# Experimental designs for testing the interactive effects of temperature and light in ecology and the problem of periodicity

D.M. Buonaiuto [1,2,a], M. Donahue[3], E.M. Wolkovich[4]

*Author affiliations:*

[1]Arnold Arboretum of Harvard University, Boston, Massachusetts, USA. ORCID: 0000-0003-4022-2591

[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

[3]Hawai'i Institute of Marine Biology, University of Hawai'i at Manoa, Kan'eohe, HI, USA.

[4]Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, Canada

[a]Corresponding author: 617.823.0687; dbuonaiuto@g.harvard.edu

# Abstract

The abstract must not exceed 350 words and should list the main results and conclusions, using FOUR simple, factual, numbered statements: 1: set the context for and purpose of the work;
2: indicate the approach and methods;
3: outline the main results;
4: identify the conclusions and the wider implications.

**Keywords:** phenology, photoperiod, thermoperiod, growth chambers, cues, light, temperature

# Introduction

Across the tree of life, temperature and light availability shape a number of important biological processes including growth and metabolic rates (MacLean & Gilchrist, 2019) sex determination (Brown *et al.*, 2014), acclimatization to seasonal environments (Hamilton *et al.*, 2016) and the timing of life cycle transitions (phenology) (Forrest & Miller-Rushing, 2010). These biological responses in turn dictate broad scale ecological processes and patterns ranging from biogeochemical cycling (Piao *et al.*, 2007) to species range limits (Chuine & Beaubien, 2001). Characterizing the specific dynamics of how these environmental factors synergistically affect biological processes across a wide range of taxa has become even more important as anthropogenic global change continues to expose organisms to novel environmental conditions (Pörtner & Farrell, 2008).

Because temperature and light availability often co-vary in the field (for example, in most temperate ecosystems, daylength and temperature both increase as the season progresses (Rosenberg, 1974)) it can be difficult to disentangle their relative contributions to biological processes. In contrast, experimental manipulations of climate variables in artificial environments can mechanistically characterize biological responses to environmental fluctuations (Ettinger *et al.*, 2020; Primack *et al.*, 2015). Growth chambers of all shapes and sizes have been used to this end (Downs, 1980) and these efforts have greatly advanced researchers' understanding of the fundamental biology of a wide variety of organisms and their ability to predict the responses to current and future climate change (Stewart *et al.*, 2013).

However, controlled environment experiments have their own challenges. Experimentalists must balance biological realism with statistical inference, experimental effort with statistical power, and account for the effects of unmanipulated or unmeasured variables (Scheiner & Gurevitch, 2001). Because biological responses to the environment are generally the product of complex interactions between multiple environmental signals (Casal, 2002), seemingly small choices about experimental design can generate significant differences in outcomes. Experimental treatments are rarely standardized among researchers even within disciplines (Wolkovich *et al.*, in Review.) and these

2

complexities may in part contribute to many discrepancies between experimental studies and observation data (Poorter *et al.*, 2016). Even with these limitations, growth chamber studies remain a powerful tool for mechanistically assessing organismic responses to the environment provided that the implications of treatment designs are well understood and well matched with the scope of the research question.

As technology advances and experiments become more complex, researchers can manipulate more variables and multiple axes of variation (e.g. temperature, amplitude, periodicity, wavelength) at the same time. Yet these efforts may tradeoff between biological realism for statistical inference. Through surveying the literature of experiments with plant phenology (timing of recurring life cycle events, e.g. leaf budburst, flowering), we found that experiments that manipulate both photo- and thermo- periodicity often to introduce a latent covariation between light and temperature experimental treatments that may misrepresent the effects of each of these environmental variables and the interaction between them. Below, we begin by briefly detailing how temperature and light treatments are generally applied in experimental phenology studies and review the minimum experimental elements required to robustly test interactions between two or more environmental variables. We then detail the problem of inference that can arise manipulating the periodicity of both temperature and light in experiments and demonstrate the extent to which this is an issue with mathematical and experimental examples. Finally, we conclude by outlining several possible solutions for overcoming these issues. While, our example deals with phenology of temperate woody plants, the issues and solutions we present below are broadly applicable to studies of any other organisms and biological processes that utilize temperature and daylength signals.

# Estimating phenological cues from experiments

Decades of experimental work in growth chambers have demonstrated that temperature (both cool temperatures in fall/winter and warming temperatures in spring) and photoperiod are the primary cues of phenology for plants in the temperate/boreal zones (Ettinger *et al.*, 2020). While exposure to cool winter temperatures (chilling) strongly impacts phenology (Laube *et al.*, 2014), we focus here on warm temperature treatments and light treatments, because controlled chilling treatments with light are uncommon ((Wolkovich *et al.*, in Review.)). Choices that about how to apply these two treatments in particular that can compromise inference on their effects, so we will focus of these two cues.

While a large variety of experimental designs have been used to study plant phenology, generally phenology experiments tend to manipulate two major axes of light and warm temperature variation:

1. <u>Intensity:</u> The amount or quality of a variable. Here we define temperature intensity as the amount of heat present in the system (measured in degrees). In the phenology literature this measurement is generally referred to as forcing. We define light intensity as the luminosity or

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

88  2. Periodicity: The interval at which the intensity of the variable is applied. Hereafter, we refer
89  to the periodicity of light as photoperiod (often used synonomously with "daylength") and
90  the periodicity of temperature as thermoperiod.

91  For phenology, photoperiodicity is generally considered the primary light cue for plants ((Way
92  & Montgomery, 2015) though see (Brelsford & Robson, 2018; Cober *et al.*, 1996) regarding light
93  intensity and phenology). For temperature, conventionally both intensity and periodicity drive
94  phenological activity and several metrics (e.g. growing degree hours, thermal sums, growing degree
95  days) that combine these two axes have been developed (Gu, 2016). This assumption is well
96  supported; under natural conditions diurnal temperature fluctuations in temperate regions can
97  be quite large in the spring, and studies have found that diurnal temperature variation strongly
98  influences plant phenology (Burghardt *et al.*, 2016). In fact, even if thermoperiodicity is not an
99  explicit treatment variable (i.e. manipulated systematically), incorporating it in experiments is
100 essential for translating experimental results into real world predictions (Chiang *et al.*, 2020).

101 Like many other biological processes, recent advances have demonstrated that plant phenological
102 responses are nonlinear, due largely to interactions between cues (Wolkovich *et al.*, in Review.;
103 Fu *et al.*, 2015), highlighting the need for experiments designed to evaluate the strength of these
104 interactions. To have the statistical power to partition the individual and interactive effects of two
105 or more variables, an experiment must:

106 1. Have at minimum of two treatment levels of at least two variables.

107 2. Treatment levels must be full factorial (Fig. 1a.). Full factorial designs are both balanced
108 (Fig. 1b.) and orthogonal (Fig. 1c.); which is to say that all possible treatment combinations
109 are applied and each treatment is independent of all others (Cheng, 2016).

110 These two critical elements may seem obvious but can be conspicuously absent from many pub-
111 lished studies. In the case of woody plant phenology, using a recently published database (OSPREE:
112 Observed spring phenological responses in experimental environments (Wolkovich *et al.*, 2019)) we
113 found that out of 152 controlled environment experiments (across 93 studies) only 64 of them ma-
114 nipulated both light and forcing cues in the same experiment and only 15 of those did so with a
115 design that was both balanced and orthogonal (see Supplement for details). This notable dearth of
116 robust tests of light and temperature interactions may stem from the common limitations of time,
117 space, and resources that experimentalists often face, but it may equally relate to a fundamental
118 issues that arises from the fact that these variables themselves are comprised of multiple axes of
119 variably.

120

## The problem of periodicticy

A common approach in phenology experiments that seems to balance prior knowledge about the underlying physiology of phenology, biological realism and experimental inferences is to vary photoperiodicity, and thermal intensity and periodicity

Yet the orthogonality of this design is based on the assumption to a 12 hour thermoperiod. If, rather the thermoperiod is coupled with the photoperiod, this is not the case because the daily mean temperature of the long/high treatment will be higher than that of the short/high treatment, and the long/low treatment slightly warmer than the short low (Fig 2.a). This is because the warmer day time temperatures are applied for different duration across the high temperature treatments. While this covariation among the photoperiod and temperature treatments is biologically realistic, it makes it statistically impossible to differentiate their independent and interactive effects on any given biological process.

This problem of inference that arises from the experimental covariation of thermo- and photoperiodicty is not limited only to studies seeking to directly compare the effects of photoperiod and forcing; it applies in any study evaluating the influence of photoperiod on biological activity, even if it is the only manipulated cue. Experimentally isolating the effect of photoperiod assumes that all other environmental variables are held constant. As in the case decribed aboved for comparing the interactive effects of photoperiod and forcing, the covariation of photoperiod and thermoperiod in an experiment where forcing was supposed to be a consistant ambient condition, would in fact yeild a situation in which the long photoperiod treatment was also receiving more, unmeasured heating that the shorter photoperiod treatments. In this case, some amount of the perceieved photoperiod effect is due to the latent, increased forcing, and the true effect of photoperiod cannot be ascertained.

Of the studies in the OSPREE database that manipulated photoperid experimental, up to 25% of them appear to include a covariation with theromoperiod. Of the studies that manipulated both photoperiod and temperature interactively, we found that around 45% of them may have this issue, suggesting that the true interactive effects of these cues on spring phenology is still quite poorly characterized. This may be in part why the relative contribution of temperature and photoperiod cues to spring phenology remains a contentious debate in the phenology literature (Koerner & Basler, 2010).

## Periodicity and inference

Studies repeatedly suggest that forcing is a more dominant cue than photoperiod for spring phenology (CHUINE *et al.*, 2010; Zohner *et al.*, 2016; Gauzere *et al.*, 2019), therefore we should *a priori* expect the covariation between photo- and thermo-period may result in an over-estimation of the photoperiod effect. We can mathematically solve for how much of an estimated photoperiod effect is due to forcing—in experiments where they covary—by making several major asssumptions. If

we assume forcing and photoperiod effects are additive, linear and there is no interaction between the two effects, then we can solve (algebraically) for the estimated effect of forcing and photoperiod given an orthogonal treatment design (see Fig. 3, Supporting info). Using estimates from one experiment that covaried forcing and photoperiod effects Flynn & Wolkovich (2018), we found roughly two-thirds of the estimated photoperiod cue could be due to forcing effects.

While we are aware of no experiments that explicitly compare the effects of co-varying vs. independent photo- and thermo- periods, two phenology experiments in our lab utilized many overlapping treatment levels and species from the same sampling sites however in one study, Flynn & Wolkovich (2018), photo- and thermo- period co-vary, while in the other Buonaiuto & Wolkovich (2021) photo- and thermo- period were varied independently. Comparing the cue estimates from these two studies offers an opportunity to test the theoretical and mathematical predictions, and further understanding the uncertainly in cue estimate due to periodicity.

We subset each dataset (publicly available at the Harvard Forest Data Archive and KNB) to include only the species shared among the two studies, and re-analyzed the data using Bayesian hierarchical models to compare difference in the photoperiod and forcing estimates (see Supplement for Methods). We found that the estimated differences in the mean response to photoperiod and forcingamong study designs were on the same order our mathematical predictions, and that the uncoupled design estimated a weaker (less negative) photoperiod effect, and stronger forcing effects than the coupled experimental design (4,**??**).

There are almost certainly other factors driving the differences between these experiments. Both were conducted in different years, sampled different individual from the population, and used different methods for applying an chilling pre-treatments (Flynn & Wolkovich, 2018; Buonaiuto & Wolkovich, 2021). However, because this comparison is well matched to our mathematical predictions and prior knowledge about how temperature and photoperiod are expected to interacting in phenology, we argue that the influence of periodicity covariation on statistical inference is apparent enough to take seriously.

# Paths Forward

Above we have systematically demonstrated that experiments that co-vary thermoperiod and photoperiod cannot robustly differentiate the individual effect of temperature and photoperiod on a spring phenology (or any other biological process) or accurately quantify their interactive influence. Given the paucity of interactive studies in the literature, it is clear that more well designed studies will be needed to better characterize the effects of these cues. Below we offer several generalized experiment designs that improve statistical orthogonality of controlled environment experiments which could be further developed and adjusted to fit the needs of experimentalists across many sub-fields of biology.

1. **Covarying photo- thermo- period with quantified uncertainty**. It may be that the experimental design that best balances environmental realism, statistical inferences and translatability to observational studies are designs that co-vary periodicity to mimic natural systems (2a.). Moving forward, researchers using this design need to be aware of the non-orthogonality of this design, and be sure to present the uncertainty surrounding their cue effect estimate, which could be done using a similar mathematical approach to the one we present in this paper (see Supplement).

2. **Manipulate photoperiod and temperature intesity with no thermoperiodicity**. This approach allows for the maintenance of statistical orthogonality across treatment combinations (2b.). The main drawback is that this design sacrifices the biological realism of diurnal temperature variation, which may make it more difficult to translate estimates from experiments to real world applications.

3. **Compensitory diurnal temperature fluctions**. There are almost unlimited pairs of integers that can reduce to the same mean (e.g. $(24 + 26)/2 = (30 + 20)/2 = 25$) and the non-orthogonality of the mean daily temperature that arises in a coupled photo-thermoperiod design could be corrected for by proportionately increasing the diurnal temperature fluctuation of the short photoperiod treatment relative to the long treatments (2c.). However, if the differences between day and night temperature has a meaningful biological effect, this introduces another confounding, non-orthogonal factor for interpreting temperature and photoperiod effects. For example, the influence of day time warming of phenology can be as much as three times stronger than proportionate night time warming (Rossi & Isabel, 2017; Meng *et al.*, 2020).

4. **Uncouple thermoperiod and photoperiod.** By varying thermoperiod and photoperiod independently (2d.), statistical orthogonality can be maintained across treatment. However, this approach may also introduce new artifacts that occur from the biological rather than statistical interactions between light and temperature. For example, there is evidence that increasing temperatures in the first two hours of daylight can be almost as effective for stimulating shoot elongation as similar temperature increases for the whole photoperiod (Erwin, 1998). With this design, treatments must inherently differ in the amount of time the warmer daytime temperature extends into the dark nighttime light regime, introducing a new axes of non-orthogonality.

In correcting one problem, each of these designs introduces another, which may in fact be an intrinsic property of any experimental manipulation. It would certainly be useful for researchers to explicitly test how cue estimates vary among these experimental designs, and which design is most useful for predicting phenology in field under current and future climate conditions. In the meantime, we hope the treatment reminds experimentalists that we must continue to be thoughtful about matching design to the goals of a study, and being transparent about uncertainty around our experimental inferences.

## Author contributions

DMB, MD and EMW conceived of the manscript; MD and EMW developed the algebraic solution; DMB performed the comparative analysis of the published studies; DMB led the writing of the manuscript. All authors contributed to writing and gave final approval for the submission.

## Data Availability

Data from the Flynn & Wolkovich (2018) study is available at the Harvard Forest Data Archevie (https://harvardforest1.fas.harvard.edu/exist/apps/datasets/showData.html?id=HF314) and from the ? study available at Knowledge Network for Biocomplexity (https://knb.ecoinformatics.org/view/doi:10.5063/F The R code used to analyse the data is available on github.

## References

Brelsford, C.C. & Robson, T.M. (2018) Blue light advances bud burst in branches of three deciduous tree species under short-day conditions. *TREES-STRUCTURE AND FUNCTION* **32**, 1157–1164.

Brown, E.E., Baumann, H. & Conover, D.O. (2014) Temperature and photoperiod effects on sex determination in a fish. *Journal of Experimental Marine Biology and Ecology* **461**, 39–43.

Buonaiuto, D.M. & Wolkovich, E.M. (2021) Differences between flower and leaf phenological responses to environmental variation drive shifts in spring phenological sequences of temperate woody plants. *Journal of Ecology* **109**, 2922–2933.

Burghardt, L.T., Runcie, D.E., Wilczek, A.M., Cooper, M.D., Roe, J.L., Welch, S.M. & Schmitt, J. (2016) Fluctuating, warm temperatures decrease the effect of a key floral repressor on flowering time in arabidopsis thaliana. *New Phytologist* **210**, 564–576.

Casal, J.J. (2002) Environmental cues affecting development. *Current Opinion in Plant Biology* **5**, 37–42.

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

Chiang, C., Bånkestad, D. & Hoch, G. (2020) Reaching natural growth: The significance of light and temperature fluctuations in plant performance in indoor growth facilities. *Plants* **9**.

Chuine, I. & Beaubien, E. (2001) Phenology is a major determinant of tree species range. *Ecology Letters* **4**, 500–510.

CHUINE, I., MORIN, X., BUGMANN, H., KÖRNER, C. & BASLER, D. (2010) ¡strong¿warming, photoperiods, and tree phenology¡/strong¿ [with response]. *Science* **329**, 277–278.

Cober, E., Tanner, J. & Voldeng, H. (1996) Soybean photoperiod-sensitivity loci respond differentially to light quality. *CROP SCIENCE* **36**, 606–610.

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

Erwin, J. (1998) Temperature and light effects on stem elongation. *JOURNAL OF THE JAPANESE SOCIETY FOR HORTICULTURAL SCIENCE* **67**, 1113–1120, Session of Plant Growth Regulation by Physical and Mechanical Stimuli at the Commemorative Symposium of the 75th Anniversary of Japanese-Society-for-Horticultural-Science, JAPAN, APR 03-04, 1998.

Ettinger, A.K., Chamberlain, C.J., Morales-Castilla, I., Buonaiuto, D.M., Flynn, D.F.B., Savas, T., Samaha, J.A. & Wolkovich, E.M. (2020) Winter temperatures predominate in spring phenological responses to warming. *Nature Climate Change* **10**, 1137–1142.

Flynn, D.F.B. & Wolkovich, E.M. (2018) Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytologist* **219**, 1353–1362.

Forrest, J. & Miller-Rushing, A.J. (2010) Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **365**, 3101–3112.

Fu, Y.S.H., Zhao, H.F., Piao, S.L., Peaucelle, M., Peng, S.S., Zhou, G.Y., Ciais, P., Huang, M.T., Menzel, A., Uelas, J.P., Song, Y., Vitasse, Y., Zeng, Z.Z. & Janssens, I.A. (2015) Declining global warming effects on the phenology of spring leaf unfolding. *Nature* **526**, 104–107.

Gauzere, J., Lucas, C., Ronce, O., Davi, H. & Chuine, I. (2019) Sensitivity analysis of tree phenology models reveals increasing sensitivity of their predictions to winter chilling temperature and photoperiod with warming climate. *Ecological Modelling* **411**, 108805.

Gu, S. (2016) Growing degree hours - a simple, accurate, and precise protocol to approximate growing heat summation for grapevines. *International Journal of Biometeorology* **60**, 1123–1134.

Hamilton, J.A., El Kayal, W., Hart, A.T., Runcie, D.E., Arango-Velez, A. & Cooke, J.E. (2016) The joint influence of photoperiod and temperature during growth cessation and development of dormancy in white spruce (Picea glauca). *Tree Physiology* **36**, 1432–1448.

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

Laube, J., Sparks, T.H., Estrella, N., Höfler, J., Ankerst, D.P. & Menzel, A. (2014) Chilling outweighs photoperiod in preventing precocious spring development. *Global Change Biology* **20**, 170–182.

MacLean, H.J. & Gilchrist, G.W. (2019) Temperature, photoperiod and life history traits in <em>drosophila subobscura</em>. *bioRxiv* p. 717967.

Meng, L., Zhou, Y., Li, X., Asrar, G.R., Mao, J., Wanamaker, A.D. & Wang, Y. (2020) Divergent responses of spring phenology to daytime and nighttime warming. *Agricultural and Forest Meteorology* **281**, 107832.

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N. & Demarty, J. (2007) Growing season extension and its impact on terrestrial carbon cycle in the northern hemisphere over the past 2 decades. *Global Biogeochemical Cycles* **21**.

Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W.H., Kleyer, M., Schurr, U. & Postma, J. (2016) Pampered inside, pestered outside? differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist* **212**, 838–855.

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

Primack, R.B., Laube, J., Gallinat, A.S. & Menzel, A. (2015) From observations to experiments in phenology research: investigating climate change impacts on trees and shrubs using dormant twigs. *ANNALS OF BOTANY* **116**, 889–897.

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

Rossi, S. & Isabel, N. (2017) Bud break responds more strongly to daytime than night-time temperature under asymmetric experimental warming. *Global Change Biology* **23**, 446–454.

Scheiner, S.M. & Gurevitch, J. (2001) *Design and analysis of ecological experiments.* Oxford University Press.

Stewart, R.I.A., Dossena, M., Bohan, D.A., Jeppesen, E., Kordas, R.L., Ledger, M.E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J.B., Suttle, B., Thompson, R., Trimmer, M. & Woodward, G. (2013) *Chapter Two - Mesocosm Experiments as a Tool for Ecological Climate-Change Research*, vol. 48, pp. 71–181. Academic Press.

Way, D.A. & Montgomery, R.A. (2015) Photoperiod constraints on tree phenology, performance and migration in a warming world. *Plant, Cell & Environment* **38**, 1725–1736.

Wolkovich, E., Chamberlin, C., Buonaiuto, D., Ettinger, A. & Morales-Castilla, I. (in Review.) Integrating experiments to predict interactive cue effects on spring phenology with warming .

Wolkovich, E.M., Ettinger, A.K., Flynn, D., Savas, T., Chamberlain, C., Buonaiuto, D. & Samaha, J. (2019) Observed spring phenology responses in experimental environments (OSPREE). doi:10.5063/F1CZ35KB.

Zohner, C.M., Benito, B.M., Svenning, J.C. & Renner, S.S. (2016) Day length unlikely to constrain climate-driven shifts in leaf-out times of northern woody plants. *Nature Climate Change* **6**, 1120–1123.
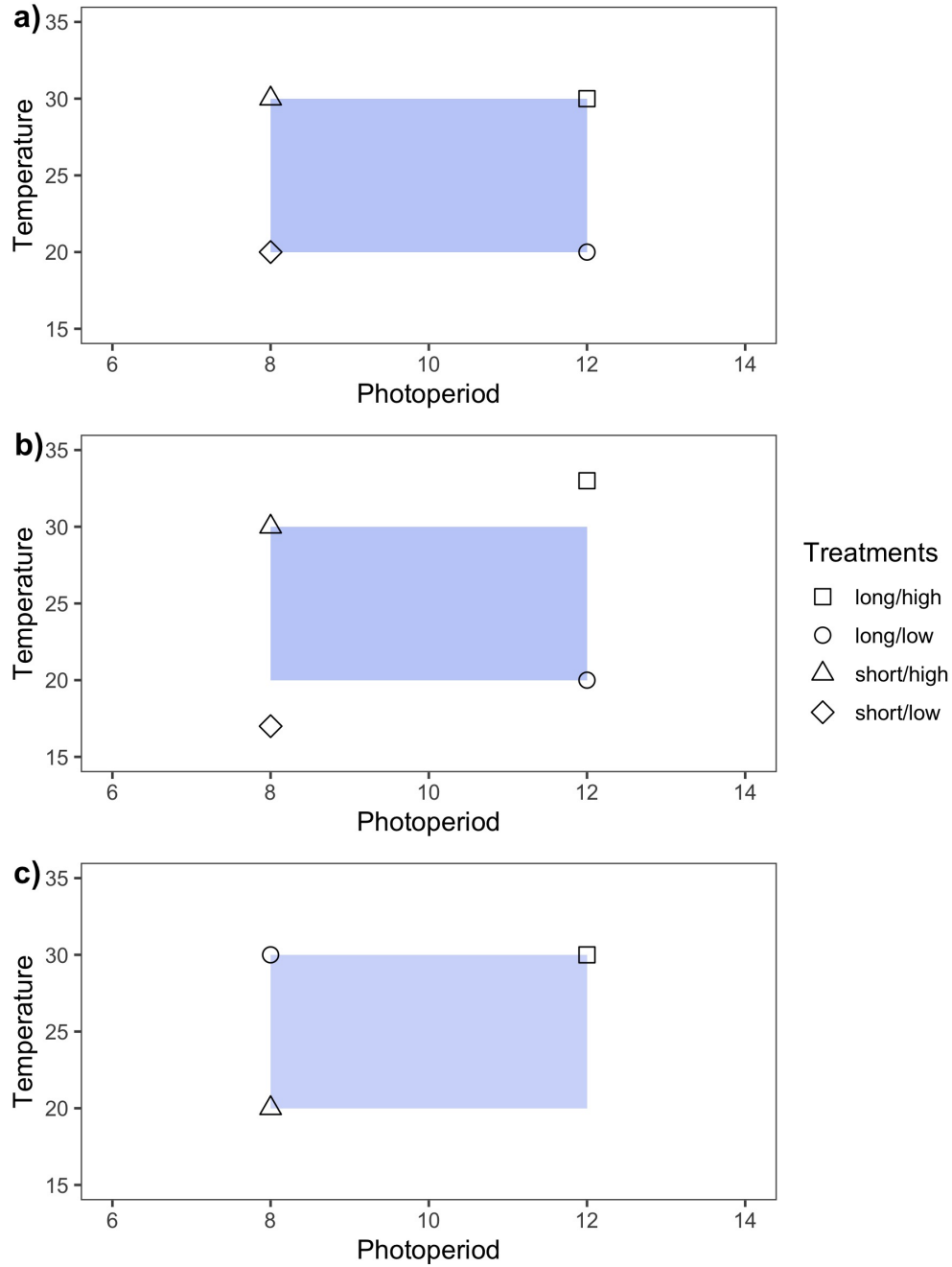
Figure 1: Idealized experimental designs demonstrate three approaches for varying temperature and light treatment level in controlled environment experiments. Design **a)** is fully factorial in that treatments levels are balanced and orthogonal. This design is appropriate for testing interactions between two or more variables. In **b)** the design is balanced both not orthogonal. Non-orthogonality in experiments often arises in experiments when there is covariation among the test variables is unaccounted for. In **c)**, the experimental design is orthogonal but unbalanced. Lack of balance in experiments often arises due to time, space or resource limitations.
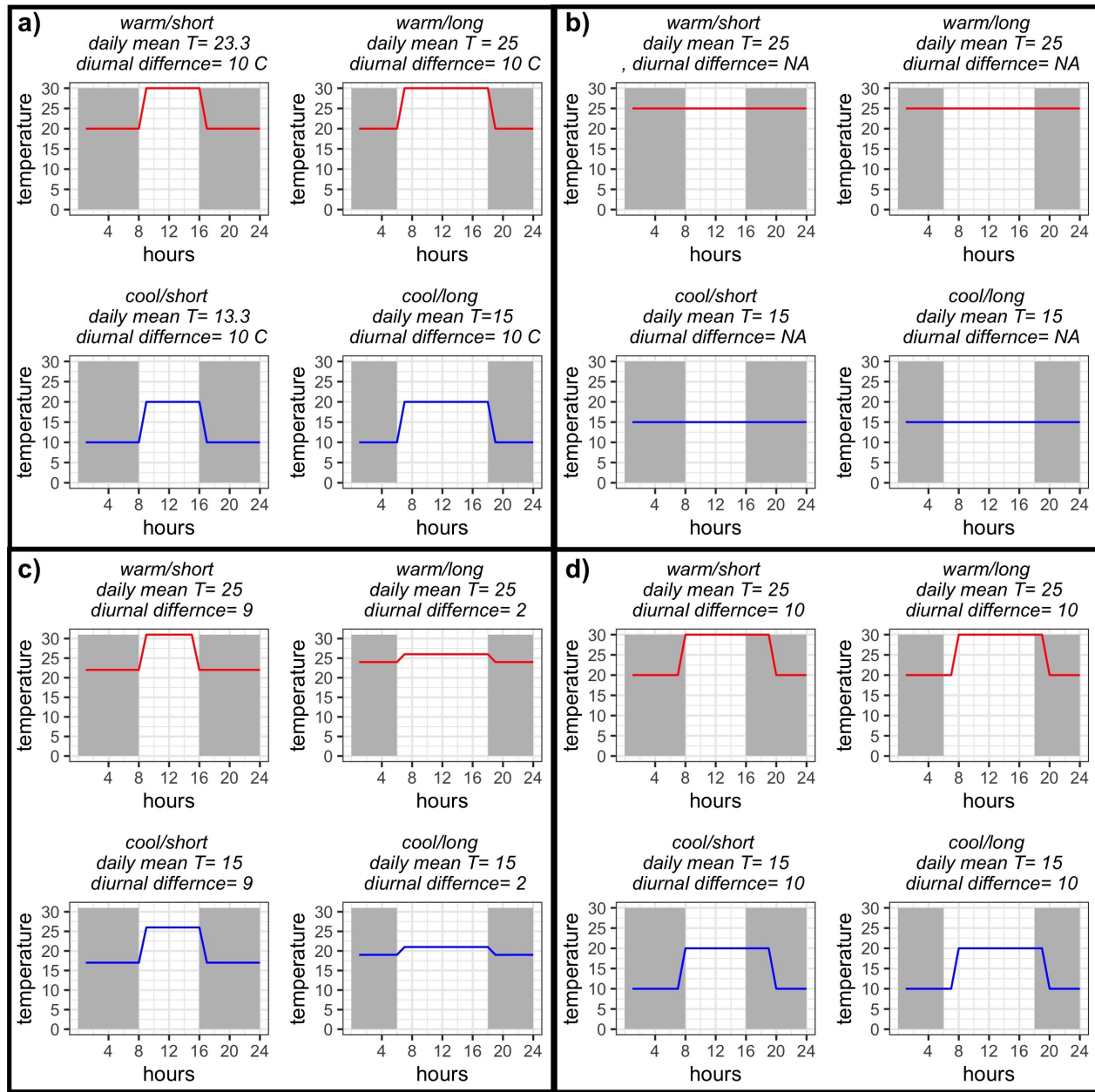
Figure 2: Conceptualized experimental designs to test temperature and daylength interactions on a biological response. In **a)** the design incorporates a standardize diurnal temperature fluctuation across all treatment. Because this thermoperiod is coupled with the photoperiod, while the same day and night temperatures are applied for the high and low temperature treatments respectively, the mean daily temperatures differ across each photoperiod treatment generating non-orthogonality. Designs **b)**,**c)** and **d)** are all designs that can correct this non-orthogonality. Design **b)** manipulated temperature intensity only (no thermoperiodicity).In **c)** photo- and thermo- periods are still are coupled but the orthogonality of mean daily temperature is maintained by proportionately varying the diurnal temperature fluctuations across treatments. In design **d)** standard diurnal temperature fluctuations are maintained but, thermoperiod and photoperiod are decoupled and varied independently, maintaining orthogonality daily mean temperatures.
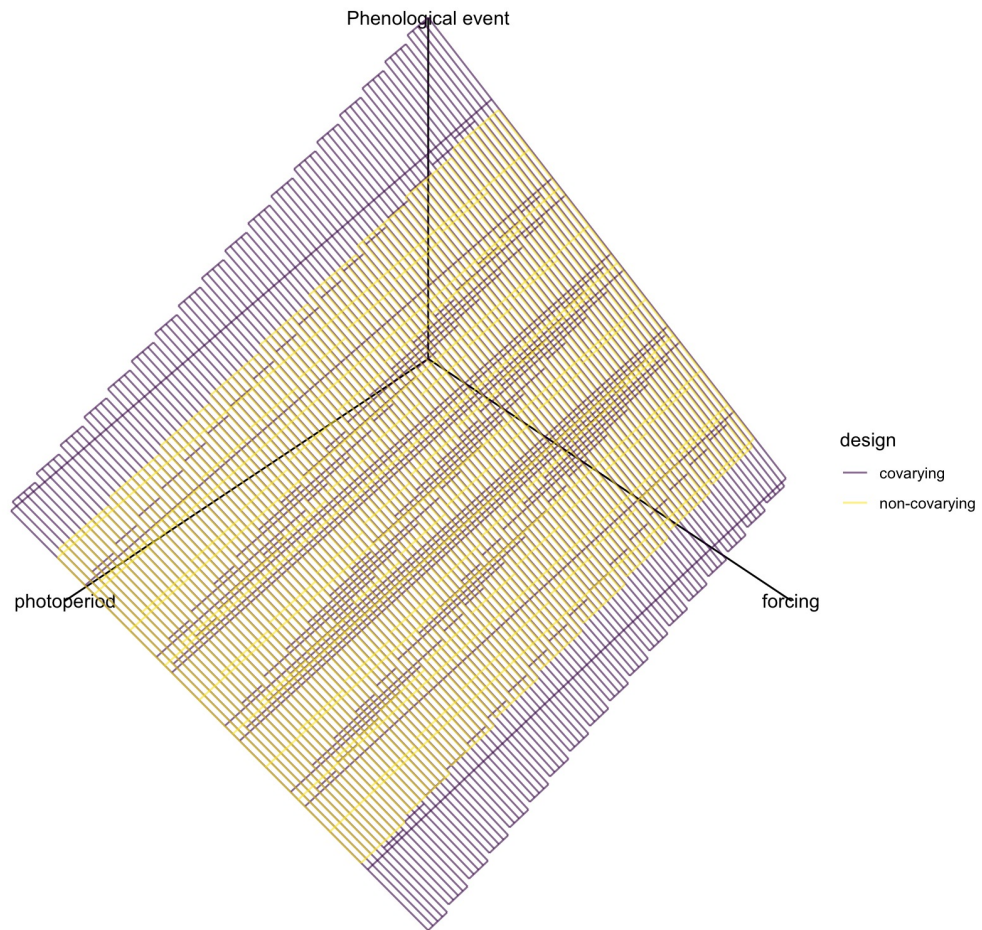
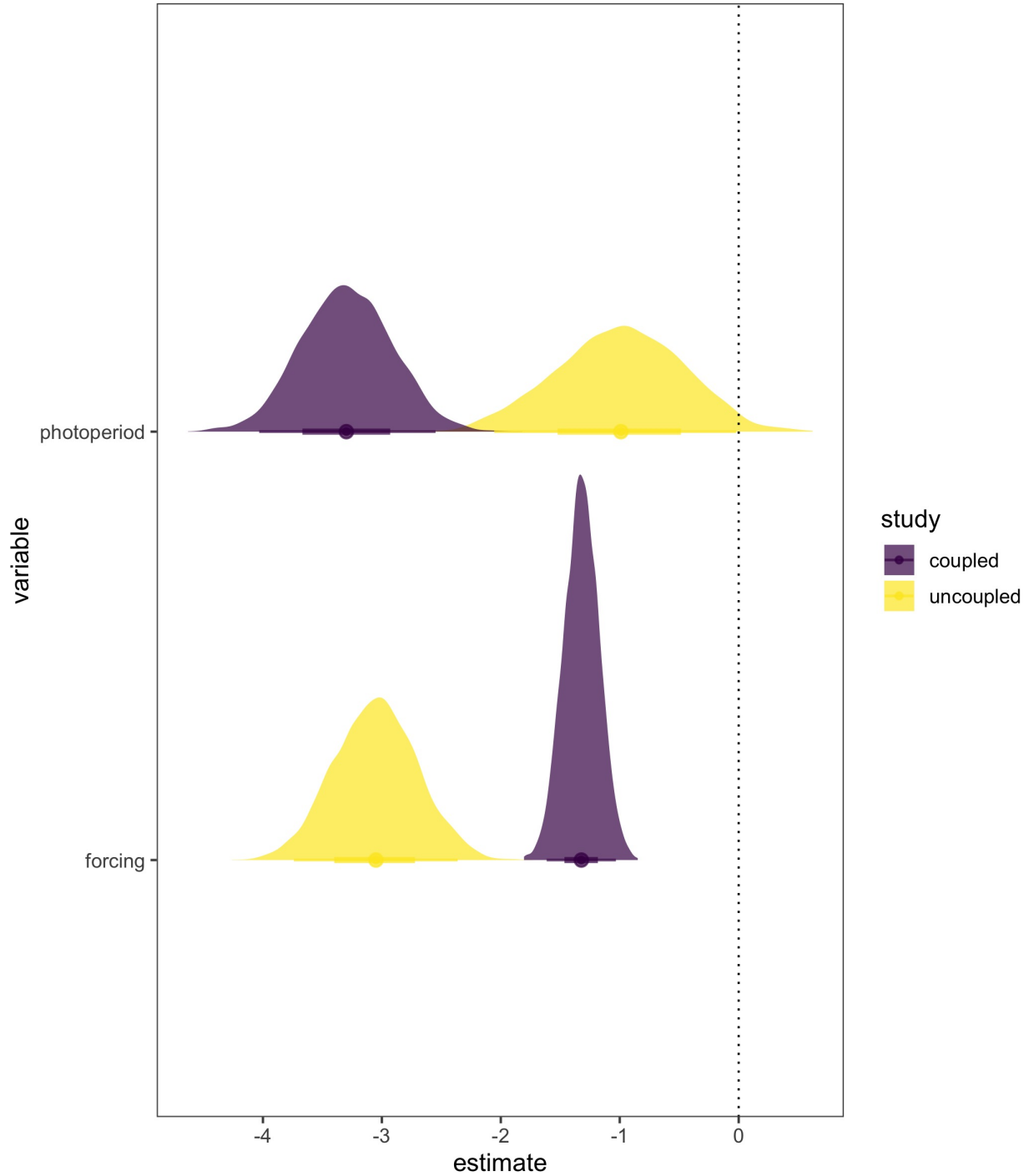Figure 3: Can we make a good 3d version of this figure?

Figure 4: Estimated effects of one hour increase in photoperiod and one degree C increase in forcing on the day of leaf expansion using an alternative methods of varying thermo-period relative to photoperiod. Dots indicate the estimated mean effect, thick and thin bars the 50% and 97.5% credible intervals respectively. The full posterior distributions are also depicted for each parameter estimate. The colors represent the study design effect, with purple representing study design in which photo- and thermo-period co-vary, and yellow representing a design where there periodicity of each variable is independent of the other.