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

January 4, 2022

1 Introduction

Across the tree of life, temperature and light availability shape a number of important biological processes including growth and metabolics 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.

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 ()) it is difficult to disentange their relative contributions to biological processes. Experimental manipulations of climate variables in artificial environments are ideal for mechanistically characterizing 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 () 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. ().

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 (). 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 (**?**) and these 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 mechansitically 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. Below we highlight a particular challenge that can arise in experiments seeking to estimate the interactive effects of temperature and light on biological outcomes when the periodicity of both variables is included. Our example deals with the phenology (timing of recurring life cycle events, e.g. leaf budburst, flowering) in temperate woody plants. However, 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.

We begin by breifly reviewing the minimum experimental elements required to robustly test interactions between two or more 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 a mathematical and experimental example. Finally, we conclude by outlining several possible solutions for overcoming these issues.

Phenology, explained

Decades experimental work in growth chambers have demonstrated that temperature (winter chilling and spring forcing) and photoperiod are the primary cues of phenology for plants in the temperate/boreal zones ().

Like many other biological processes, recent advances have demonstrated that plant phenological responses are nonlinear, due largely to interactions between cues (), highlighting the need for experiments to be designed to evaluate the strength of these interactions. For a brief overview of how temperature and light interact to influence phenology write this.

In order to have the statistical power to partition the individual and interactive effects of two or more variables in an experiment, one must:

1. Have at minimum of two treatment levels of at least two variables ().
2. Treatment levels must be full factorial (Fig. **??**a.). Full factorial designs are both balanced (Fig. **??**b.) and orthogonal (Fig. **??**c.); which is to say that all possible treatment combinations are applied and each treatment is independent of all others ().

These two critical elements may seem obvious but can be conspicuously absent from many published studies. In the case of woody plant phenology, using a recently published database (OSPREE: Observed spring phenolgoical responses in experimental enviornments ()) we found that out of 152 controlled environment experiments (across 93 studes) (), only 64 of them manipulated both light and temperature cues in the same experiment and only 15 of those did so with a design that was

both balanced and orthogonal (see Supplement for details). This notable dearth of robust tests of light and temperature interactions may stem from the common limitations of time, space, and resources that experimentalists often face (), but it may equally relate to a fundamental issues that arises from the fact that these variables themselves are comprised of multiple axes of variably.
Phenology experiments tend to manipulated two major axes of light and temperature variation:

1. Intensity: 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 irradiance present in the system (measured in lumens or watts).

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

(sentence here backing down from above statment and aknowldge there are other axes especiall with light) For phenology, it is generally accepted that photoperiodicity is the primary light cue to which plants respond ((() though see () regarding light intensity and phenology). For temperature, conventionally both intensity and periodicity drive phenological activity () and several metrics, (e.g. growing degree hours, thermal sums, growing degree days) that combine these two axes have been developed (). This assumption is well supported; under natural conditions diurnal temperature fluctuations in temperate regions can be quite large in the spring, and several studies have found that diurnal temperature variation strongly influences plant phenology (). In fact, several studies suggest that the even if thermoperiodicity is not an explicit treatment variable (i.e. manipulated systematically), incorporating it in experiments is essential for translating experimental results into real world predictions ().

It follows that a common approach in phenology experiments that seems to balance prior knowledge, biological realism and experimental inferences is to vary photoperiodicity, and thermal intensity and periodicity (). For example, a basic experiment might include a long (12 hours) and short (8 hours) photoperiod treatment and a high (30/20 C day/night) and low (20/10 C day/night) temperature treatment. Note that in this case, the thermoperiodicity is not an explicit treatment (both high and low temperature treatments employ a diurnal fluctuation of 10 C), and is simply incorporated in the design to enhance biological realism. At first glance, this design appears to meet the criteria of a full factorial design, multiple treatment levels that are balanced and orthogonal, with mean high/low temperature treatments (25 and 15 C respectively) and long/short photoperiod treatments applied in all possible combinations.

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 **??**.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 releastic, it make it statistically impossible to differentiate their independent and interactive effects on any given biological process.

Of the studies in the OSPREE database 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 relatively contribution of temperature and photoperiod cues to spring phenology remains a contentious debate in the phenology literature ().

Theoretical evidence

To estimate the potential impact of this experimental artifact on estimations of cue effects, we used the simple geometry of a plane to calculate the approximate maximum amount of the forcing (temperature) and photoperiod effect estimates that could potentially be mis-attributed due to the latent non-orthogonality of in a coupled therm0-photoperiod design. We based our calculations on estimates from the a study by in which the forcing and photoperiod effects on spring phenology were estimated to be approximately -9.5 and -4.5 Δ day of leaf out / Δ cue level. Our calcuations suggest that as much about 3.0 units of of these effects could be mis-attributed from a coupled photo-theromoperiod experiment (for the full calculations see the Supplement.) Because forcing is expected to be the more dominant cue (), we would expect the covariation of these variables may result in an over-estimation of the photoperiod effect and a weaker interaction estimate.

Experimental comparison

Our estimate of "how much of each cue estimate could be mis-attributed to the covariation of thermo- and photoperiod" can be rephrased as a maximum prediction of the expected difference in effect size estimates from a coupled photo- and thermo-period experiment and an uncouple one in which diurnal photoperiod and thermoperiod are varied independently. While we are aware of no experiments that explicitly compare these different designs, another later study by **?** utilized many overlapping treatment levels and species from the same sampling site and several treatment levels with the Flynn & Wolkovich (2018) study, but authors decoupled photo- and thermo-period, allowing for a reasonable comparison.

We subset each dataset (publicly available at HF 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, forcing and their interactions amoung study designs were on the same order our mathmatical predictions, and that the un-coupled design estimated a weaker (less negative) photoperiod effect, and stronger forcing and interaction effects than the coupled experimental design (3).

It is worth saying that there may be other factors driving the differences between these experiments. For example, both were conducted in different years, sampled different individual from the population, and used different methods for applying an addition temperature pre-treatment ,chilling (see Supplement). However, becasue 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 .

2 Paths Forward

In the sections above we have systematically demonstrated that experiments which 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 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. **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.

2. **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).

3. **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 on of these designs introduces another, which may in fact be an intrinsic property of any experimental manipulation. This fact should caution experimentalists to continue to think carefully about our designs and perhaps most importantly, remind us to be humble in our inference, and think critically about what is, and isn't accounted for in our work. **Instead, author should carefulyl consider the goals of a study.**

References

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.

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

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

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.

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.

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

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.

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.

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.

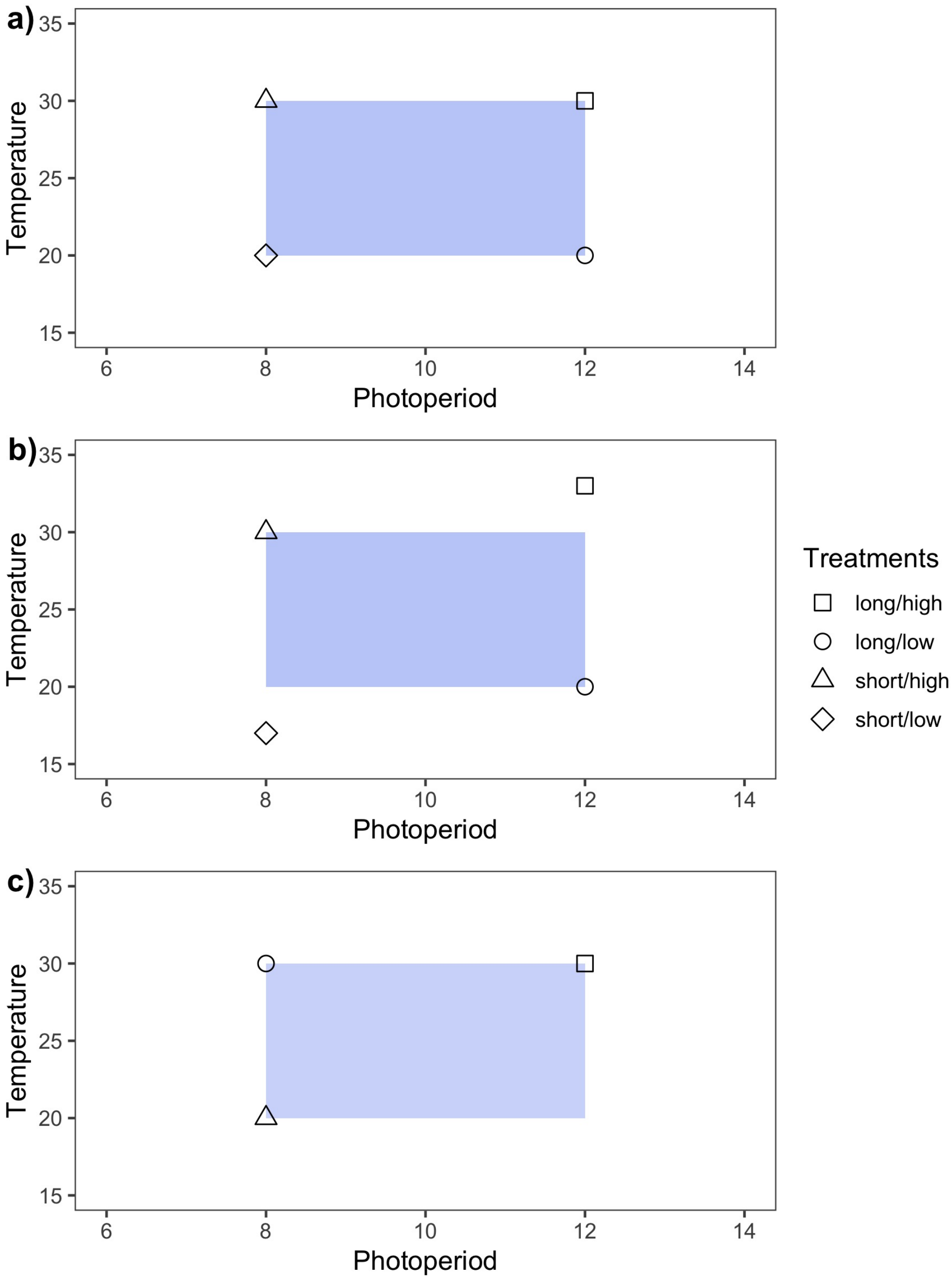

Figure 1: Idealized experimental designs demonstrate three approaches for varying temperature and light treatment level in controlled environment experiments. Design **a)** is fully factorial in that treatments levels are balanced and orthogonal. This design is appropriate for testing interactions between 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.

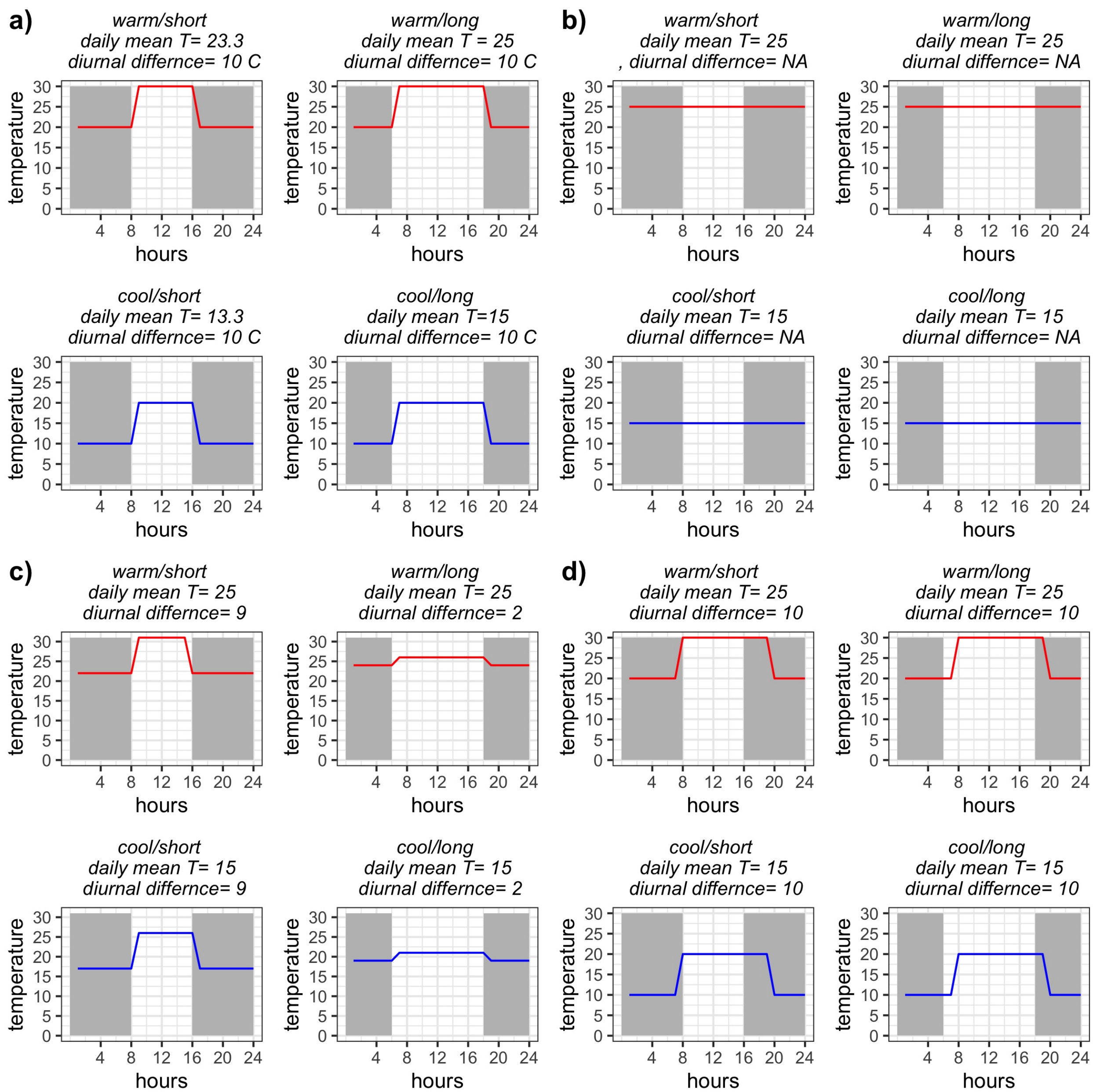

Figure 2: Conceptualized experimental designs to test temperature and daylength interactions on a biological response. In **a)** the design incorporates a standardize diurnal temperature fluctuation across all treatment. Because this thermoperiod is coupled with the photoperiod, while the same day and 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.

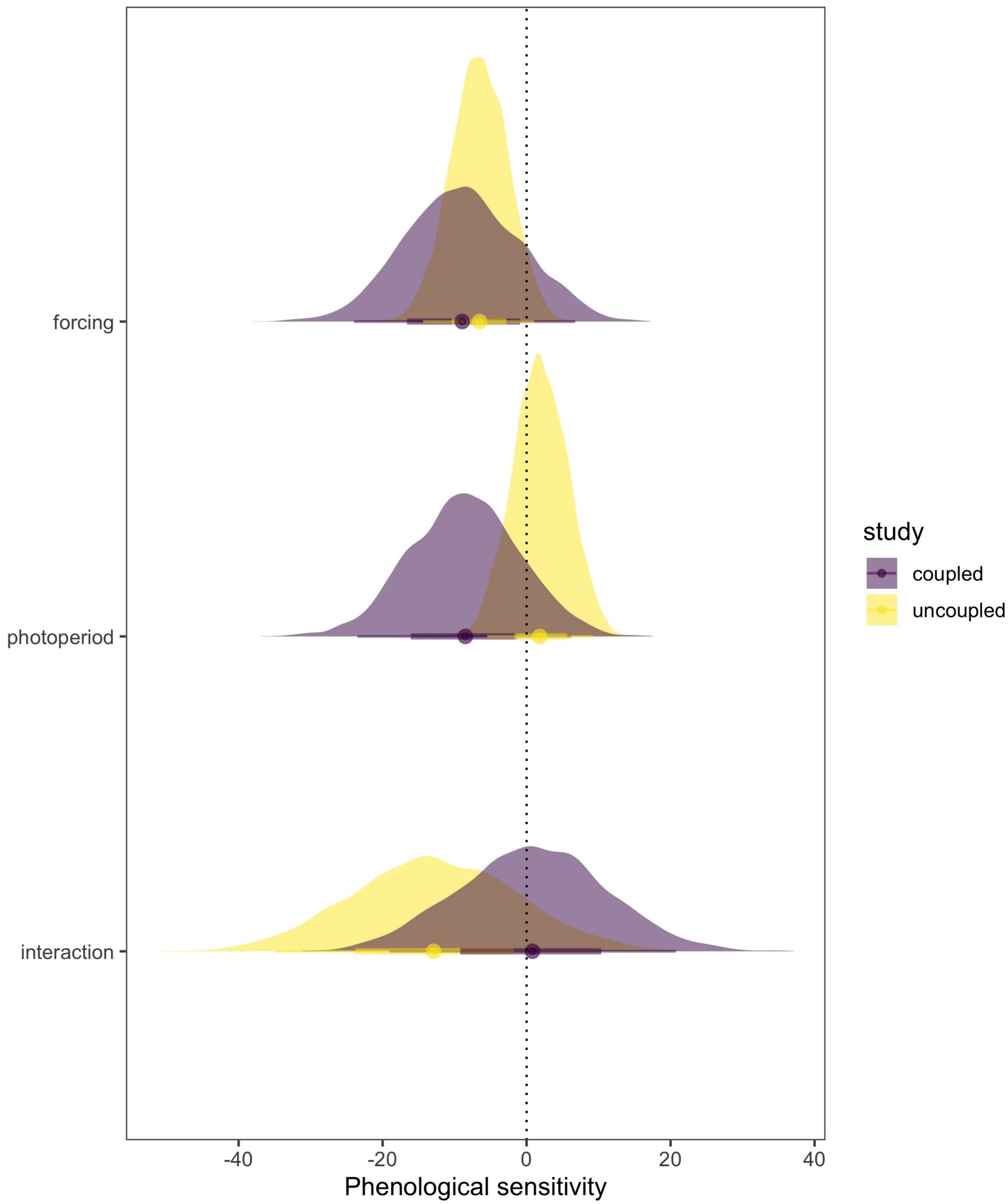

Figure 3: Need a caption but basically, estimates differs in expected ways