1 **Abstract**

2 1. Temperature and light cues interact to control many biological processes. Experiments give

3 researchers the ability to manipulate these environmental cues independently, and can be

4 designed to robustly quantify their individual and interactive effects on any particular bio-

5 logical activity. Such experiments have produced important insights into the environmental

6 controls on numerous biological processes in both plant and animal taxa across terrestrial and

7 aquatic environments. Testing the interactive effects of multiple environmental cues, how-

8 ever, requires experimental treatments to be fully independent; any unmeasured experimental

9 covariation among treatments can result in incorrect conclusions.

10 2. Using a database of controlled environment experiments on the spring phenology of woody

11 plants as a case study, we highlight how a common experimental set-up, designed to parse

12 the interactive effects of temperature and photoperiod on time to budburst, introduces a

13 latent experimental covariation of these treatments by coupling photo- and thermo- period-

14 icity. Using data simulations, algebraic corrections and a comparative analysis of published

15 experiments, we demonstrate how this unmeasured experimental covariation biases statistical

16 inference regarding the relative contribution of light and temperature cues to phenological

17 variation.

18 3. We identify this experimental covariation in more than 40% of published phenology studies

19 that manipulate photoperiod. Our analyses demonstrate that the coupling of thermo- and

20 photo- periodicity results in the overestimation of the effect of photoperiod, the underesti-

21 mation of forcing effects, and misleading conclusions about their interactions on phenology.

22 This may, in part, explain why the significance of photoperiod cues for spring phenology is

23 currently debated in the literature.

24 4. Accurate forecasting of how varying environmental conditions will impact the dynamics of

25 biological events requires accurately quantifying cue responses. To this end, we present several

26 options for statistical corrections and alternative experimental designs that can provide more

27 robust estimates of the relative effects of temperature and photoperiod on phenology, and

28 many other biological processes controlled by temperature and light.

29 **Keywords:** forcing, full-factorial, growth chamber, light, phenology, photoperiod, temperature,

30 thermoperiod

31

32 **Introduction**

33 Across the tree of life, temperature and light availability shape a number of important biological

34 processes including growth and metabolic rates ([MacLean & Gilchrist](#_bookmark30), [2019](#_bookmark30)), sex determination

35 ([Brown *et al.*](#_bookmark5), [2014](#_bookmark5)), acclimatization to seasonal environments ([Hamilton *et al.*](#_bookmark24), [2016](#_bookmark24)) and the

36 timing of life cycle transitions (i.e., phenology, [Forrest & Miller-Rushing](#_bookmark20), [2010](#_bookmark20)). These biological

37 responses in turn dictate broad-scale ecological processes and patterns ranging from biogeochemical

38 cycling ([Piao *et al.*](#_bookmark32), [2007](#_bookmark32)) to species range limits ([Chuine & Beaubien](#_bookmark12), [2001](#_bookmark12)). Characterizing the

39 specific dynamics of how these environmental factors synergistically affect biological processes across

40 a wide range of taxa has become even more important as anthropogenic global change continues to

41 expose organisms to novel environmental conditions ([Pörtner & Farrell](#_bookmark34), [2008](#_bookmark34)).

42 Because temperature and light availability often co-vary in the field (for example, in most temperate

43 ecosystems, daylength and temperature both increase as the season progresses, [Rosenberg](#_bookmark36), [1974](#_bookmark36)),

44 it can be diﬀicult to disentangle their relative contributions to biological processes. In contrast,

45 experimental manipulations of climate variables in artificial environments can mechanistically char-

46 acterize biological responses to environmental fluctuations ([Ettinger *et al.*](#_bookmark18), [2020](#_bookmark18); [Primack *et al.*](#_bookmark35),

47 [2015](#_bookmark35)). Researchers have used controlled environments of all shapes and sizes to this end ([Downs](#_bookmark16),

48 [1980](#_bookmark16)); these efforts have greatly advanced our collective understanding of the fundamental biology

49 of a wide variety of organisms and ability to predict ecological and evolutionary responses to current

50 and future climate change ([Stewart *et al.*](#_bookmark40), [2013](#_bookmark40)).

51 However, controlled environment experiments have their own challenges. Experimentalists must

52 balance biological realism with robust inference, experimental effort with statistical power, and

53 account for the effects of unmanipulated or unmeasured variables ([Scheiner & Gurevitch](#_bookmark38), [2001](#_bookmark38)).

54 Because biological responses to the environment are generally the product of complex interactions

55 between multiple environmental signals ([Casal](#_bookmark8), [2002](#_bookmark8)), seemingly small choices about experimental

56 designs can generate significant differences in outcomes. Experimental treatments are rarely stan-

57 dardized among researchers, even within disciplines ([Wolkovich *et al.*](#_bookmark43), [2022](#_bookmark43)), and these complexities

58 may in part contribute to the many discrepancies between experimental studies and observational

59 data ([Poorter *et al.*](#_bookmark33), [2016](#_bookmark33)). Even with these limitations, controlled environment studies remain

60 a powerful tool to mechanistically assess organismic responses to the environment, provided that

61 the implications of treatment designs are well understood and well matched with the scope of the

62 research question.

63 As technology advances and experiments become more complex, researchers can manipulate more

64 variables and multiple axes of variation (e.g., temperature, amplitude, periodicity, wavelength) at

65 the same time. Yet these efforts may present a tradeoff between biological realism and robust in-

66 ference. Through investigating the literature on experiments with plant phenology, we show that

67 experiments that manipulate both photo- and thermo- periodicity often introduce a latent experi-

68 mental covariation between light and temperature treatments, which may misrepresent the effects

69 of each of these environmental variables and the interaction between them. We begin by briefly

70 detailing how temperature and light treatments are generally applied in phenology experiments and

71 review the minimum experimental elements required to robustly test interactions between two or

72 more environmental variables. We then detail the problem of inference that can arise when manip-

73 ulating the periodicity of both temperature and light in experiments, and demonstrate the extent

74 to which this is an issue through data simulations, a mathematical correction, and a comparative

75 analysis of published experiments. Finally, we conclude by outlining methods that correct for this

76 issue, along with alternative experimental designs that can overcome the problem of periodicity.

77 While our case study deals with phenology of temperate woody plants, it provides insights into a

[78](#_bookmark0) number of other systems with parallel issues. Studies of aquatic algae ([Xu *et al.*](#_bookmark46), [2019](#_bookmark46)), insects ([And-](#_bookmark0)

79 [uaga *et al.*](#_bookmark0), [2018](#_bookmark0)), amphibians ([Wright & Bruni](#_bookmark45), [2004](#_bookmark45)) and fish ([López‐Olmeda & Sánchez‐Vázquez](#_bookmark29),

80 [2009](#_bookmark29)) have similarly struggled to disentangle the effects of thermo- and photo- periodicity. Thus,

81 we believe the potential problems and solutions we present here are broadly applicable to studies

82 on other organisms and biological processes that utilize temperature and light signals.

# 83 Case study: Estimating phenological cues from experiments

84 Decades of experimental work in controlled environments have demonstrated that temperature

85 (both cool temperatures in fall/winter and warming temperatures in spring) and photoperiod are

86 the primary phenological cues for plants in the temperate/boreal zones ([Ettinger *et al.*](#_bookmark18), [2020](#_bookmark18)). While

87 exposure to cool winter temperatures (chilling) strongly impacts phenology ([Laube *et al.*](#_bookmark28), [2014](#_bookmark28)), we

88 focus here on warm temperature and light treatments, because controlled chilling treatments with

89 light are uncommon ([Wolkovich *et al.*](#_bookmark43), [2022](#_bookmark43)). Choices about how to apply warm temperature and

90 light treatments, in particular, can compromise inference on their effects, so we focus on these two

91 cues.

92 While a large variety of experimental designs have been used to study plant phenology, generally

93 experiments tend to manipulate two major axes of light and warm temperature variation:

94 1. Intensity: The amount or quality of a variable. Here we define temperature intensity as the

95 amount of heat present in the system (measured in degrees). In the phenology literature this

96 measurement is generally referred to as forcing. We define light intensity as the luminosity or

97 irradiance present in the system (measured in lumens or watts).

98 2. Periodicity: The interval at which the intensity of the variable is applied. Hereafter, we refer

99 to the periodicity of light as photoperiod (often used synonymously with “daylength”) and

100 the periodicity of temperature as thermoperiod.

[101](#_bookmark42) For phenology, photoperiodicity is generally considered the primary light cue for plants ([Way &](#_bookmark42)

102 [Montgomery](#_bookmark42), [2015](#_bookmark42)), (though regarding light intensity and phenology see [Brelsford & Robson](#_bookmark4), [2018](#_bookmark4);

103 [Cober *et al.*](#_bookmark14), [1996](#_bookmark14)). For temperature, conventionally both intensity and periodicity drive pheno-

104 logical activity and several metrics (e.g. growing degree hours, thermal sums, growing degree days)

105 that combine these two axes have been developed ([Gu](#_bookmark23), [2016](#_bookmark23)). The importance of thermo-intensity

106 and periodicity is well supported; under natural conditions diurnal temperature fluctuations in tem-

107 perate regions can be quite large in the spring, and studies have found that diurnal temperature

108 variation strongly influences plant phenology ([Burghardt *et al.*](#_bookmark7), [2016](#_bookmark7)). In fact, even if thermoperi-

109 odicity is not an explicit treatment variable (i.e., manipulated systematically), incorporating it in

[110](#_bookmark11) experiments can be essential for translating experimental results into real world predictions ([Chiang](#_bookmark11)

111 [*et al.*](#_bookmark11), [2020](#_bookmark11)).

112 Like many other biological processes, recent advances have demonstrated that plant phenological

113 responses are nonlinear, due largely to interactions between cues ([Wolkovich *et al.*](#_bookmark43), [2022](#_bookmark43); [Fu *et al.*](#_bookmark21),

114 [2015](#_bookmark21)), highlighting the need for experiments designed to evaluate the strength of these interactions.

115 To have the statistical power to partition the individual and interactive effects of two or more

116 variables, an experiment must:

117 1. Have a minimum of two treatment levels of at least two variables.

118 2. Treatment levels must be full factorial (Fig. [1](#_bookmark49)a.). Full factorial designs are both balanced

119 (Fig. [1](#_bookmark49)b.) and orthogonal (Fig. [1](#_bookmark49)c.); meaning that all possible treatment combinations are

120 applied and each treatment is independent of all others ([Cheng](#_bookmark9), [2016](#_bookmark9)).

121 These two critical elements may seem obvious but are conspicuously absent from many published

122 studies. Of the 136 studies contained in a published database of woody plant phenological ex-

123 periments (OSPREE: Observed Spring Phenological Responses in Experimental Environments,

124 [Wolkovich *et al.*](#_bookmark44), [2019](#_bookmark44)), a recent study by [Wolkovich *et al.*](#_bookmark43)([2022](#_bookmark43)) found that only 37% of the

125 studies that manipulated more than one variable did so with a design that was both balanced and

126 orthogonal. But even experiments that are designed to be full factorial frequently violate the as-

127 sumption of orthogonality when both photo- and thermo- periodicity are built into experiments.

128 We detail this problem below.

# 129 The problem of periodicity

130 A common approach in phenology experiments that seems to balance prior knowledge about the

131 underlying physiology of phenology, biological realism and experimental inference is to vary pho-

[132](#_bookmark37) toperiodicity, and thermal intensity and periodicity (e.g., [Flynn & Wolkovich](#_bookmark19), [2018](#_bookmark19); [Sanz-Perez](#_bookmark37)

133 [*et al.*](#_bookmark37), [2009](#_bookmark37); [Basler & Körner](#_bookmark1), [2014](#_bookmark1)). This design could include any number of treatment levels for

134 each variable (e.g., 8, 12, 16...*n* hours of photoperiod and 20/10*◦*C, 22/12*◦*C, 25/15*◦*C...*nday/nnight*

135 day/night temperatures). To consider a simple example, we use a hypothetical experiment with

136 two treatment levels of each variable. Consider a basic experiment that includes, at minimum, a

137 long (16 hours) and short (8 hours) photoperiod treatment and a high (25/15*◦*C day/night) and

138 low (20/10*◦*C day/night) forcing treatment. In this case, the thermoperiodicity is not an explicit

139 treatment (both high and low temperature treatments use a diurnal fluctuation of 10 *◦*C), and is

140 simply incorporated in the design to enhance biological realism. At first glance, this design appears

141 to meet the criteria of a full factorial design, multiple treatment levels that are balanced and or-

142 thogonal, with high/low temperature treatments (mean 20*◦*C and 15*◦*C respectively) and long/short

143 photoperiod treatments applied in all possible combinations.

144 Yet the orthogonality of this design is based on the assumption of a 12 hour thermoperiod. If, rather

145 the thermoperiod is coupled with the photoperiod, the temperature treatment is non-orthogonal

146 because the daily mean temperature of the long/high treatment will be higher than that of the

147 short/high treatment, and the long/low treatment slightly warmer than the short/low. We refer

148 to this experimental set-up as a coupled design (i.e. thermoperiod and photoperiod are coupled

149 with each other). Coupled designs introduce an experimental covariation between photoperiod and

150 forcing treatments. This experimental covariation is clearly illustrated when temperature treatment

151 levels are converted to thermals sums. We calculate thermal sums (also called growing degree hours),

152 by multiplying hourly temperatures above a certain base temperature threshold by the number of

153 hours for which they are applied over a 24 hour period ([Parent *et al.*](#_bookmark31), [2019](#_bookmark31)). For example, given a

154 base temperature of 0*◦*C, a low forcing treatment of 20/10*◦*C day/night accrues 400 thermal units

155 per 24 hours when crossed with the long (16 hour) photoperiod treatment and only 320 thermal units

156 when crossed with the short (8 hour) photoperiod treatment. While this experimental covariation

157 among the photoperiod and temperature treatments is biologically realistic, it makes it statistically

158 impossible to differentiate the independent and interactive effects of temperature and photoperiod

159 on any given biological process.

160 This problem of inference that arises from the experimental covariation of thermo- and photo-

161 periodicity is not limited only to studies seeking to directly compare the effects of photoperiod and

162 forcing; it applies in any study evaluating the influence of photoperiod on biological activity, even if

163 it is the only manipulated cue. Experimentally isolating the effect of photoperiod assumes that all

164 other environmental variables are held constant. Similar to the case described above, the coupling

165 of photoperiod and thermoperiod in an experiment where forcing is intended to be a consistent,

166 background condition (e.g., two or more levels of photoperiod treatments (e.g. 8, 12, and 16 hours),

167 all at a background temperature of 20/10*◦*C day/night) would yield a situation in which longer

168 photoperiod treatments were also receiving more—unmeasured—heating than shorter photoperiod

169 treatments. In this case, some amount of the perceived photoperiod effect is due to the latent,

170 increased forcing, and the experiment will not isolate the true effect of photoperiod.

171 We queried the OSPREE database to identify experiments that applied different day and night

172 temperatures in their studies without designating diurnal temperature variation as an explicit ex-

173 perimental treatment. Of the 51 experiments in the OSPREE database that manipulated pho-

174 toperiod experimentally, up to 43% of them appear to include an experimental covariation with

175 thermoperiod. Of the 18 studies that manipulated both photoperiod and temperature interactively,

176 we found that up to 55% of them appear to have this issue, suggesting that the true interactive

177 effects of these cues on spring phenology is quite poorly characterized. This may be in part why

178 the relative contribution of temperature and photoperiod cues to spring phenology remains a con-

[179](#_bookmark27) tentious debate in the phenology literature ([Koerner & Basler](#_bookmark26), [2010](#_bookmark26); [Chuine *et al.*](#_bookmark13), [2010](#_bookmark13); [Körner &](#_bookmark27)

180 [Basler](#_bookmark27), [2010](#_bookmark27)).

# 181 Periodicity and inference

182 If the lack of orthogonality introduced to experiments when photoperiod and thermoperiod are

183 coupled is overlooked, regression models will always overestimate the photoperiod effect and un-

184 derestimate the forcing effect (Fig. [2](#_bookmark50)a,b.). This is because forcing is the variable with latent,

185 unmeasured variation. In the case of phenology, this is particularly significant because studies

186 repeatedly suggest that forcing is a more dominant cue than photoperiod for spring phenology

187 ([Chuine *et al.*](#_bookmark13), [2010](#_bookmark13); [Zohner *et al.*](#_bookmark48), [2016](#_bookmark48); [Gauzere *et al.*](#_bookmark22), [2019](#_bookmark22)). The influence of this experimen-

188 tal covariation of periodicity on generating incorrect estimates of temperate and photoperiod cue

189 effect-sizes pervades experiments with any number of treatment levels (see Fig. S1 for an example

190 with three treatment levels of forcing and photoperiod), and may be even more diﬀicult to identify

191 as experimental complexity increases.

192 If experiments are designed to quantify the interaction between photoperiod and forcing, here too,

193 the experimental covariation of periodicity will result in an erroneous estimation of the interaction.

194 (Fig. [2](#_bookmark50)c,d.). Our simulation depicts a particularly troublesome case where a true sub-additive

195 interaction is interpreted as a supra-additive one (Fig. [2](#_bookmark50)c,d.), however, other outcomes are possible.

196 Experimental covariation of light and temperature treatments due to coupling thermo- and photo-

197 periodicity will generally result in the incorrect estimation of the interaction term, but the exact

198 nature of this statistical issue depends on the sign and strength of the interaction.

199 We can attempt to estimate how much of a photoperiod effect is due to forcing in experiments where

200 they covary by making several major assumptions. First we assume that forcing and photoperiod

201 effects are additive and linear (i.e., there is no interaction). While this may not be true in nature, it

202 gives us insight into the potential effect of the experimental covariation of periodicity by allowing us

203 to solve algebraically for the separate effects of forcing and photoperiod. We replace the qualitative

204 factor (high/low forcing) by the quantitative effect of forcing (thermal sums) to properly account

205 for the difference in forcing between short and long photoperiods (see Supporting Information:

206 Estimating the effects of experimental periodicity covariance mathematically). Using the data from

207 one experiment that experimentally coupled thermo- and photo-period, [Flynn & Wolkovich](#_bookmark19) ([2018](#_bookmark19)),

208 we found that 33% of the published photoperiod effect of 4.5 days could be due to forcing.

209 Our algebraic solution cannot be as readily applied in experiments that assume photoperiod and

210 forcing interact. However, we can generally assess the scope of the problem of inference due to

211 experimental covariation of periodicity by comparing studies that used a coupled design to those with

212 alternative approaches. While we are aware of no experiments that explicitly compare the effects

213 of experimentally coupling vs. uncoupling photo- and thermo- periods, we identified two phenology

214 experiments that utilized many overlapping treatment levels and species from the same sampling

215 sites, however in one study, [Flynn & Wolkovich](#_bookmark19) ([2018](#_bookmark19)), photo- and thermo- period experimentally

216 co-vary, while in the other, [Buonaiuto & Wolkovich](#_bookmark6) ([2021](#_bookmark6)), photo- and thermo- period were varied

217 independently (see Supporting Information: Modeling Methods for details on treatment similarities

218 and differences between the studies). Comparing the cue estimates from these two studies offers

219 an opportunity to test our theoretical and mathematical predictions, and further understand the

220 uncertainty in cue estimates due to coupled periodicities.

221 We subset each dataset to include only the species shared among the two studies, and re-analyzed

222 the data using Bayesian hierarchical models to compare the difference in the photoperiod, forcing

223 and interaction estimates (see Supporting Information: Modeling Methods). We found that the

224 estimated differences in the mean response to photoperiod and forcing and their interactions among

225 study designs were on the same order as our predictions above for misestimated cue effects due to

226 experimental covariation between light and temperature treatments. We estimated a substantially

227 weaker (less negative) photoperiod effect, and marginally stronger forcing effect for the uncoupled

228 vs. coupled experimental design (Fig. [3](#_bookmark51)). The interaction term we estimated for the uncoupled

229 design was negative, suggesting the interaction between photoperiod and forcing is supra-additive,

230 while the estimated interactive effect from the coupled design was sub-additive (Fig. [3](#_bookmark51)).

231 Unlike in our simulations (Fig. [2](#_bookmark50)), in this comparison we cannot assess what the “true” effects

232 of these variables are. There are almost certainly other factors driving the differences between

233 these experiments. Both were conducted in different years, sampled different individuals from the

234 population, and used different methods for applying chilling pre-treatments ([Flynn & Wolkovich](#_bookmark19),

235 [2018](#_bookmark19); [Buonaiuto & Wolkovich](#_bookmark6), [2021](#_bookmark6)). However, because this comparison is well matched to our

236 predictions and prior knowledge about how temperature and photoperiod are expected to interact-

237 ing in phenology, we argue that the influence of experimental covariation on statistical inference is

238 apparent enough to take seriously.

239

240 **Paths Forward**

241 We have demonstrated that experiments that coupled thermoperiod and photoperiod cannot ro-

242 bustly differentiate the individual or interactive effects of temperature and photoperiod on spring

243 phenology (or any other biological process) due to an unmeasured experimental covariation among

244 temperature and light treatments. Given the paucity of interactive studies in the literature, it is

245 clear that we need more well designed studies to better characterize the effects of these cues. At

246 the same time, there are straight-forward statistical approaches for accounting for this experimental

247 covariation that can be adopted immediately and will rapidly improve scope of inference possible

248 with controlled environment experiments. Below we detail these approaches, and then offer several

249 generalized experimental designs that improve statistical orthogonality of controlled environment

250 experiments, which could be further developed and adjusted to fit the needs of experimentalists

251 across many sub-fields of ecology and evolutionary biology.

## 252 Effect-size inference and statistical corrections:

253 It may be that the experimental design that best balances environmental realism, statistical infer-

254 ence and translatability to observational studies are designs that continue to couple periodicity to

255 mimic natural systems. Fundamentally, simply recognizing the issues that arise when thermo- and

256 photoperiods are experimentally co-varied and accounting for this in interpreting effect-sizes and

257 reporting uncertainty is a powerful start for improving inference from experiments. This aware-

258 ness can be applied both forward and backward: to future experiments that seek to understand

259 the interactive effects of temperature and photoperiod and to synthesizing and interpreting the

260 near-century’s worth of research in this area that has already been published.

261 The growing awareness of this issue has prompted the development of several simple statistical

262 corrections to deal with this experimental covariation. Some recent studies adjust temperature

263 effect-sizes in their statistical models by daylength treatments (e.g., [Ettinger *et al.*](#_bookmark18), [2020](#_bookmark18)). For

264 example, if a day/night forcing treatment of 25/15*◦*C was applied in conjunction with 8 and 16 hour

265 photoperiod treatment levels, the mean daily temperature of the forcing level can be weighted by

266 hours for which it was applied, in this case resulting in forcing treatments with mean temperatures of

267 18.3*◦*C and 21.6*◦*C respectively. This approach does not remove the covariation between temperature

268 and photoperiod (i.e., the higher photoperiod treatment is still getting more heat that the lower

269 photoperiod treatment, e.g, Fig. [1](#_bookmark49)b), but this covariation is no longer latent, and can be accounted

270 for in a regression model. This simple approach could be adopted by any experimentalist and will

271 substantially increase the utility of such experiments for ecological forecasting.

## 272 Experimental Re-designs:

273 For researchers interested in taking on this problem of periodicity head-on in their experiments, there

274 are several experimental designs that can either eliminate the problem of experimental covariation

275 of photo- and thermo-period entirely, or more robustly address it at the experimental stage. Below,

276 we provide general details about alternative experimental designs with representative examples in

277 Fig. [4](#_bookmark52). As in our previous examples, here we depict experiments with two levels of two experimental

278 variables (photoperiod and temperature), but importantly, these generalized schemes can be readily

279 adapted for experiments with any number of treatment levels number of variables, as long as they

280 are full-factorial. Further, these designs could be adapted for any experimental variables for which

281 both intensity and periodicity can be manipulated (e.g. light, humidity, heat or freeze shock etc).

282 1. **Manipulate photoperiod and temperature intensity with no thermoperiodicity**.

283 The simplest way to evaluate the individual and combined effects of temperature and pho-

284 toperiod in experimental settings is to remove thermoperiodicity from studies entirely by

285 maintaining constant day/night temperatures within temperature treatments (Fig. [4](#_bookmark52)a.). This

286 approach allows for the maintenance of statistical orthogonality across treatment combina-

287 tions. The main drawback is that this design sacrifices the realism of diurnal temperature

288 variation, which may make it more diﬀicult to translate estimates from experiments to real-

289 world applications. However, many aspects of physiology and development do not appear to

290 respond explicitly to diurnal temperature variation (e.g., [Hellmers](#_bookmark25), [1966](#_bookmark25); [Warrington *et al.*](#_bookmark41),

291 [1977](#_bookmark41); [Bhatt *et al.*](#_bookmark3), [2019](#_bookmark3)), so in many cases this experimental simplification may be worth-

292 while to improve inference on the overall individual and combined effect of temperature and

293 photoperiod on biological processes.

294 2. **Uncouple thermoperiod and photoperiod**. By varying thermoperiod and photoperiod

295 independently, statistical orthogonality can be maintained across treatments. For example,

296 a study could apply photoperiod treatment levels of 8 vs. 16 hour day/night with a 12

297 hour thermoperiod regime across temperature treatments (Fig. [4](#_bookmark52)b.). While this approach

298 allows for more robust evaluation of cue effects and interactions, uncoupling photoperiod

299 and thermoperiod can require newer and more expensive technologies which many not be

300 widely available. Further, this approach may also introduce new artifacts that occur from

301 the biological rather than statistical interactions between light and temperature ([Chew *et al.*](#_bookmark10),

302 [2012](#_bookmark10)). There is evidence that increasing temperatures in the first two hours of daylight can

303 be almost as effective for stimulating shoot elongation as similar temperature increases for

304 the whole photoperiod ([Erwin](#_bookmark17), [1998](#_bookmark17)). With this design, treatments must inherently differ in

305 the amount of time the warmer daytime temperature extends into the dark, nighttime light

306 regime (or vice versa), introducing a new axis of non-orthogonality to consider.

307 3. **Include thermoperiodicity as an explicit experimental treatment**. In many study

308 systems in which both photoperiod and thermoperiod influence biological processes, experi-

309 mentalists often include thermoperiodicity as an explicit experimental treatment with both

310 constant and varying day/night temperatures applied as separate treatment variables (e.g.,

311 [Zaslavski *et al.*](#_bookmark47), [1995](#_bookmark47)). Such experiments are the best way to assess the comparative im-

312 portance of temperature intensity and periodicity, which would provide important insights

313 towards parsing the relative strength of temperature and light cues and their interactions,

314 (Fig. [4](#_bookmark52)c.). However, executing such an experiment with a full-factorial design would sub-

315 stantially increase the size of a study and its experimental effort and, given the availability of

316 the statistical tools and simpler experimental designs we discuss above, may only be worth-

317 while when researchers are explicitly interested in the relative contributions of periodicity and

318 intensity in a particular study system.

319 Any of these designs can be implemented for any number of treatment levels, and incorporate

320 additional manipulated variables. This is particularly important in the context of global change

321 because shifts in the magnitude, and even direction, of climate change will vary spatially, and in some

322 systems, two level studies may struggle to estimate interactions due to under-sampled treatment

323 levels ([Collins *et al.*](#_bookmark15), [2022](#_bookmark15)). Researchers seeking to capture such environmental dynamics could

324 adapt these design to include at least a “control”, “increase” and “decrease” treatment level for

325 experimental variables. Additionally, with increasing numbers of treatments and treatment levels,

[326](#_bookmark39) full-factorial experiments can be implement with a *response surface methodology*, (e.g, [Schubert](#_bookmark39)

327 [*et al.*](#_bookmark39), [2009](#_bookmark39); [Begoude *et al.*](#_bookmark2), [2007](#_bookmark2)); a powerful approach for detecting interactions and non-linearities

328 among variables. Experimentalists should leverage their knowledge of the natural history of their

329 study organisms, historical climate observations and climate change projections to determine the

330 treatment levels that are most appropriate for their study.

331 In correcting one problem, each of the designs that we outlined above introduces another, which

332 may in fact be an intrinsic property of any experimental manipulation. It would be useful for

333 researchers to explicitly test how cue estimates vary among experimental designs, and which design

334 is most useful for predicting biological responses to environmental cues in the field under current

335 and future climate conditions. In the meantime, we hope that this issue is a reminder that, as

336 experimentalists, we must continue to be thoughtful about matching our experimental designs to

337 the goals of a study, and be transparent about uncertainty around our experimental inference.

338 **References**

339 Anduaga, A.M., Nagy, D., Costa, R. & Kyriacou, C.P. (2018) Diapause in drosophila melanogaster

340 – photoperiodicity, cold tolerance and metabolites. *Journal of Insect Physiology* **105**, 46–53.

341 Basler, D. & Körner, C. (2014) Photoperiod and temperature responses of bud swelling and bud

342 burst in four temperate forest tree species. *Tree physiology* **34**, 377–388.

343 Begoude, B.A.D., Lahlali, R., Friel, D., Tondje, P.R. & Jijakli, M.H. (2007) Response surface

344 methodology study of the combined effects of temperature, ph, and aw on the growth rate of

345 trichoderma asperellum. *Journal of Applied Microbiology* **103**, 845–854.

346 Bhatt, A., Bhat, N.R., Suleiman, M.K. & Santo, A. (2019) Effects of storage, mucilage presence,

347 photoperiod, thermoperiod and salinity on germination of farsetia aegyptia turra (brassicaceae)

348 seeds: implications for restoration and seed banks in arabian desert. *Plant Biosystems - An*

349 *International Journal Dealing with all Aspects of Plant Biology* **153**, 280–287.

350 Brelsford, C.C. & Robson, T.M. (2018) Blue light advances bud burst in branches of three deciduous

351 tree species under short-day conditions. *Trees-Structure and Function* **32**, 1157–1164.

352 Brown, E.E., Baumann, H. & Conover, D.O. (2014) Temperature and photoperiod effects on sex

353 determination in a fish. *Journal of Experimental Marine Biology and Ecology* **461**, 39–43.

354 Buonaiuto, D.M. & Wolkovich, E.M. (2021) Differences between flower and leaf phenological re-

355 sponses to environmental variation drive shifts in spring phenological sequences of temperate

356 woody plants. *Journal of Ecology* **109**, 2922–2933.

357 Burghardt, L.T., Runcie, D.E., Wilczek, A.M., Cooper, M.D., Roe, J.L., Welch, S.M. & Schmitt, J.

358 (2016) Fluctuating, warm temperatures decrease the effect of a key floral repressor on flowering

359 time in arabidopsis thaliana. *New Phytologist* **210**, 564–576.

360 Casal, J.J. (2002) Environmental cues affecting development. *Current Opinion in Plant Biology* **5**,

361 37–42.

362 Cheng, C.S. (2016) *Theory of Factorial Design*. Chapman and Hall/CRC.

363 Chew, Y.H., Wilczek, A.M., Williams, M., Welch, S.M., Schmitt, J. & Halliday, K.J. (2012) An

364 augmented arabidopsis phenology model reveals seasonal temperature control of flowering time.

365 *New Phytologist* **194**, 654–665.

366 Chiang, C., Bånkestad, D. & Hoch, G. (2020) Reaching natural growth: The significance of light

367 and temperature fluctuations in plant performance in indoor growth facilities. *Plants* **9**.

368 Chuine, I. & Beaubien, E. (2001) Phenology is a major determinant of tree species range. *Ecology*

369 *Letters* **4**, 500–510.

370 Chuine, I., Morin, X., Bugmann, H., Koerner, C. & Basler, D. (2010) Warming, photoperiods, and

371 tree phenology. *Science* **329**, 277–278.

372 Cober, E., Tanner, J. & Voldeng, H. (1996) Soybean photoperiod-sensitivity loci respond differen-

373 tially to light quality. *Crop Science* **36**, 606–610.

374 Collins, S., Whittaker, H. & Thomas, M.K. (2022) The need for unrealistic experiments in global

375 change biology. *Current Opinion in Microbiology* **68**, 102151.

376 Downs, R.J. (1980) Phytotrons. *Botanical Review* **46**, 447–489.

377 Erwin, J. (1998) Temperature and light effects on stem elongation. *Journal of the Japanses Society*

378 *For Horticultural Science* **67**, 1113–1120, Session of Plant Growth Regulation by Physical and

379 Mechanical Stimuli at the Commemorative Symposium of the 75th Anniversary of Japanese-

380 Society-for-Horticultural-Science, JAPAN, APR 03-04, 1998.

381 Ettinger, A.K., Chamberlain, C.J., Morales-Castilla, I., Buonaiuto, D.M., Flynn, D.F.B., Savas, T.,

382 Samaha, J.A. & Wolkovich, E.M. (2020) Winter temperatures predominate in spring phenological

383 responses to warming. *Nature Climate Change* **10**, 1137–1142.

384 Flynn, D.F.B. & Wolkovich, E.M. (2018) Temperature and photoperiod drive spring phenology

385 across all species in a temperate forest community. *New Phytologist* **219**, 1353–1362.

386 Forrest, J. & Miller-Rushing, A.J. (2010) Toward a synthetic understanding of the role of phenology

387 in ecology and evolution. *Philosophical Transactions of the Royal Society of London B: Biological*

388 *Sciences* **365**, 3101–3112.

389 Fu, Y.S.H., Zhao, H.F., Piao, S.L., Peaucelle, M., Peng, S.S., Zhou, G.Y., Ciais, P., Huang, M.T.,

390 Menzel, A., Uelas, J.P., Song, Y., Vitasse, Y., Zeng, Z.Z. & Janssens, I.A. (2015) Declining global

391 warming effects on the phenology of spring leaf unfolding. *Nature* **526**, 104–107.

392 Gauzere, J., Lucas, C., Ronce, O., Davi, H. & Chuine, I. (2019) Sensitivity analysis of tree phenol-

393 ogy models reveals increasing sensitivity of their predictions to winter chilling temperature and

394 photoperiod with warming climate. *Ecological Modelling* **411**, 108805.

395 Gu, S. (2016) Growing degree hours - a simple, accurate, and precise protocol to approximate

396 growing heat summation for grapevines. *International Journal of Biometeorology* **60**, 1123–1134.

397 Hamilton, J.A., El Kayal, W., Hart, A.T., Runcie, D.E., Arango-Velez, A. & Cooke, J.E. (2016)

398 The joint influence of photoperiod and temperature during growth cessation and development of

399 dormancy in white spruce (Picea glauca). *Tree Physiology* **36**, 1432–1448.

400 Hellmers, H. (1966) Growth response of redwood seedlings to thermoperiodism. *Forest Science* **12**,

401 276–283.

402 Koerner, C. & Basler, D. (2010) Phenology under global warming. *Science* **327**, 1461–1462.

403 Körner, C. & Basler, D. (2010) Response—warming, photoperiods, and tree phenology. *Science*

404 **329**, 278–278.

405 Laube, J., Sparks, T.H., Estrella, N., Höfler, J., Ankerst, D.P. & Menzel, A. (2014) Chilling out-

406 weighs photoperiod in preventing precocious spring development. *Global Change Biology* **20**,

407 170–182.

408 López‐Olmeda, J.F. & Sánchez‐Vázquez, F.J. (2009) Zebrafish temperature selection and synchro-

409 nization of locomotor activity circadian rhythm to ahemeral cycles of light and temperature.

410 *Chronobiology International* **26**, 200–218.

411 MacLean, H.J. & Gilchrist, G.W. (2019) Temperature, photoperiod and life history traits in

412 drosophila subobscura. *bioRxiv* p. 717967.

413 Parent, B., Millet, E.J. & Tardieu, F. (2019) The use of thermal time in plant studies has a sound

414 theoretical basis provided that confounding effects are avoided. *Journal of Experimental Botany*

415 **70**, 2359–2370.

416 Piao, S., Friedlingstein, P., Ciais, P., Viovy, N. & Demarty, J. (2007) Growing season extension and

417 its impact on terrestrial carbon cycle in the northern hemisphere over the past 2 decades. *Global*

418 *Biogeochemical Cycles* **21**.

419 Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W.H., Kleyer, M.,

420 Schurr, U. & Postma, J. (2016) Pampered inside, pestered outside? differences and similarities

421 between plants growing in controlled conditions and in the field. *New Phytologist* **212**, 838–855.

422 Pörtner, H.O. & Farrell, A.P. (2008) Physiology and climate change. *Science* **322**, 690–692.

423 Primack, R.B., Laube, J., Gallinat, A.S. & Menzel, A. (2015) From observations to experiments

424 in phenology research: investigating climate change impacts on trees and shrubs using dormant

425 twigs. *Annals of Botany* **116**, 889–897.

426 Rosenberg, N.J. (1974) *Microclimate: the biological environment*. Wiley, New York.

427 Sanz-Perez, V., Castro-Diez, P. & Valladares, F. (2009) Differential and interactive effects of tem-

428 perature and photoperiod on budburst and carbon reserves in two co-occurring Mediterranean

429 oaks. *Plant Biology* **11**, 142–51.

430 Scheiner, S.M. & Gurevitch, J. (2001) *Design and analysis of ecological experiments*. Oxford Uni-

431 versity Press.

432 Schubert, M., Mourad, S., Fink, S. & Schwarze, F. (2009) Ecophysiological responses of the biocon-

433 trol agent trichoderma atroviride (t-15603.1) to combined environmental parameters. *Biological*

434 *Control* **49**, 84–90.

435 Stewart, R.I.A., Dossena, M., Bohan, D.A., Jeppesen, E., Kordas, R.L., Ledger, M.E., Meerhoff,

436 M., Moss, B., Mulder, C., Shurin, J.B., Suttle, B., Thompson, R., Trimmer, M. & Woodward, G.

437 (2013) *Chapter Two - Mesocosm Experiments as a Tool for Ecological Climate-Change Research*,

438 vol. 48, pp. 71–181. Academic Press.

439 Warrington, I., Peet, M., Patterson, D., Bunce, J., Haslemore, R. & Hellmers, H. (1977) Growth

440 and physiological responses of soybean under various thermoperiods. *Functional Plant Biology* **4**,

441 371–380.

442 Way, D.A. & Montgomery, R.A. (2015) Photoperiod constraints on tree phenology, performance

443 and migration in a warming world. *Plant, Cell & Environment* **38**, 1725–1736.

444 Wolkovich, E.M., Chamberlain, C.J., Buonaiuto, D.M., Ettinger, A.K. & Morales-Castilla, I. (2022)

445 Integrating experiments to predict interactive cue effects on spring phenology with warming. *New*

446 *Phytologist* **235**, 1719–1728.

447 Wolkovich, E.M., Ettinger, A.K., Flynn, D., Savas, T., Chamberlain, C., Buonaiuto, D. &

448 Samaha, J. (2019) Observed spring phenology responses in experimental environments (OS-

449 PREE). doi:10.5063/F1CZ35KB.

450 Wright, M.L. & Bruni, N.K. (2004) Influence of the photocycle and thermocycle on rhythms of

451 plasma thyroxine and plasma and ocular melatonin in late metamorphic stages of the bullfrog

452 tadpole, rana catesbeiana. *Comparative Biochemistry and Physiology Part A: Molecular Inte-*

453 *grative Physiology* **139**, 33–40.

454 Xu, D., Song, X., Li, F., Pan, X., Zhu, W. & Zhang, X. (2019) Effects of desiccation, diurnal

455 temperature changes and irradiance on archeospore production of pyropia yezoensis. *Aquaculture*

456 **509**, 167–170.

457 Zaslavski, V.A., Zinovjeva, K.B., Reznik, S.Y. & Umarova, T.Y. (1995) Effect of photoperiod and

458 thermoperiod on the eclosion rhythm of trichogramma evanescens. *Entomologia Experimentalis*

459 *et Applicata* **74**, 99–104.

460 Zohner, C.M., Benito, B.M., Svenning, J.C. & Renner, S.S. (2016) Day length unlikely to constrain

461 climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change* **6**, 1120–

462 1123.

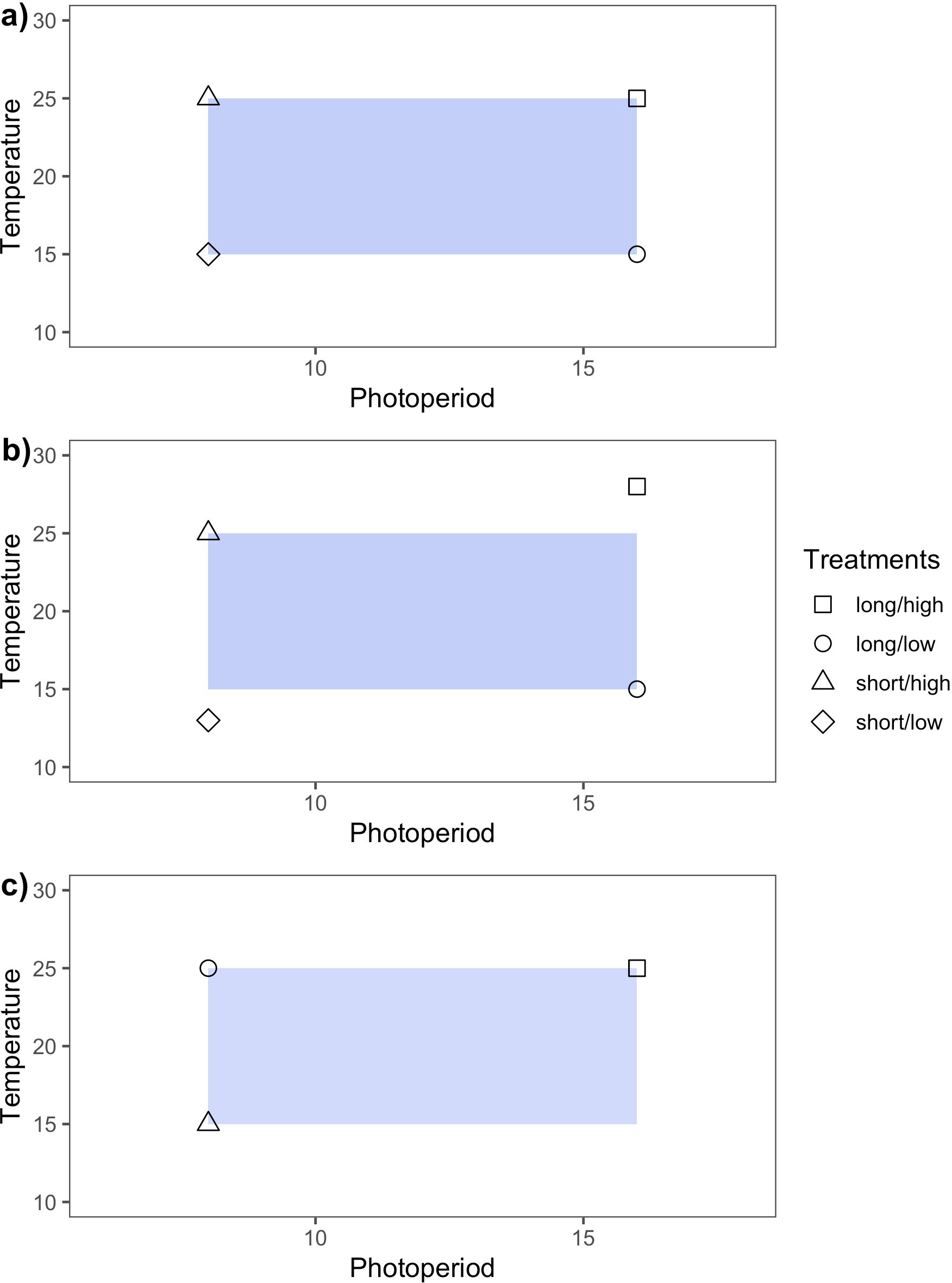


Figure 1: Idealized experimental designs demonstrate three approaches for varying temperature and light treatment levels in controlled environment experiments. Design **a)** is full factorial in that treatments levels are balanced and orthogonal. This design is appropriate for testing interactions between two (or more) variables. In **b)** the design is balanced but not orthogonal. Non-orthogonality in experiments can arise when experimental covariation among the manipulated variables is not accounted for. In **c)**, the experimental design is orthogonal but unbalanced. Lack of balance in experiments often arises due to time, space or resource limitations.

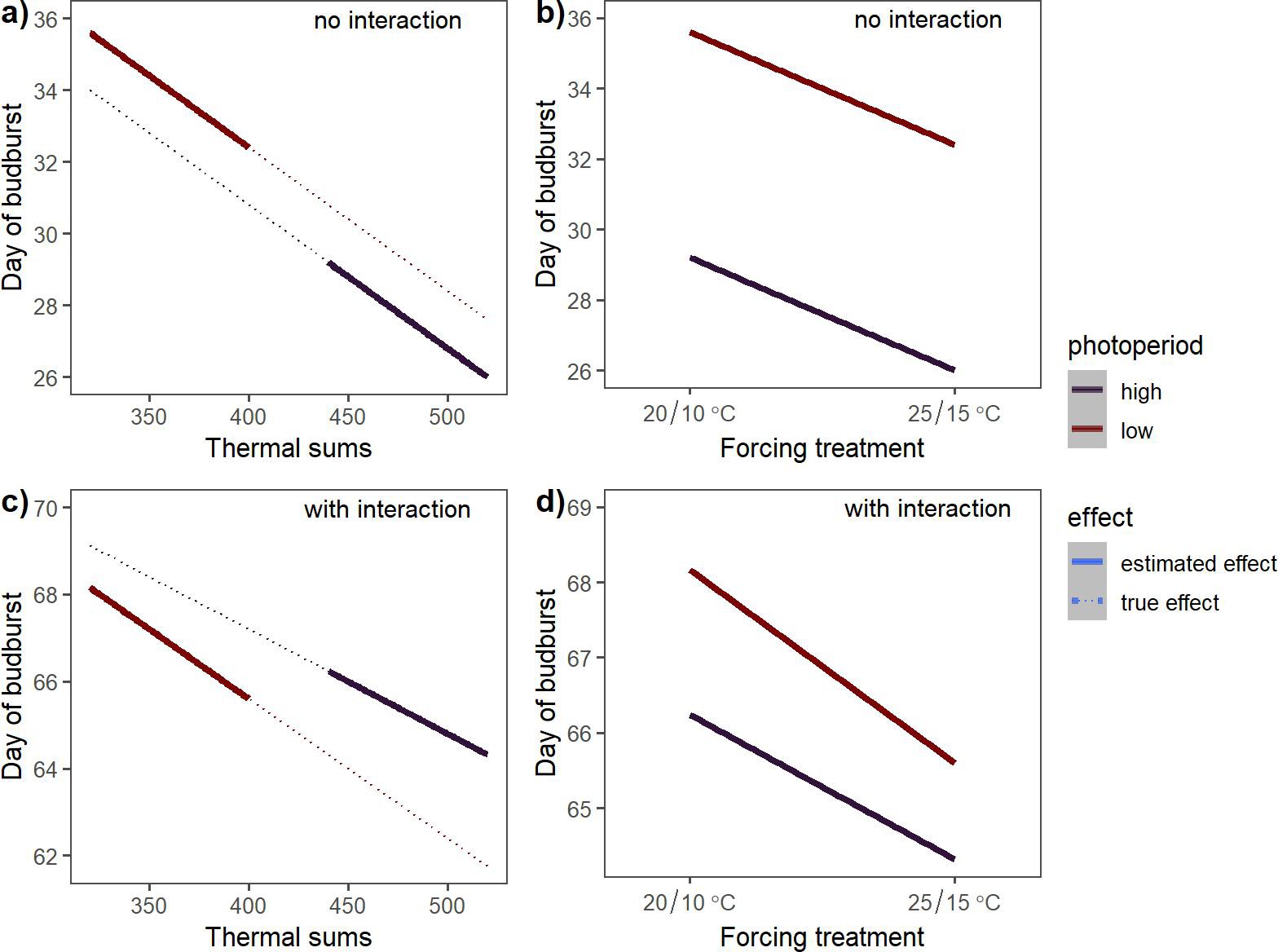


Figure 2: Estimated effects of photoperiod and forcing on spring phenology based on a simulated experiment in which the coupling of photoperiod and thermoperiod introduce an experimental covariation between the temperature and light treatments. The dotted lines in **a)** and **c)** depict the true effects of forcing at each photoperiod level, and the solid lines depict the estimated effects.

**a)** depicts a scenario where forcing and photoperiod effects do not interact, while **c)** includes an interactive effect. **b)** and **d)** depict the estimated effects of forcing and photoperiod if the experimental covariation due to periodicity coupling in **a)** and **c)**, respectively, is unacknowledged. For an example of this principle in experiments with more than two treatment levels see Fig. S1.

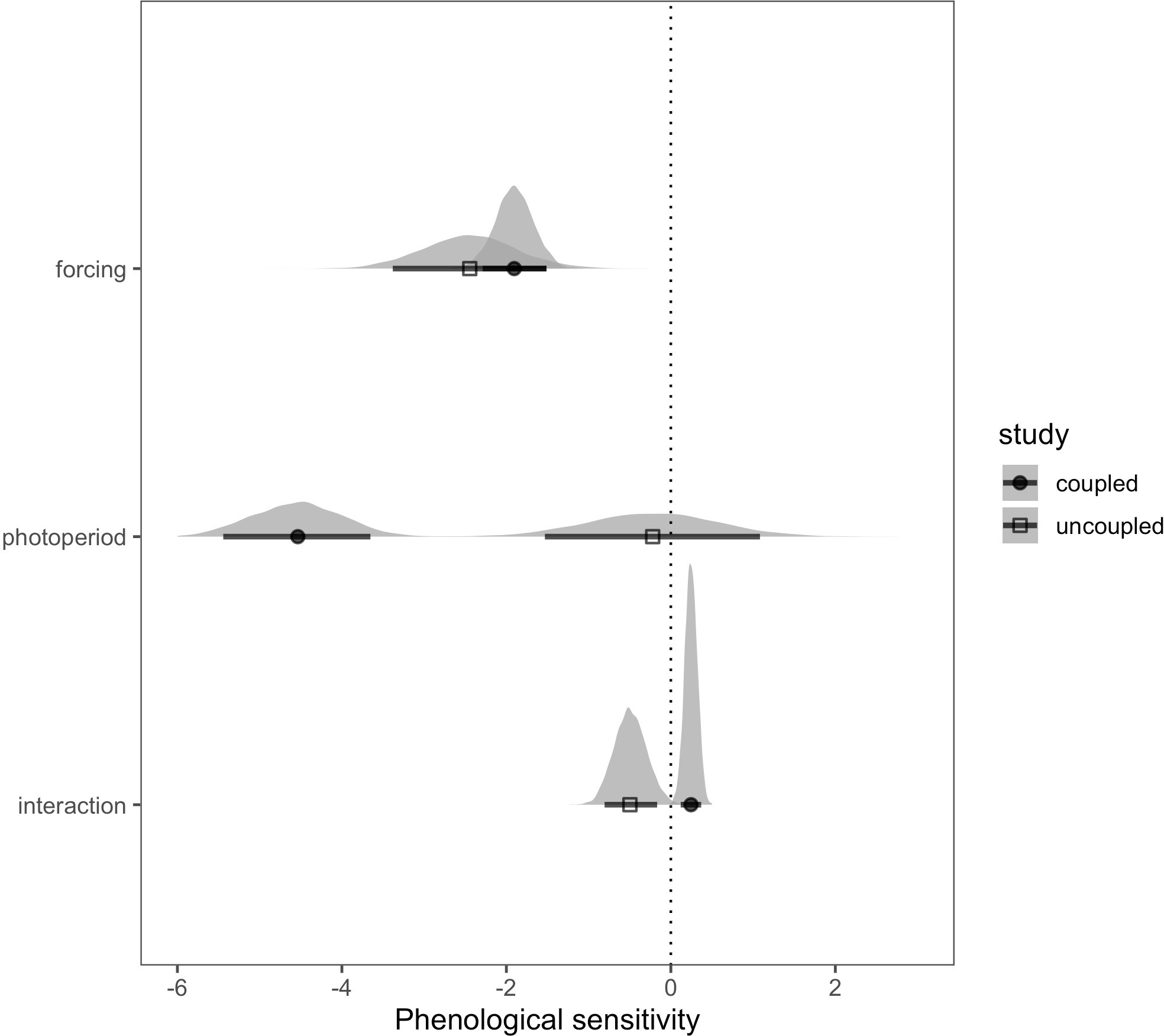


Figure 3: Estimated phenological sensitivity (∆ day of leaf expansion/∆ unit increase in cue level), using alternative methods of varying thermoperiod relative to photoperiod. Points indicate the estimated mean effect and bars the 90% uncertainty intervals. The full posterior distributions for each parameter are also depicted as an additional display of uncertainty. The coupled thermo- photo- period design is from [Flynn & Wolkovich](#_bookmark19) ([2018](#_bookmark19)) and the uncoupled design is from [Buonaiuto](#_bookmark6) [& Wolkovich](#_bookmark6) ([2021](#_bookmark6)).

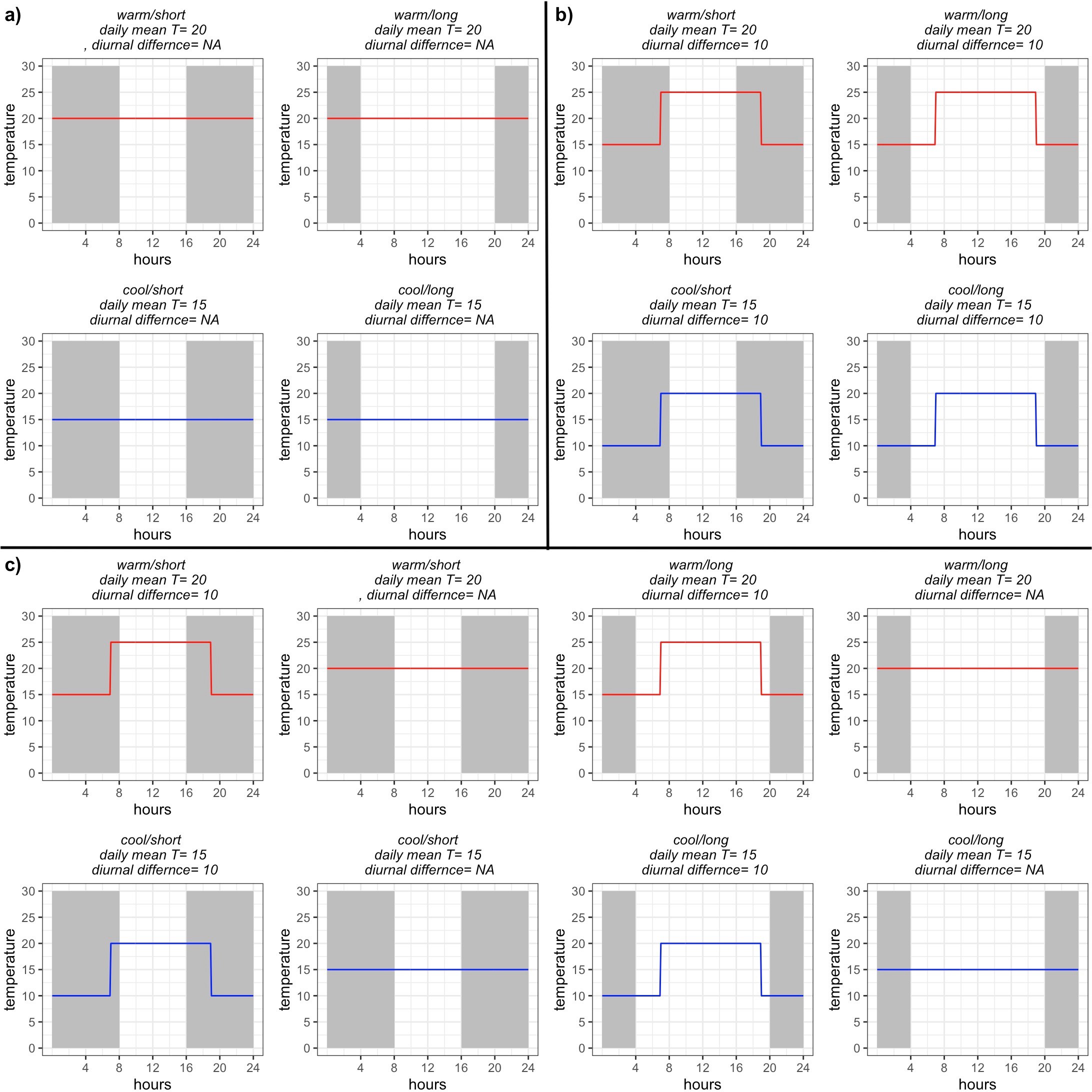


Figure 4: Conceptualized experimental designs to test temperature and daylength interactions on a biological response. Design **a)** manipulates temperature intensity only (no thermoperiodicity). In design **b)**, consistent diurnal temperature fluctuations are maintained but thermoperiod and pho- toperiod are decoupled and varied independently, maintaining orthogonality in daily temperature treatments. In **c)**, thermoperiod is included as an additional, explicit experimental treatments to evaluate the individual and interactive effect of photoperiod, thermoperiod and temperature.