. Scalar invariance is supported under Rutkowski and Svetina's (2017) cutoffs, since the CFI and RMSEA did not worsen at all, and, in fact, improved moderately (Table 1). (The $\Delta\chi 2$ test would have us reject the scalar model, but it is overly sensitive to large sample sizes (Bentler & Bonett, 1980).) Loadings of various models can be seen in Table S2.

Table 1. Testing for measurement invariance. *Rows 1-6 report the fits of each subpopulation individually, row 7 shows the configural multi-group confirmatory factor analysis (MG-CFA), and row 8 shows the scalar model (skipping the metric model, as is common when working with the WLSMV estimator). Due to the effect of large sample sizes on Δχ2, we use changes in alternative fit indices to test for scalar invariance.*

	Sample size	CFI	RMSEA	SRMR	Δχ2
Asian	1469	0.961	0.062	0.058	-
Black	259	0.985	0.042	0.057	-
Hispanic	660	0.986	0.050	0.059	-
Non-PEER	251	0.992	0.039	0.052	-
PEER	875	0.979	0.054	0.056	-
White	2832	0.977	0.041	0.048	-
Configural MG-CFA	6346	0.979	0.051	0.054	-
Scalar MG-CFA	6346	0.982	0.040	0.054	Δχ2 = 449.745, DoF = 370, p = 0.0028

Since we achieved scalar invariance, we were able meaningfully compare means between groups (Rocabado et al., 2020). Any difference between the overall mean and the mean for a specific racial identity was likely to be because of differences in experiences students have in class, rather than differences in the way the scale was working for students with different racial identities. Using a two-sided Wald test, we compared latent means for each construct to the overall mean for all respondents. This allowed us to test whether students of different racial identities experienced different levels of competition, belonging, and control (RQ3). To report standardized effect sizes, we constrained the overall mean of the population to be 0 and we calculated a modified Cohen's (1988) d: d=(μ_race - μ_overall) / σ_pooled, where σ_pooled accounts for within-group variation of all racial identities included in the analysis.

Structural equation modeling

We fit structural equation models to a) characterize the strength and explanatory power of the hypothesized relationships between perceived competition, sense of belonging, sense of control, and

exam z-score (RQ1) and b) compare the paths between these variables for students with different racial identities (RQ2).

As we did when fitting CFAs, we first fit a single-group model (not considering students' racial identities) using clustering by class and the WLSMV estimator. Our structural model included the effects of perceived competition on sense of belonging, sense of control and performance and the effects of sense of belonging and sense of control on performance, allowing us to measure the direct effect of perceived competition on performance, as well as the indirect effect through our two potential mediators (sense of control and sense of belonging). We allowed sense of control and sense of belonging to be correlated with one another. (See Fig. 1 for a visual representation of the effects we modeled.) We used the z-score of students' average exam grade, henceforth "exam z-score," as the performance outcome of interest. We report the standardized path coefficients, so that the effect is stated in terms of standard deviations (SD) of the dependent variable per SD of the independent variable.

To test whether path coefficients vary among students with different racial identities, we then fit two multi-group SEMs using our established scalar measurement model: one model constrained all path coefficients to be equal among racial groups, while the other did not constrain any of the path coefficients. We compared our constrained and unconstrained model with a $\Delta\chi 2$ test to determine whether constraining the path coefficients significantly worsened the model fit (Kline, 2015). If the model can be constrained without reducing model fit, this suggests that the relationships between perceived competition, sense of belonging, sense of control, and exam z-score do not vary significantly for students with different racial identities. In the case of our multi-group SEMs, we report the raw coefficients since standardization in Mplus occurs at the group level, so it does not use the same scale for students with different racial identities.

Results

RQ1. Does perceived competition affect students' performance directly or only indirectly, mediated through sense of belonging and/or sense of control?

We ran a single-group SEM with all students, which had a similar fit to our single-group measurement model (CFI=0.966, RMSEA=0.037, SRMR=0.051). Overall, our model explained 9.1% of

the variation in exam z-score, 14.7% of the variation in sense of belonging, and 28.1% of the variation in sense of control (Table 2).

[Table 2 should be inserted here]

We found that perceived competition significantly and substantially decreased sense of belonging and sense of control (Fig. 1, Table 2). Sense of control significantly increased exam z-score, and the indirect path from perceived competition to exam z-score via sense of control was significant, with a standardized coefficient of -0.169, meaning that a 1-standard deviation (SD) increase in perceived competition is correlated with a 0.169-SD decrease in exam z-score (Table 2, unstandardized β=-0.090, $p<0.001$). This indirect effect represents the decrement to exam z-score that is explained by perceived competition's negative effect on sense of control and its coefficient is the product of the two components that make up the indirect effect (-0.531*0.318=-0.169, Fig. 1). Meanwhile, the negative effect of perceived competition on sense of belonging was strong and significant (standardized β=-0.383, $p<0.001$), but the association between sense of belonging and exam z-score was near zero (standardized β=0.007, $p=0.748$), leading to a negligible indirect effect of perceived competition on exam z-score via sense of belonging. Notably, sense of control and sense of belonging were significantly correlated with each other ($r=0.523$, $\beta=1.414$, $p<0.001$), meaning that 27% (0.523^2) of the variation in one latent variable can be explained by the variation in the other. Also, the direct effect of perceived competition on exam z-score was positive, though not significant (standardized β=0.045**,** $p=0.176$, Fig. 1, Table 2). The total standardized effect of perceived competition on exam z-score (through all mediators and the direct path) was significant, though relatively small, at -0.127 (β=-0.066**,** $p<0.001$). To validate our estimates of indirect effects, we fit the same model using bootstrapping and obtained similar confidence intervals (Table S4; Shrout and Bolger, 2002). We also ran single-group SEMs with only Biology classes (25 classes and 3,941 students) and without the collapsing of response categories necessary for multi-group SEMs and the results remain qualitatively similar (see the quantitative results in Fig. S3-S4).

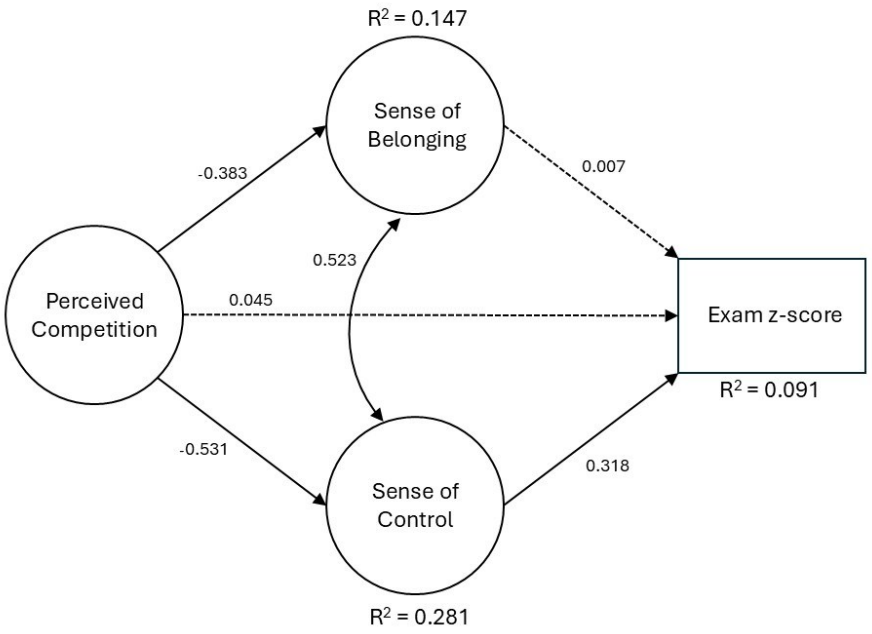

Fig. 1. *Non-disaggregated SEM results. The relationships between our variables of interest with coefficients standardized by y and x. For example, a 1-SD increase in sense of control leads to a 0.318-SD increase in exam z-score. Solid lines represent statistically significant relationships (p<0.05), while dashed lines are not statistically significant. R^2 values for endogenous variables are reported.*

RQ2. In what ways do the relationships between perceived competition, sense of belonging, sense of control, and performance vary for students with different racial identities?

Using our scalar model as the measurement model, we fit two multi-group SEMs with the same paths included as in our single-group SEM (RQ1). The first model did not constrain any path coefficients and the second model constrained all path coefficients to be equal among the racial groups. We used a $\Delta\chi 2$ test to compare the two models, and there was no significant difference ($\Delta\chi 2$=22.822, DoF=25, p=0.5880), indicating that constraining the paths did not worsen the overall fit. We present both models (Table 2). Unsurprisingly, our constrained model was very similar to our single-group SEM (RQ1). One important difference, however, is that the positive direct effect from perceived competition to exam z-score was significant (though approximately the same magnitude as the direct effect of perceived competition on exam z-score in the single-group SEM).

Moving from an unconstrained to a constrained multi-group SEM did not substantially change our explanatory power (i.e., R^2) for most racial groups, with the exception of our two smallest groups: for Black students (n=259), we saw a nearly two-fold drop in R^2 from 0.118 to 0.066, and for Non-PEER

students (n=251), we saw a modest increase in R^2 from 0.106 to 0.121. It is difficult to know if this demonstrates instability of findings due to small sample sizes, or if constraining the paths to be equal is not suitable for all racial groups despite the non-significant $\Delta\chi 2$ test.

Regardless of choice of model (constrained or unconstrained), we found that in all groups there was a negative total association of perceived competition with exam z-score, driven by the mediator of sense of control. We also found a consistent negative relationship between perceived competition and sense of belonging, although sense of belonging generally did not significantly impact exam z-score. In the unconstrained model, we did observe some differences between students with different racial identities in magnitude and significance of some effects. For example, for Black students, the total negative effect of perceived competition on performance was not statistically significant, although the magnitude of the total negative effect was comparatively large (unstandardized β=-0.084, p=0.064). Also, Hispanic students experienced a significant positive indirect effect of perceived competition on exam z-score via sense of belonging (unstandardized β=0.015, p=0.016), counterintuitively suggesting that a decreased sense of belonging is beneficial to performance. We are skeptical of this result and note that, while statistically significant, it is much smaller in magnitude than the negative indirect effect via sense of control. In the unconstrained model there was generally a small positive but nonsignificant direct effect of perceived competition on exam z-score, though the magnitude of the effect ranged from -0.004 to 0.034 for students with different racial identities. Moving to the constrained model, this positive direct effect became significant. The coefficients of our constrained model were very similar to the single-group SEM (Table 2).

In this section, we have presented the constrained and unconstrained multi-group models. However, since the constrained model is supported and is quite similar to the single-group model from RQ1, we conclude that the relationships between our variables of interest do not vary significantly for students of different racial identities. Thus, when drawing conclusions from our study, we focus on the single-group model in which students' racial identity was not explicitly included.

RQ3. Do students with different racial identities perceive different levels of competition, belonging, and control?

Racial identity did not appear to explain a lot of variation in perceived competition, sense of belonging, and sense of control. In most cases, the means for the various student racial identities hover around the overall population mean (Fig. 2). That said, compared to the overall mean, Black students reported significantly less belonging (d=-0.225, p=0.012), while Asian students reported significantly more competition (d=0.218, p=0.028) and significantly less control (d=-0.165, p=0.007) (Fig. 2, Table S3). Interestingly, though not statistically significantly, Black students reported a lower sense of control than average (d=-0.151, p=0.066).

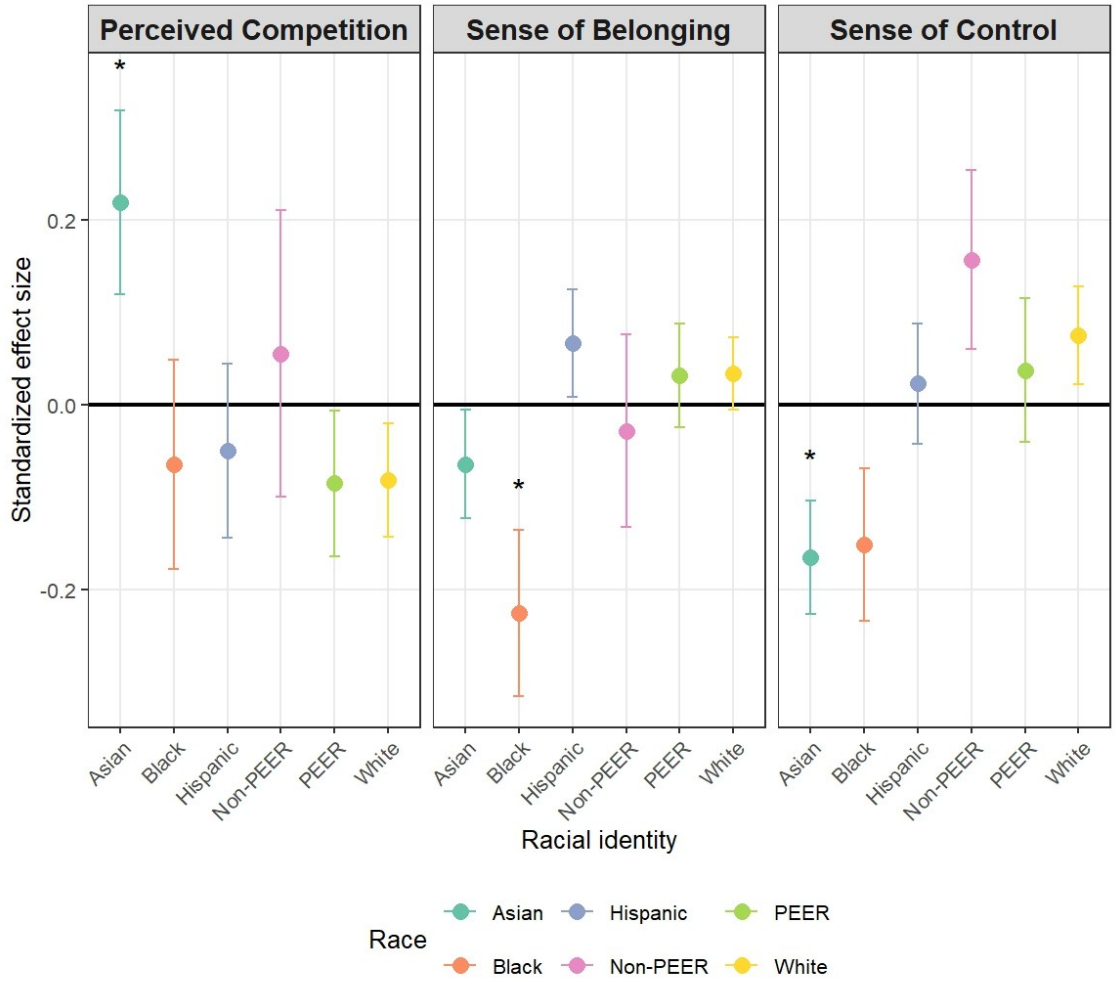

Fig. 2. *Latent variable estimates by racial identity. Levels of the latent variables – perceived competition, sense of belonging, and sense of control – for students with different racial identities generally do not significantly differ from the overall mean. Dots show the means of the latent variables for students of each racial identity when the mean of the entire student population is constrained to be 0. Negative values*

(dots below the dotted line) indicate a lower mean than the overall population and positive values indicate a higher mean. Error bars represent standard errors, and the asterisk denotes a statistically significant difference from the overall mean ($p<0.05$). Standardized effect sizes are calculated: $d=(\mu_{race} - \mu_{overall}) / \sigma_{pooled}$.

Discussion

Summary

We sought to answer three research questions related to competition in introductory undergraduate STEM classes:

RQ1. Does perceived competition affect students' performance directly or only indirectly, mediated through sense of belonging and/or sense of control?

RQ2. In what ways do the relationships between perceived competition, sense of belonging, sense of control, and performance vary for students with different racial identities?

RQ3. Do students with different racial identities perceive different levels of competition, belonging, and control?

We found that the largest impact by far of perceived competition on performance was the indirect effect via sense of control (Fig. 1, Table 2). Specifically, perceived competition reduced sense of control, which in turn reduced exam z-score. We did not find that the relationships between perceived competition, sense of belonging, sense of control, and exam z-score varied by racial identity. We found significant inequities among students with different racial identities in experiences of competition, control, and belonging (Fig. 2, Table S3). Specifically, Black students reported a lower sense of belonging than the overall mean of all students, and Asian students reported a higher level of perceived competition and a lower sense of control than the overall mean. In sum, regardless of race, competition erodes students' sense of belonging and the negative effect of competition on student performance is mediated through sense of control.

RQ1. Perceived competition affects students' performance indirectly, mediated through sense of control.

We found that the only statistically significant effect of perceived competition on exam z-score was an indirect effect through sense of control (Fig. 1, Table 2). This finding is supported by past theoretical and empirical research on sense of control and motivation (Pekrun, 2006; Frenzel et al., 2007; Choi et al., 2025; You et al., 2011; Stupnisky et al., 2007; Respondek et al., 2017, Perry et al., 2001).

Specifically, Pekrun's (2006) Control-Value Theory postulates that students' sense of control in the classroom, as well as the level of value they place on learning the material, elicits emotional responses which can either help or hinder motivation. High levels of perceived control (and value) trigger beneficial emotions, while low levels trigger negative emotions, which reduce motivation, engagement, and, ultimately, academic performance (Pekrun, 2006; Frenzel et al., 2007). Perceived competition may take away students' sense of control by making their success feel contingent on other students' failure rather than just on their own efforts (Frenzel et al., 2007). Our results constitute another piece of evidence that competition in the classroom demotivates students by reducing their sense of control, ultimately harming student performance. Consequently, instructors may wish to reduce competition in their classrooms, and, when the reduction of competition is not possible, instructors can also mitigate some of the damage by restoring students' sense of control.

We did not find a significant indirect effect through the mediator of sense of belonging: perceived competition harmed sense of belonging, but there was not a significant association between sense of belonging and exam z-score. This is surprising and not in line with previous findings. For example, Canning et al. (2020) found that perceived competition reduced student performance via increased imposter feelings. Indeed, sense of belonging is generally theorized to be important for student success (Strayhorn, 2018).

So why did we fail to find an effect of reduced sense of belonging on student outcomes, where other studies have? There are several possibilities. In our study, we used an established survey on social belonging (PERTs, n.d.); however, there are other ways of conceptualizing belonging that relate more directly to learning and performance. For example, on their 30-point math belonging scale, Good et al. (2008) have a "Trust" subscale of belongingness (i.e. "I trust that I do not have to constantly prove myself" and "Even when I do poorly, I trust my instructors to have faith in my potential."). Students' responses to questions such as these might have more proximal implications for performance, while our questions (i.e., "I feel accepted in this class") may have more distal implications for performance. We also specifically used exam z-score as our measure of performance, but a lowered sense of belonging likely has negative effects on outcomes we did not measure. For example, Rainey et al. (2018) found that student persistence in STEM majors was negatively impacted by a low sense of belonging. Finally, unlike other

studies, we tested sense of belonging and sense of control within the same model. Since the two are substantially correlated (r=0.522, p<0.001), it is possible that, once sense of control is accounted for, the effect of sense of belonging on performance that others have found shrinks or disappears.

The total effect of perceived competition on performance was significantly negative. Adding all the pathways (direct and indirect) together, we found that a 1-SD increase in perceived competition was associated with a -0.127 decrease in exam z-score (in other words, r = -0.127). Using the formula $d = 2r/(1-r^2)^{1/2}$, this is equivalent to a Cohen's d effect size of -0.26 (Rosenthal and DiMatteo, 2001). Comparing this to other effect sizes on college student's performance from meta-analyses can help contextualize these numbers (Schneider and Preckel, 2017). In a meta-analysis, Richardson and colleagues (2012) looked at many correlates of students' college GPA, including personality traits, motivational factors, self-regulatory learning strategies, students' approaches to learning, and psychosocial contextual influences. Our effect size is similar to that of student stress levels (r=0.13, 95% CI = [-0.19, -0.06]), a little bit lower than that of meta-cognition (r=0.18, 95% CI= [0.10, 0.26]), and a little bit higher than that of social support (r=0.08, 95% CI = [0.03, 0.12]) (Richardson et al., 2012). Similarly, a meta-analysis of growth mindset interventions found an overall effect of d = 0.14 (95% CI = [0.06, .22]), but with significant heterogeneity by student subpopulation (Burnette et al., 2022). Finally, a meta-analysis of active learning in post-secondary STEM classes concluded that the positive effect of active learning on exams was nearly twice as large an effect size as the effect we found (g = 0.47, where Hedges' g (Gurevitch and Hedges, 1999) was used to account for variation between studies in the meta-analysis, but it can be interpreted similarly to Cohen's d; Freeman et al., 2014). Keeping in mind that our study was observational (while the latter two of these meta-analyses considered interventions, which are likely to result in slightly lower effects (Kraft, 2020)), our total effect size of perceived competition on exam z-score is modest and not negligible.

In addition to finding a negative overall effect of perceived competition on performance, we found that perceived competition significantly harms both sense of control and sense of belonging, which is concerning in and of itself. The negative effect of perceived competition on sense of belonging can be intuitively explained: high levels of competition signal a culture where it is assumed that not everyone can succeed in STEM. Students, education scholars, and others with a vested interest in higher education

have a variety of terms and theories surrounding this culture – “weed-out culture” (Baldwin, 2009; Bok, 2006; Weston et al., 2019), “fixed mindset” (Dweck, 2006), “culture of genius” (Bowman, 2023), an “agreement of scarcity” (Imad et al., 2023) – but most agree that it is harmful for students’ sense of belonging (Seymour and Hunter, 2019). Competition also harms sense of control, as we have detailed above. In short, competition is leading to negative experiences in the classroom. Many STEM instructors no doubt want their classes to foster a sense of belonging and increase students’ sense of agency, so these negative effects of competition may be especially concerning to them.

RQ2. The relationships between perceived competition, sense of belonging, sense of control, and performance do not vary significantly among students with different racial identities.

Since we found that constraining the multi-group SEM was supported, we draw our conclusions from the single-group SEM, where race was not accounted for (Table 2). Thus, in our sample, the effects of perceived competition on sense of belonging, sense of control, and exam z-score were the same for all students, regardless of their racial identity. This was unexpected, and there are several potential explanations for this finding. First, much of the background literature which inspired our hypothesis was not quantitative and did not make comparisons between students with different racial identities (McGee, 2016; Tatapudy et al., 2025; Dost, 2024). Second, perhaps competition leads to racial inequities in dependent variables we did not measure. Our study found no disproportionate effects of perceived competition on sense of belonging, sense of control, or exam z-score, but it is possible that competition contributes to other racial inequities in the classroom. For example, Posselt and Lipson (2016) found that competition disproportionately harmed the mental health of Black and Latinx students. Finally, it may be that we did not find differences in the relationships between our variables due to unequal and/or low sample sizes for students with different racial identities or due to how we measured our latent variables.

As QuantCrit reminds us, low sample sizes and the difficulty of achieving statistical significance are known issues when disaggregating by race (Gillborn et al., 2018). In particular, we had a somewhat small sample size of Black students (n=259, on the low end for an SEM), and a larger sample size might have helped us to more accurately characterize the relationships between perceived competition, sense of belonging, and sense of control for these students. For transparency and to allow for a fuller interpretation of our data, we share the fits of our unconstrained model (Table 2). While magnitudes of the

effect of perceived competition on sense of belonging, sense of control, and exam z-score vary somewhat between students with different racial identities in our unconstrained model, the finding that competition is harmful for students of all racial identities is clear and important. Our results are just one contribution, and given important caveats (see below), do not undermine past and future qualitative and quantitative studies of the impacts of competition on racially minoritized students.

RQ3. A few racial inequities in perceived competition, sense of belonging, and sense of control were present.

We detected inequities in levels of perceived competition, sense of belonging and sense of control (Fig. 2). We found that Black students experienced a lower sense of belonging in these introductory STEM classrooms than students on average (the overall mean). This seems to be unrelated to competition but is presumably related to the racism Black students experience in education, including, but not limited to, microaggressions, stereotypes, othering, lack of representation, and other forms of discrimination, all of which send the message that they do not belong (McGee and Martin, 2011; Harper, 2015; Randolph, 2017; Dortch and Patel, 2017). We also found that Asian students perceived more competition and less control in the classroom than the overall mean. This may be related to the pressures of being treated as a "model minority" (Museus and Park, 2015). When it comes to competition, Asian students could be viewed by other students as "the one to beat," and indeed being treated as a "model minority" often entails being treated as a competitor or threat to the in-group (Fiske et al., 2002). Regarding sense of control, when instructors, peers, or others believe that STEM classes should automatically be "easy" for Asian students, these students may feel unable to meet these expectations (McGee et al., 2017). For Hispanic students, we did not find inequities in perceived competition, sense of control, and/or sense of belonging, and this might be because a sizable portion (~50%) of the Hispanic students in our sample were at Hispanic Serving Institutions (HSIs) (Murphy, 2023). Perhaps practices of HSIs and/or not being such a numerical minority in the classroom helped to address inequities in classroom experience that may be more prevalent for Hispanic students at PWIs (Cuellar, 2014). One nearly significant difference – lower sense of control among Black students – warrants further investigation.

Caveats and limitations

Despite this study's massive overall sample size, and our ability to disaggregate our data to six racial identity groups, we still remain limited by low sample sizes for students with some minoritized racial identities (Table 2). Given the sample size recommendations for SEMs (Kline, 2023; Pendergast et al., 2017, Wang and Wang, 2019), we had acceptable but low sample sizes for groups of students with some racial identities (Black and non-PEER students, n≈250). Additionally, the differences in sample sizes for different groups (approximately 10-fold from smallest to largest group sample size) may make it difficult to detect non-invariance or differences in path coefficients (Chen, 2007; Yoon and Lai, 2018). Our low sample sizes reflect the ongoing history of oppression in the United States and the exclusionary nature of higher education (Patel, 2021).

Furthermore, although our large dataset allowed us to represent racial identity beyond just a binary categorization of only "PEER" and "non-PEER," we still were not able to fully honor every student's identity. We did not have large enough sample sizes to separately represent Indigenous students, Middle Eastern students, Pacific Islander students, or the specific identities of mixed-race students. We also were not able to look at the intersections of race and gender (or race and first-generation status) because, under our current approach of multi-group SEMs, this would require us to split our groups further, resulting in even lower sample sizes.

Our R^2 values and effect sizes are also somewhat small (Table 2, top row), but this is not wholly surprising for a simple model in this field. For example, Canning et al. (2020) found similar (albeit a little bit smaller) effect sizes to ours when looking at the indirect effect of perceived competition (via imposter feelings) on course grades, course attendance, and course engagement. Additionally, considering that many factors are likely to inform exam z-score, sense of belonging, and sense of control, our modest R^2 values make sense, and still represent some explanatory power. Likewise, the effect size for the effect of perceived competition on exam z-score may be small according to Cohen's (1988) cut-off of <0.3, but more recent reviews in education and psychology suggest that these cut-offs may be unrealistically high for this type of study (Kraft, 2020; Gignac and Sozodorai, 2016). Ultimately, our effect size is medium-to-small – but certainly not negligible – by field-specific standards. Unlike the indirect effect of perceived competition on exam z-score, the associations between perceived competition and sense of control and

between perceived competition and sense of belonging both had sizable effect sizes even by Cohen's standards.

Additionally, our conclusions about competition come with some caveats. First, the students in our sample generally perceived a low level of competition (Table S1), and this may not be representative of all introductory STEM classrooms. Indeed, our classrooms were selected because they were active learning classrooms and because the instructors were interested in helping advance education research, and these selection effects may lead our sample to suggest a lower level of competition than is actually present in most introductory STEM classes. Also, the institutions in our study tend to be research-intensive state schools and are all four-year colleges. Competition and its effects should also be studied in other contexts, such as two-year colleges. Second, our dataset was mostly Biology courses (25 out of the 46 total classes were biology), so conclusions may be more relevant to biology contexts and less applicable to other disciplines (i.e., computer science) which were not as well represented in this study. Exploring the effects of the discipline of the classroom, as well as the intended major of the student, could be an informative follow-up study. Third, our analysis also operates on the within-classroom level: although we cluster by class and look at a student's average exam score relative to other students in the same class, a more complex model than ours could represent the student level and the classroom level separately (and even the institution level) and partition the variation between these levels, but this would require more data that was collected for this purpose.

Recommendations for instructors

The link between instructor practices and student perceptions of competition is no doubt real, but not entirely straightforward. The topic of this study is not classroom interventions; however, we do find that perceived competition in the classroom harms student performance (as well as their senses of control and belonging). Instructors may wish to mitigate this harm, so we briefly summarize interventions that have been studied in the past. Of note, perceptions of competition are subjective: two students in the same class might disagree on whether the class is competitive. Competition also is multi-faceted: it comes not just from instructor behavior but also peer behavior, institutional norms, and post-college realities such as medical school admissions and the job market. All this is to say that the responsibility to

reduce competition does not fall solely to instructors, but with knowledge of the multi-faceted nature of competition and its harmful effects, instructors can work to set a non-competitive tone in their classes.

In light of the negative effects of competition on students' performance and classroom experience, what can instructors do to reduce competition in their classrooms? Tatapudy et al. (in rev.) suggest key areas: messaging and course policies (including structure and assessment approach). One particularly fruitful message may be about growth mindset (Kroeper et al., 2022), wherein success is framed malleable and attained through hard work. Similarly, instructors can promote a mastery goal structure (Meece et al., 2006). Instructors can encourage their students to focus on improvement and effort rather than beating the average and can reward these behaviors with praise and credit. For example, flexibility for late submissions (Basu et al., 2024), grading for growth (Clark and Talbert, 2023), specifications gradings (Katzman et al. 2021), and avoiding norm-referenced grading (Hughes et al., 2014) could all be useful in promoting a non-competitive environment. Second, in their course design, instructors can encourage collaboration through structured group work and other opportunities for students to learn from one another in class (Ghadirian et al., 2014; Johnson et al., 2014). This intentional group work may short circuit competitive behaviors by instead building social interdependence between students (Johnson and Johnson, 2008).

Our findings suggest that restoring students' sense of control could be another avenue to combat the negative effects of competition on performance. The literature has several suggestions for instructors to improve student sense of control (summarized nicely in Respondek et al., 2017), including normalizing setbacks as a part of learning and encouraging students to adopt a growth mindset (Perry et al., 1993; Ruthig et al., 2009), offering guidance about how to study the material and prepare for tests (Stupnisky et al., 2007; Ruthig et al., 2009), offering clear rubrics and syllabi (Stupnisky et al., 2008), and offering assessment and feedback early on in the course (Tinto, 2010).

Conclusion

Leveraging a large survey dataset, we characterized the relationships between perceived competition, sense of control, sense of belonging and performance in introductory STEM classrooms. We found that perceived competition is detrimental to students' sense of belonging and sense of control. Interestingly, the negative impacts of perceived competition on course performance (measured by exam

z-score) occurred only via an indirect effect mediated by sense of control. When students find a class competitive, it can impede their learning and worsen their experience in the classroom. We found that these relationships are consistent for students of all racial identities. We hope these findings will motivate instructors to understand and combat competition in their classrooms.

List of abbreviations

CFA = confirmatory factor analysis

HSI = Hispanic serving institution

MG- = multi-group (e.g., MG-CFA)

PEER = persons excluded because of ethnicity and race

PWI = predominantly White institution

QuantCrit = Quantitative Critical Race Theory

SEM = structural equation modeling

SG- = single-group (e.g., SG-SEM)

STEM = science, technology, engineering and mathematics

SD = standard deviation(s)

Declarations

Availability of data and materials

We do not share the data for this study in order to preserve the anonymity of the participants. All code for analysis (including Mplus input and output files and some R files used for data processing and figure generation) can be viewed at https://github.com/TheobaldLab/CompetitionSEM.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the National Science Foundation (NSF DUE-2012792 and 2420369).

Authors' contributions

Sarah Eddy and Elli Theobald obtained the funding for this project. Sarah Eddy, Elli Theobald, Yoon Ha Choi, and Shangmou Xu obtained and prepared the data. Julia Smith, Sumitra Tatapudy, and Elli Theobald conceived of the project with input from Sarah Eddy. Julia Smith led the data analysis and writing with regular input from Sumitra Tatapudy and Elli Theobald. All authors reviewed the manuscript.

Acknowledgements

We would like to thank Linnea Bostrom for her work coordinating with instructors to obtain the data. We would like to thank members of the Theobald lab for friendly reviews: Carly Busch, Annie Colgan, Madison Mueller, and Rachel Potter. Furthermore, we would like to thank the first author's doctoral committee – Lauren Buckley, Berry Brosi, Emily Carrington, and Andrew Berdahl – for feedback on the project and/or manuscript. We would like thank Yale Quan at the Center for Social Science Computation and Research at the University of Washington for advice regarding structural equation modeling. We would like to thank the Center for Studies in Demography and Ecology at the University of Washington for access to Mplus. We would like to thank Scott Freeman for ad hoc consultations on this and related projects. We would also like to thank the NSF for funding this project, the instructors who contributed to our dataset, and two anonymous reviewers for feedback on the manuscript.

Tables (that are too large for in-text)

Table 2. SEM results. *A table of coefficients, R^2 values, and sample sizes from various SEMs. comp = competition, bel = belonging, ctrl = control. The first two rows are for a single-group model (SG-SEM) that includes all students and doesn't take race into account. The next row contains the path coefficients for the constrained multi-group model. The final six rows are from an unconstrained multi-group SEM (MG-SEM), where coefficients were allowed to vary among racial groups. All rows except the first one present raw, unstandardized coefficients to allow comparisons among racial identities. A second column for R^2 allows us to compare the R^2 for each racial group under the unconstrained model and the constrained one. * indicates significance at alpha = 0.05.*

		Effects of competition on exam z-score				Intermediate/other coefficients					Fit and sample size						
		Direct	Indirect (bel)	Indirect (ctrl)	Total	Comp on bel	Comp on ctrl	Bel on exam	Ctrl on exam	Bel with ctrl	R^2			R^2 constrained			N
											exam	bel	ctrl	exam	bel	ctrl	
SG-SEM	All (std.)	0.045	-0.003	**-0.169***	**-0.127***	**-0.383***	**-0.531***	0.007	**0.318***	**0.523***	0.091	0.147	0.281	-			6463
	All	0.023	-0.001	**-0.088***	**-0.066***	**-0.468***	**-0.410***	0.003	**0.215***	**1.414***	0.091	0.147	0.281	-			6463
Constrained MG-SEM	All	**0.020***	0.002	**-0.085***	**-0.063***	**-0.481***	**-0.362***	-0.003	**0.236***	(Varies)	Varies, see R^2 constrained columns			-			6346
Unconstrained MG-SEM	Asian	-0.004	0.000	**-0.081***	**-0.084***	**-0.491***	**-0.357***	0.000	**0.226***	**1.290***	0.091	0.171	0.314	0.080	0.159	0.318	1469
	Black	0.029	0.005	**-0.118***	-0.084	**-0.468***	**-0.370***	-0.010	**0.320***	**1.114***	0.118	0.155	0.332	0.066	0.153	0.323	259
	Hispanic	0.013	**0.015***	**-0.084***	**-0.056***	**-0.397***	**-0.313***	**-0.038***	**0.268***	**1.897***	0.089	0.103	0.207	0.087	0.113	0.232	660
	Non-PEER	0.023	-0.010	**-0.077***	**-0.064***	**-0.593***	**-0.399***	0.017	**0.193***	**1.658***	0.106	0.203	0.254	0.121	0.164	0.253	251
	PEER	0.034	0.010	**-0.094***	**-0.050***	**-0.421***	**-0.377***	-0.025	**0.249***	**1.394***	0.084	0.107	0.294	0.087	0.118	0.285	875
	White	0.030	-0.005	**-0.084***	**-0.059***	**-0.473***	**-0.387***	0.011	**0.217***	**1.746***	0.123	0.128	0.246	0.124	0.128	0.241	2832

Bibliography

Asai, D. J. (2020). Race Matters. *Cell*, *181*(4), 754–757. https://doi.org/10.1016/j.cell.2020.03.044

Arnold, T., Flaherty, K. E., Voss, K. E., & Mowen, J. C. (2009). Role Stressors and Retail Performance: The Role of Perceived Competitive Climate. *Journal of Retailing*, *85*(2), 194–205. https://doi.org/10.1016/j.jretai.2009.02.002

Baldwin, R. G. (2009). The climate for undergraduate teaching and learning in STEM fields. *New Directions for Teaching and Learning*, *2009*(117), 9–17. https://doi.org/10.1002/tl.340

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. Macmillan.

Bardach, L., Oczlon, S., Pietschnig, J., & Lüftenegger, M. (2020). Has achievement goal theory been right? A meta-analysis of the relation between goal structures and personal achievement goals. *Journal of Educational Psychology*, *112*(6), 1197–1220. https://doi.org/10.1037/edu0000419

Basu, D., Najjar, N., & Gosline, K. (2024). Late or Not Late—Student Perception of Performance, Confidence, and Sense of Belonging When Late Submission Is Accepted. *2024 IEEE Frontiers in Education Conference (FIE)*, 1–9. https://doi.org/10.1109/FIE61694.2024.10893379

Bello, B. D. C. (2018). *Exploring Latina and Hispanic Female Students' Sense of Belonging in STEM Majors Following a Belonging Intervention* [Ph.D., University of California, Santa Barbara]. https://www.proquest.com/docview/2137542674/abstract/63A3541DB0D24B55PQ/1

Bergin, D. A. (1995). Effects of a Mastery Versus Competitive Motivation Situation on Learning. *The Journal of Experimental Education*, *63*(4), 303–314. https://doi.org/10.1080/00220973.1995.9943466

Bok, D. (2019). *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More - New Edition*. Princeton University Press. https://doi.org/10.1515/9781400831333

Bowman, L. L. (2023). Genius culture: How we influence student identity in STEM. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1159417

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research, Second Edition*. Guilford Publications.

Burke, A., Okrent, A., Hale, K., & Gough, N. (2022). *The State of U.S. Science & Engineering 2022. National Science Board Science & Engineering Indicators. NSB-2022-1*. National Science Foundation. https://eric.ed.gov/?id=ED619338

Burnette, J., Billingsley, J., Banks, G., Knouse, L., Hoyt, C., Pollack, J., & Simon, S. (2022). A Systematic Review and Meta-Analysis of Growth Mindset Interventions: For Whom, How, and Why Might Such Interventions Work? *Psychological Bulletin*, *149*, 174–205. https://doi.org/10.1037/bul0000368

Canning, E. A., LaCosse, J., Kroeper, K. M., & Murphy, M. C. (2020). Feeling Like an Imposter: The Effect of Perceived Classroom Competition on the Daily Psychological Experiences of First-Generation College Students. *Social Psychological and Personality Science*, *11*(5), 647–657. https://doi.org/10.1177/1948550619882032

Castillo, W., & Gillborn, D. (2023). *How to "QuantCrit:" Practices and Questions for Education Data Researchers and Users. EdWorkingPaper No. 22-546*. Annenberg Institute for School Reform at Brown University. https://eric.ed.gov/?id=ED672027

Chan, J. C. Y., & Lam, S. (2008). Effects of competition on students' self-efficacy in vicarious learning. *British Journal of Educational Psychology*, *78*(1), 95–108. https://doi.org/10.1348/000709907X185509

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Choi, Y. H., Theobald, E., Velasco, V., & Eddy, S. L. (2025). Exploring how course social and cultural environmental features influence student engagement in STEM active learning courses: A control–value theory approach. *International Journal of STEM Education*, *12*(1), 4. https://doi.org/10.1186/s40594-025-00526-6

Clark, D., & Talbert, R. (2023). *Grading for Growth: A Guide to Alternative Grading Practices that Promote Authentic Learning and Student Engagement in Higher Education*. Routledge. https://doi.org/10.4324/9781003445043

Cohen, J. (1988). "The effect size," in *Statistical Power Analysis for the Behavioural Science*, 2nd Edition. ed L. Erlbaum. Taylor and Francis Inc, 79–80.

Cuellar, M. (2014). The Impact of Hispanic-Serving Institutions (HSIs), Emerging HSIs, and Non-HSIs on Latina/o Academic Self-Concept. *The Review of Higher Education*, *37*(4), 499–530.

DiStefano, C., Shi, D., & Morgan, G. B. (2021). Collapsing Categories is Often More Advantageous than Modeling Sparse Data: Investigations in the CFA Framework. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 237–249. https://doi.org/10.1080/10705511.2020.1803073

Dortch, D., & Patel, C. (2017). Black Undergraduate Women and Their Sense of Belonging in STEM at Predominantly White Institutions. *NASPA Journal About Women in Higher Education*, *10*(2), 202–215. https://doi.org/10.1080/19407882.2017.1331854

Dost, G. (2024). Students' perspectives on the 'STEM belonging' concept at A-level, undergraduate, and postgraduate levels: An examination of gender and ethnicity in student descriptions. *International Journal of STEM Education*, *11*(1), 12. https://doi.org/10.1186/s40594-024-00472-9

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. Random House Publishing Group.

Eddy, S. L., Brownell, S. E., Thummaphan, P., Lan, M.-C., & Wenderoth, M. P. (2015). Caution, Student Experience May Vary: Social Identities Impact a Student's Experience in Peer Discussions. *CBE—Life Sciences Education*, *14*(4), ar45. https://doi.org/10.1187/cbe.15-05-0108

Elliot, A. J. (2020). Competition and achievement outcomes: A hierarchical motivational analysis. *Motivation Science*, *6*(1), 3–11. https://doi.org/10.1037/mot0000164

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*(1), 218–232. https://doi.org/10.1037/0022-3514.72.1.218

Fematt, V. L., Puente, M., Garcia, K. A., & Mireles-Rios, R. (2024). "You Didn't Go by Choice!": Exposing Institutional Barriers Leading to Latinx STEM Pushout at a Hispanic-Serving Research Institution. *Education Sciences*, *14*(9), 979. https://doi.org/10.3390/educsci14090979

Fines, B. G. (1997). Competition and the Curve. *University of Missouri Kansas City Law Review*, *65*(4), 879–915.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Flora, D. B. (2020). Your Coefficient Alpha Is Probably Wrong, but Which Coefficient Omega Is Right? A Tutorial on Using R to Obtain Better Reliability Estimates. *Advances in Methods and Practices in Psychological Science*, *3*(4), 484–501. https://doi.oCrg/10.1177/2515245920951747

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, *17*(5), 478–493. https://doi.org/10.1016/j.learninstruc.2007.09.001

Garibay, J. C. (2015). STEM students' social agency and views on working for social change: Are STEM disciplines developing socially and civically responsible students? *Journal of Research in Science Teaching*, *52*(5), 610–632. https://doi.org/10.1002/tea.21203

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From Gatekeeping to Engagement: A Multicontextual, Mixed Method Study of Student Academic Engagement in Introductory STEM Courses. *Research in Higher Education*, *53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

Ghadirian, H., Fauzi Mohd Ayub, A., Daud Silong, A., Binti Abu Bakar, K., & Mohammad Hossein Zadeh, A. (2014). Knowledge Sharing Behaviour among Students in Learning Environments: A Review of Literature. *Asian Social Science*, *10*(4), p38. https://doi.org/10.5539/ass.v10n4p38

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gillborn, D., Warmington, P., & Demack, S. (2018). QuantCrit: Education, policy, 'Big Data' and principles for a critical race theory of statistics. *Race Ethnicity and Education*, *21*(2), 158–179. https://doi.org/10.1080/13613324.2017.1377417

Gonzalez-DeHass, A. R., Furner, J. M., Vásquez-Colina, M. D., & Morris, J. D. (2024). Undergraduate Students' Math Anxiety: The Role of Mindset, Achievement Goals, and Parents. *International Journal of Science & Mathematics Education*, *22*(5), 1037–1056. https://doi.org/10.1007/s10763-023-10416-4

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. https://doi.org/10.1037/a0026659

Graham, J. W. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Green, S. B., & Yang, Y. (2009). Reliability of Summed Item Scores Using Structural Equation Modeling: An Alternative to Coefficient Alpha. *Psychometrika*, *74*(1), 155–167. https://doi.org/10.1007/s11336-008-9099-3

Gurevitch, J., & Hedges, L. V. (1999). Statistical Issues in Ecological Meta-Analyses. *Ecology*, *80*(4), 1142–1149. https://doi.org/10.1890/0012-9658(1999)080%255B1142:SIIEMA%255D2.0.CO;2

Harper, S. R. (2015). Black Male College Achievers and Resistant Responses to Racist Stereotypes at Predominantly White Colleges and Universities. *Harvard Educational Review*, *85*(4), 646–674. https://doi.org/10.17763/0017-8055.85.4.646

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of Learning Skills Interventions on Student Learning: A Meta-Analysis. Review of Educational Research, 66(2), 99. https://doi.org/10.2307/1170605

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, B., Hurtado, S., & Eagan, K. (2014). Driving up or dialing down competition in introductory STEM courses: Individual and classroom level factors. *Association for the Study of Higher Education*, Retrieved on November 25, 2025, from https://www.heri.ucla.edu/nih/downloads/ASHE2014-Competition-in-Introductory-STEM-Courses.pdf

Imad, M., Reder, M., & Rose, M. (2023). Recasting the agreements to re-humanize STEM education. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1193477

Jackson, M. C., Galvez, G., Landa, I., Buonora, P., & Thoman, D. B. (2016). Science That Matters: The Importance of a Cultural Connection in Underrepresented Students' Science Pursuit. *CBE—Life Sciences Education*, *15*(3), ar42. https://doi.org/10.1187/cbe.16-01-0067

Johnson, D. W., & Johnson, R. T. (2008). Social Interdependence Theory and Cooperative Learning: The Teacher's Role. In R. M. Gillies, A. F. Ashman, & J. Terwel (Eds.), *The Teacher's Role in Implementing Cooperative Learning in the Classroom* (pp. 9–37). Springer US. https://doi.org/10.1007/978-0-387-70892-8_1

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative Learning: Improving University Instruction by Basing Practice on Validated Theory. *Journal on Excellence in College Teaching*, *25*, 85–118.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, *89*(1), 47–62. https://doi.org/10.1037/0033-2909.89.1.47

Kalkbrenner, M. T. (2024). Choosing Between Cronbach's Coefficient Alpha, McDonald's Coefficient Omega, and Coefficient H: Confidence Intervals and the Advantages and Drawbacks of Interpretive Guidelines. *Measurement and Evaluation in Counseling and Development*, *57*(2), 93–105. https://doi.org/10.1080/07481756.2023.2283637

Katzman, S. D., Hurst-Kennedy, J., Barrera, A., Talley, J., Javazon, E., Diaz, M., & Anzovino, M. E. (2021). The Effect of Specifications Grading on Students' Learning and Attitudes in an Undergraduate-Level Cell Biology Course. *Journal of Microbiology & Biology Education*, *22*(3), e00200-21. https://doi.org/10.1128/jmbe.00200-21

Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling*. Guilford Publications.

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Kroeper, K. M., Fried, A. C., & Murphy, M. C. (2022). Towards fostering growth mindset classrooms: Identifying teaching behaviors that signal instructors' fixed and growth mindsets beliefs to students. *Social Psychology of Education*, *25*(2), 371–398. https://doi.org/10.1007/s11218-022-09689-4

McGee, E. O. (2016). Devalued Black and Latino Racial Identities: A By-Product of STEM College Culture? *American Educational Research Journal*, *53*(6), 1626–1662. https://doi.org/10.3102/0002831216676572

McGee, E., & Bentley, L. (2017). The Equity Ethic: Black and Latinx College Students Reengineering Their STEM Careers toward Justice. *American Journal of Education*, *124*(1), 1–36. https://doi.org/10.1086/693954

McGee, E. O., & Martin, D. B. (2011). "You Would Not Believe What I Have to Go Through to Prove My Intellectual Value!" Stereotype Management Among Academically Successful Black Mathematics and Engineering Students. *American Educational Research Journal*, *48*(6), 1347–1389. https://doi.org/10.3102/0002831211423972

McGee, E. O., Thakore, B. K., & LaBlance, S. S. (2017). The burden of being "model": Racialized experiences of Asian STEM college students. *Journal of Diversity in Higher Education*, *10*(3), 253–270. https://doi.org/10.1037/dhe0000022

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom Goal Structure, Student Motivation, and Academic Achievement. *Annual Review of Psychology*, *57*(Volume 57, 2006), 487–503. https://doi.org/10.1146/annurev.psych.56.091103.070258

Murayama, K., & Elliot, A. J. (2012). The competition–performance relation: A meta-analytic review and test of the opposing processes model of competition and performance. *Psychological Bulletin*, *138*(6), 1035–1070. https://doi.org/10.1037/a0028324

Murphy, J. P. (2023). Hispanic Serving Institutions after a Quarter Century: Charting the Course for the Future. *Journal of Latinos and Education*, *22*(5), 1929–1938. https://doi.org/10.1080/15348431.2022.2069110

Murphy, M. C., & Zirkel, S. (2015). Race and Belonging in School: How Anticipated and Experienced Belonging Affect Choice, Persistence, and Performance. *Teachers College Record*, *117*(12), 1–40. https://doi.org/10.1177/016146811511701204

Museus, S. D., & Park, J. J. (2015). The Continuing Significance of Racism in the Lives of Asian American College Students. *Journal of College Student Development*, *56*(6), 551–569.

Muthen, B., & Asparouhov, T. (2022). *Latent Variable Analysis With Categorical Outcomes: Multiple-Group And Growth Modeling In Mplus*.

Nguyen, A., & Lewis, C. M. (2020). Competitive Enrollment Policies in Computing Departments Negatively Predict First-Year Students' Sense of Belonging, Self-Efficacy, and Perception of Department. *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, 685–691. https://doi.org/10.1145/3328778.3366805

Olson, S., & Riordan, D. G. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. *Executive Office of the President*. https://eric.ed.gov/?id=ed541511

Ong, M., Smith, J. M., & Ko, L. T. (2018). Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *Journal of Research in Science Teaching*, *55*(2), 206–245. https://doi.org/10.1002/tea.21417

Patel, L. (2021). *No Study Without Struggle: Confronting Settler Colonialism in Higher Education*. Beacon Press.

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, *60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002

Perry, R. P., Hechter, F. J., Menec, V. H., & Weinberg, L. E. (1993). Enhancing achievement motivation and performance in college students: An attributional retraining perspective. *Research in Higher Education*, *34*(6), 687–723. https://doi.org/10.1007/BF00992156

Perry, R. P., Hladkyj, S., Pekrun, R. H., & Pelletier, S. T. (2001). Academic control and action control in the achievement of college students: A longitudinal field study. *Journal of Educational Psychology*, *93*(4), 776–789. https://doi.org/10.1037/0022-0663.93.4.776

PERTS. (n.d.). "*Advance Student Success* | *Ascend"* Perts.Net. Retrieved September 12, 2025, from https://www.perts.net/ascend

Peteet, B. J., Montgomery, L., & Weekes, J. C. (2015). Predictors of Imposter Phenomenon among Talented Ethnic Minority Undergraduate Students. *The Journal of Negro Education*, *84*(2), 175–186. https://doi.org/10.7709/jnegroeducation.84.2.0175

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92*(3), 544–555. https://doi.org/10.1037/0022-0663.92.3.544

Posselt, J. R., & Lipson, S. K. (2016). Competition, Anxiety, and Depression in the College Classroom: Variations by Student Identity and Field of Study. *Journal of College Student Development*, *57*(8), 973–989.

Qin, Z., Johnson, D. W., & Johnson, R. T. (1995). Cooperative Versus Competitive Efforts and Problem Solving. *Review of Educational Research*, *65*(2), 129–143. https://doi.org/10.3102/00346543065002129

Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education*, *5*(1), 10. https://doi.org/10.1186/s40594-018-0115-6

Randolph, L. L. (2017). *Understanding Black Undergraduate Females' Sense of Belonging at a Predominantly White Institution* [Ed.D., Temple University]. https://www.proquest.com/docview/1906740390/abstract/C9393F7D7B34C9APQ/1

Respondek, L., Seufert, T., Stupnisky, R., & Nett, U. E. (2017). Perceived Academic Control and Academic Emotions Predict Undergraduate University Student Success: Examining Effects on Dropout Intention and Achievement. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00243

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Rocabado, G. A., Komperda, R., Lewis, J. E., & Barbera, J. (2020). Addressing diversity and inclusion through group comparisons: A primer on measurement invariance testing. *Chemistry Education Research and Practice*, *21*(3), 969–988. https://doi.org/10.1039/D0RP00025F

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews. *Annual Review of Psychology*, *52*(1), 59–82. https://doi.org/10.1146/annurev.psych.52.1.59

Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, *134*(2), 223–246. https://doi.org/10.1037/0033-2909.134.2.223

Ruthig, J. C., Haynes, T. L., Stupnisky, R. H., & Perry, R. P. (2009). Perceived Academic Control: Mediating the effects of optimism and social support on college students' psychological health. *Social Psychology of Education*, *12*(2), 233–249. https://doi.org/10.1007/s11218-008-9079-6

Rutkowski, L., & Svetina, D. (2014). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educational and Psychological Measurement*, *74*(1), 31–57. https://doi.org/10.1177/0013164413498257

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. Psychological Bulletin, 143(6), 565–600. https://doi.org/10.1037/bul0000098

Senko, C. (2019). When do mastery and performance goals facilitate academic achievement? *Contemporary Educational Psychology*, *59*, 101795. https://doi.org/10.1016/j.cedpsych.2019.101795

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement Goal Theory at the Crossroads: Old Controversies, Current Challenges, and New Directions. *Educational Psychologist*, *46*(1), 26–47. https://doi.org/10.1080/00461520.2011.538646

Seymour, E., & Hunter, A.-B. (Eds.). (2019). *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education*. Springer International Publishing. https://doi.org/10.1007/978-3-030-25304-2

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422

Smeding, A., Darnon, C., Souchal, C., Toczek-Capelle, M.-C., & Butera, F. (2013). Reducing the Socio-Economic Status Achievement Gap at University by Promoting Mastery-Oriented Assessment. *PLOS ONE*, *8*(8), e71678. https://doi.org/10.1371/journal.pone.0071678

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stornes, T., & Bru, E. (2011). Perceived motivational climates and self-reported emotional and behavioural problems among Norwegian secondary school students. *School Psychology International*, *32*(4), 425–438. https://doi.org/10.1177/0143034310397280

Strayhorn, T. L. (2018). *College Students' Sense of Belonging: A Key to Educational Success for All Students* (2nd ed.). Routledge. https://doi.org/10.4324/9781315297293

Strenta, C. A., Elliott, R., Adair, R., Matier, M., & Scott, J. (1994). Choosing and leaving science in highly selective institutions. *Research in Higher Education*, *35*(5), 513–547. https://doi.org/10.1007/BF02497086

Stupnisky, R. H., Renaud, R. D., Daniels, L. M., Haynes, T. L., & Perry, R. P. (2008). The Interrelation of First-Year College Students' Critical Thinking Disposition, Perceived Academic Control, and Academic Achievement. *Research in Higher Education*, *49*(6), 513–530. https://doi.org/10.1007/s11162-008-9093-8

Stupnisky, R. H., Renaud, R. D., Perry, R. P., Ruthig, J. C., Haynes, T. L., & Clifton, R. A. (2007). Comparing self-esteem and perceived control as predictors of first-year college students' academic achievement. *Social Psychology of Education*, *10*(3), 303–330. https://doi.org/10.1007/s11218-007-9020-4

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-Group Invariance with Categorical Outcomes Using Updated Guidelines: An Illustration Using Mplus and the lavaan/semTools Packages. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 111–130. https://doi.org/10.1080/10705511.2019.1602776

Tatapudy, S., Boukouzis, K., Kayal, I., Socko, R., Rajesh, M., Theobald, E., Not just the curve: Student experiences of competition in introductory STEM classrooms (in revision).

Tatapudy, S., Boukouzis, K., Kayal, I. S., Socko, R., Rajesh, M., & Theobald, E. J. (2025). Perceived Inequities in STEM Classes Make Them Feel Competitive. *CBE—Life Sciences Education*, *24*(4), ar45. https://doi.org/10.1187/cbe.24-02-0084

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483. https://doi.org/10.1073Ar/pnas.1916903117

Thoman, D. B., Muragishi, G. A., & Smith, J. L. (2017). Research Microcultures as Socialization Contexts for Underrepresented Science Students. *Psychological Science*, *28*(6), 760–773. https://doi.org/10.1177/0956797617694865

Tinto, V. (2010). From Theory to Action: Exploring the Institutional Conditions for Student Retention. In J. C. Smart (Ed.), *Higher Education: Handbook of Theory and Research: Volume 25* (pp. 51–89). Springer Netherlands. https://doi.org/10.1007/978-90-481-8598-6_2

Urhahne, D., & Wijnia, L. (2023). Theories of Motivation in Education: An Integrative Framework. *Educational Psychology Review*, *35*(2), 45. https://doi.org/10.1007/s10648-023-09767-9

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

von Hippel, P. T. (2007). Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. *Sociological Methodology*, *37*(1), 83–117. https://doi.org/10.1111/j.1467-9531.2007.00180.x

Wallwey, C. J. (2022). *Transcending Engineering's 'Weed-Out' Culture by Investigating the Impacts of Classroom Feedback* [The Ohio State University]. https://etd.ohiolink.edu/acprod/odb_etd/etd/r/1501/10?clear=10&p10_accession_num=osu1650376124141154

Walton, G. M., & Cohen, G. L. (2011). A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students. *Science*, *331*(6023), 1447–1451. https://doi.org/10.1126/science.1198364

Wang, J., & Wang, X. (2019). *Structural Equation Modeling: Applications Using Mplus*. John Wiley & Sons.

Weis, L. (2016). Positioning for Elite and Quasi-Elite Colleges and Universities in the United States: Parent and Student Strategies for "Maintaining Advantage" in New Economic and Postsecondary Context. In S. Slaughter & B. J. Taylor (Eds.), *Higher Education, Stratification, and Workforce Development: Competitive Advantage in Europe, the US, and Canada* (pp. 271–288). Springer International Publishing. https://doi.org/10.1007/978-3-319-21512-9_14

Weston, T. J., Seymour, E., Koch, A. K., & Drake, B. M. (2019). Weed-Out Classes and Their Consequences. In E. Seymour & A.-B. Hunter (Eds.), *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education* (pp. 197–243). Springer International Publishing. https://doi.org/10.1007/978-3-030-25304-2_7

Yoon, M., & Lai, M. H. C. (2018). Testing Factorial Invariance With Unbalanced Samples. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(2), 201–213. https://doi.org/10.1080/10705511.2017.1387859

You, S., Hong, S., & Ho, H. (2011). Longitudinal Effects of Perceived Control on Academic Achievement. *The Journal of Educational Research*, *104*(4), 253–266. https://doi.org/10.1080/00220671003733807

Zuberi, T. (2001). *Thicker Than Blood: How Racial Statistics Lie*. U of Minnesota Press.

Supplement

The University of Washington determined this study to be exempt from IRB review (University of Washington Study 00010826). For most institutions, we obtained approvals and permissions from either departmental, college, or university administrators, or Institutional Review Boards, following recommendation by the University of Washington Human Subjects Division. For the institutions for whom review was necessary, this study was considered exempt under the following study numbers: Eastern Michigan University UHSRC-FY21-22-139, Florida International University IRB-20-0370-AM01, New Mexico State University 2206001150, University of Minnesota STUDY00020526, University of South Florida Study003772, Washington University in St. Louis 202112131, and Eastern New Mexico University 2324013.

Table S1. *Questions and response statistics from our survey regarding competition, belonging, and control. Means are based on the original data, before the collapsing of sparse cells.*

Question	Wording	Mean response	Mode response	Mode frequency	Notes
comp_1	The professor seems to pit students against each other in a competitive manner in this class.	1.7	1	56%	Collapsed 4, 5 and 6
comp_2	Students tend to be very competitive with each other in this class.	2.3	1	35%	Collapsed 5 and 6
comp_3	Students are encouraged to outperform others.	2.1	1	40%	
comp_4	The professor frequently compares students to each other.	1.5	1	66%	Collapsed 4, 5, and 6
comp_5	The amount of recognition students get in this class depends on how their performance compares to other students.	1.9	1	50%	Collapsed 5 and 6
comp_6	Everybody is concerned with finishing at the top of the class.	2.6	1	29%	Collapsed 5 and 6
comp_7	Students in this class frequently compare their performance to each other.	2.8	1	26%	
comp_8	The grading in this course is competitive.	2.8	1	27%	

bel_1	I feel comfortable in this class.	4.8	5	44%	Collapsed 1 and 2
bel_2	I feel accepted in this class.	4.9	5	48%	Collapsed 1,2 and 3
bel_3	I feel like I can be myself in this class.	4.8	5	44%	Collapsed 1 and 2
bel_4	I feel like I belong in this class.	4.7	5	41%	Collapsed 1 and 2
ctrl_1	I have a great deal of control over my academic performance in this class.	4.7	5	41%	Collapsed 1 and 2
ctrl_2	The more effort I put into this class, the better I do.	5.0	6	38%	Collapsed 1 and 2
ctrl_3	No matter what I do, I can't seem to do well in this course.*	4.1	5	29%	
ctrl_4	I see myself as largely responsible for my performance in this course.	4.9	5	45%	Collapsed 1, 2, and 3
ctrl_5	How well I do in this course is often the "luck of the draw".*	4.5	5	34%	Collapsed 1 and 2
ctrl_6	There is little I can do about my performance in this course.*	4.8	5	39%	Collapsed 1, 2, and 3
ctrl_7	When I do poorly on an assignment or assessment in this course, it's usually because I haven't given it my best effort.	4.1	5	32%	
ctrl_8	My grades are basically determined by things beyond my control and there is little I can do to change that.*	4.7	5	35%	Collapsed 1 and 2

*Negatively coded, but we reversed the responses (7 - original response)

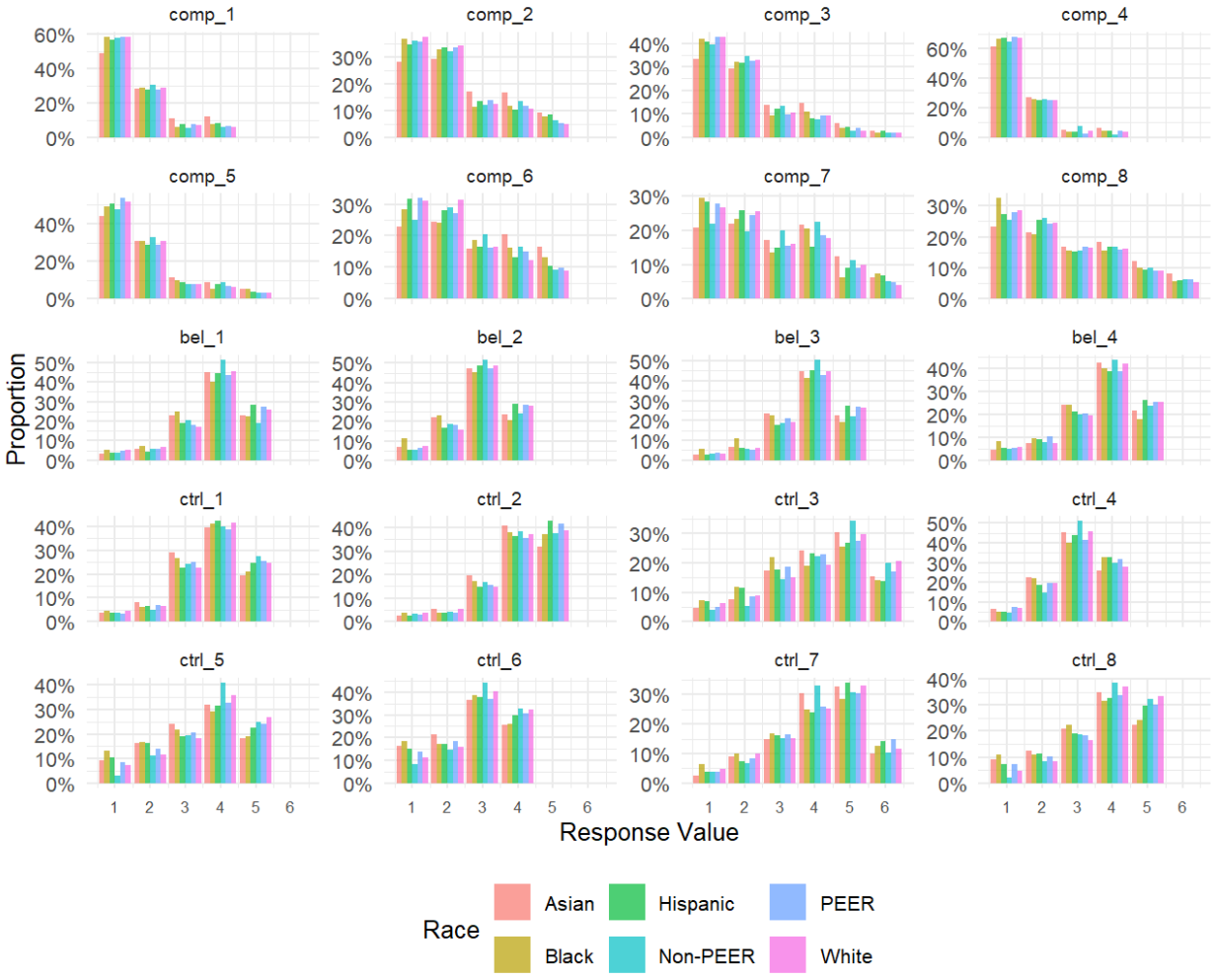

Fig. S1. *Survey responses. Distribution of responses (after collapsing sparse cells) to each survey question disaggregated by students' racial identities.*

Table S2. *Loadings of CFA models. Loadings of CFAs using the WLSMV estimator. All (SG-CFA) indicates the measurement model for all students, not disaggregated by race. The following six columns are the loadings for each racial group within a single model of configural invariance. The final column represents the loadings constrained to be equal among groups in the scalar model. comp_1 = competition item 1, bel = belonging, ctrl = control. Correlated residuals are not constrained to be equal among racial categories in the scalar model.*

	All (SG-CFA)	All (SG-CFA, std.)	Asian (config)	Black (config)	Hispanic (config)	Non-PEER (config)	PEER (config)	White (config)	All (scalar)
comp_1	1.000	0.887	1.000	1.000	1.000	1.000	1.000	1.000	1.000
comp_2	0.627	0.769	0.688	0.621	0.872	0.873	0.570	0.656	0.682
comp_3	0.894	0.864	0.884	1.102	0.998	1.120	0.990	1.029	0.896
comp_4	0.945	0.875	0.786	0.825	1.171	0.997	1.135	0.983	0.789
comp_5	0.689	0.797	0.648	0.634	0.753	0.804	0.742	0.731	0.643
comp_6	0.419	0.626	0.485	0.466	0.541	0.433	0.501	0.400	0.469
comp_7	0.395	0.604	0.398	0.436	0.521	0.424	0.477	0.422	0.396
comp_8	0.550	0.726	0.601	0.655	0.599	0.713	0.669	0.604	0.596
bel_1	1.000	0.919	1.000	1.000	1.000	1.000	1.000	1.000	1.000
bel_2	1.193	0.941	1.194	1.440	1.523	1.389	1.305	1.161	1.182
bel_3	0.978	0.915	1.060	1.356	1.522	1.321	1.181	0.922	1.057
bel_4	0.899	0.902	0.967	0.962	0.867	1.005	1.019	0.892	0.970
ctrl_1	1.000	0.829	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ctrl_2	0.884	0.795	0.984	1.216	1.198	0.783	0.885	0.881	0.994
ctrl_3	0.629	0.682	0.794	0.664	0.515	0.791	0.559	0.652	0.767
ctrl_4	0.629	0.661	0.559	0.763	0.539	0.554	0.730	0.635	0.561
ctrl_5	0.600	0.665	0.637	0.489	0.508	0.707	0.506	0.646	0.613
ctrl_6	0.767	0.751	0.879	0.794	0.683	0.760	0.667	0.811	0.835
ctrl_7	0.388	0.499	0.387	0.519	0.377	0.434	0.392	0.422	0.417
ctrl_8	0.647	0.692	0.684	0.637	0.594	0.613	0.556	0.719	0.629
ctrl_3 w/ ctrl_5	0.394	0.394	0.390	0.460	0.439	0.057	0.481	0.363	Varies
ctrl_3 w/ ctrl_6	0.370	0.370	0.300	0.440	0.436	0.222	0.394	0.362	-
ctrl_3 w/ ctrl_8	0.276	0.276	0.285	0.229	0.348	-0.055	0.350	0.231	-
ctrl_5 w/ ctrl_6	0.561	0.561	0.541	0.604	0.655	0.388	0.603	0.544	-
ctrl_5 w/ ctrl_8	0.479	0.479	0.492	0.525	0.462	0.190	0.549	0.444	-
ctrl_6 w/ ctrl_8	0.490	0.490	0.532	0.443	0.487	0.362	0.534	0.445	-
comp_2 w/ comp_6	0.327	0.327	0.257	0.254	0.277	0.243	0.305	0.332	-
comp_2 w/ comp_7	0.362	0.362	0.379	0.245	0.295	0.309	0.413	0.314	-

comp_6 w/ comp_7	0.572	0.572	0.597	0.610	0.550	0.535	0.522	0.555	-
comp w/ bel	-1.712	-0.384	-1.981	-1.402	-1.053	-1.436	-1.135	-1.544	-
comp w/ ctrl	-1.514	-0.533	-1.455	-1.294	-1.065	-1.244	-1.346	-1.402	-
bel w/ ctrl	2.118	0.614	1.985	1.215	2.096	2.497	1.659	2.249	-

Table S3. *Comparison of latent means to grand mean. A three-factor scalar CFA produced latent means for competition, belonging, and control. These unstandardized latent means (different from in Fig. 2) are compared to the grand mean (i.e., the weighted mean of all these groups or the mean of all students in the dataset), so a negative value means that the group mean is below the grand mean. In parentheses are the values standardized by the overall standard deviation of the entire sample (also different from in Fig. 2). * indicates significance at alpha = 0.05.*

	Asian	*Black*	*Hispanic*	*Non-PEER*	*PEER*	*White*
Mean competition	***0.437* (0.229)***	*-0.129 (-0.067)*	*-0.100 (-0.052)*	*0.110 (0.058)*	*-0.170 (-0.089)*	*-0.163 (-0.085)*
Mean belonging	*-0.164 (-0.076)*	***-0.577* (-0.267)***	*0.171 (0.079)*	*-0.072 (0.033)*	*0.081 (0.038)*	*0.086 (0.040)*
Mean control	***-0.240* (-0.192)***	*-0.220 (-0.176)*	*0.033 (0.026)*	*0.228 (0.182)*	*0.054 (0.043)*	*0.109 (0.087)*

Table S4. Indirect effect 95% CI sensitivity analysis. *Indirect effects in our main model (Model 1) were calculated using the delta method: c=a*b. However, although a and b are normally distributed, the product of a and b may not be normal, thus the standard errors and p value may not be accurate. We did a sensitivity analysis by using bootstrapped confidence intervals (1000 iterations) for the standardized indirect effects (Model 3). However, we could not bootstrap while using clustering, so we compared to a no cluster + delta model (Model 2). These three models are generally in agreement. Model 2 and Model 3 are especially in agreement, indicating that bootstrapping led to similar results to the delta method in this case.*

Model	**Indirect via Belonging**	**Indirect via Control**	**Total Indirect**
Model 1: Clustered + Delta	[-0.019, 0.014]	[-0.206, -0.132]	[-0.204, -0.139]
Model 2: No Cluster + Delta	[-0.006, 0.019]	[-0.207, -0.167]	[-0.197, -0.165]
Model 3: No Cluster + Bootstrap	[-0.006, 0.021]	[-0.211, -0.163]	[-0.199, -0.163]

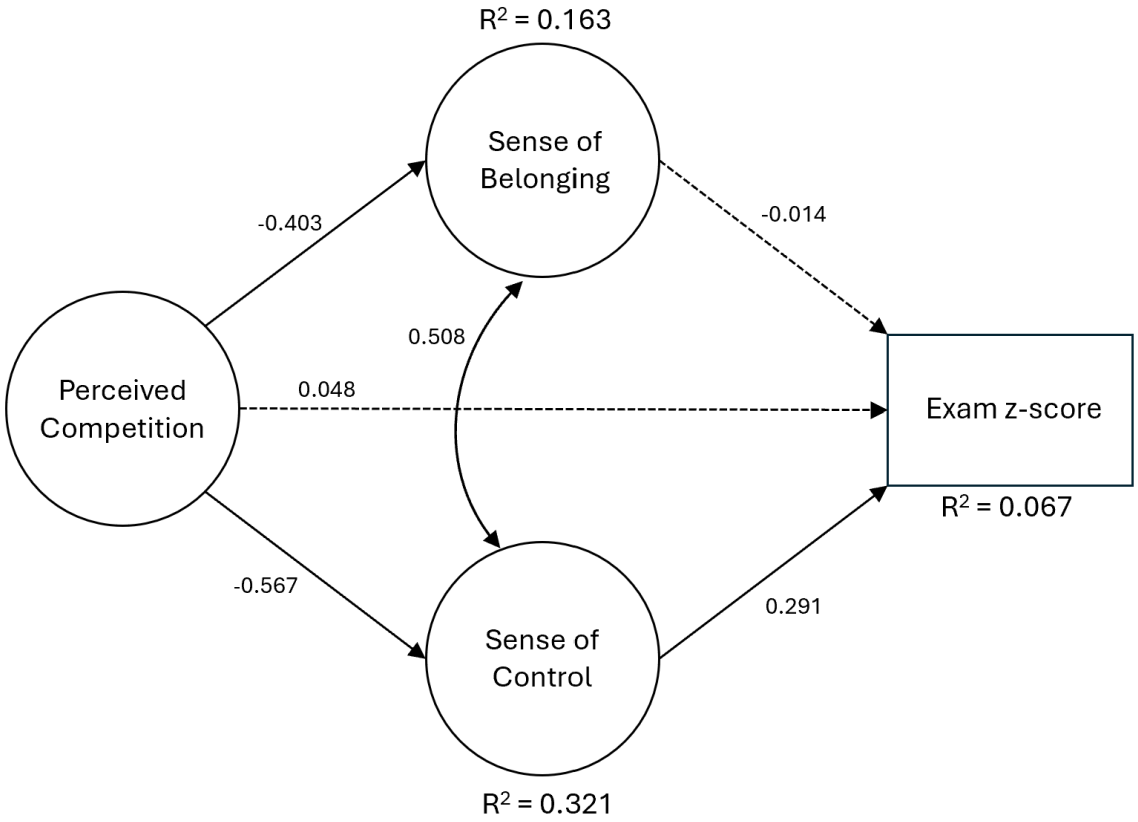

Fig. S2. *Non-disaggregated SEM results (Biology classes only). Like Fig. 1 in the main text but the SEM was only on biology classes.*

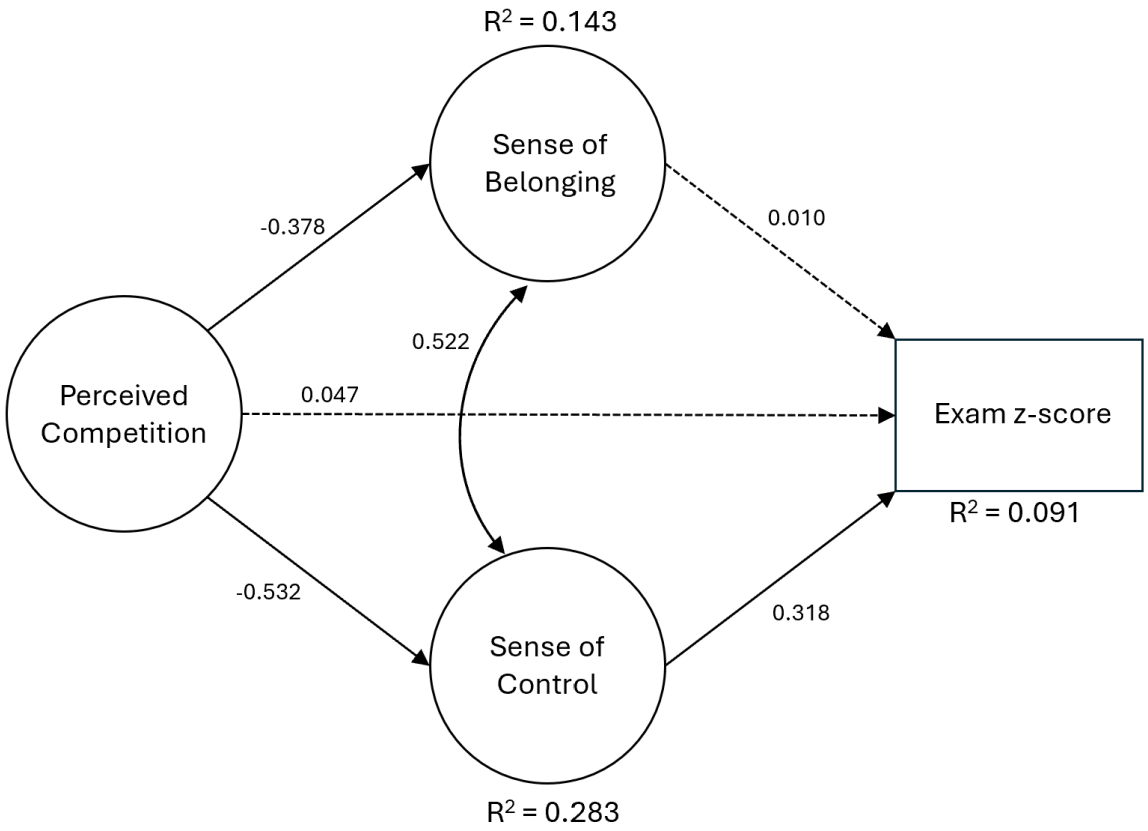

Fig. S3. *Survey responses. Like Fig. 1 in the main text but no response categories were collapsed.*