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

October 7, 2021

1 Introduction

Temperature and light dictate numerous biological processes including....*list here*(). These cues often interact in complex ways, and a major goal of biology is to quantify their both individual and interactive effects on biological activity (). This effort has only become more critical in recent decades as such measurements are essential for accurately predicting organisms' response to anthropogenic global change and informing biodiversity conservation planning and practices ().

It is difficult to evaluate the individual effects of temperature and light cues and their interactions in *in situ* observational studies because temperature and light variables tend to co-vary in nature (). However, well designed experiments in controlled environments (e.g. growth chambers, greenhouses, mesocosm) can be used to manipulate temperature and light variables independently, allowing for robust comparisons of their individual and interactive contributions to biological processes. Indeed, this approach has provided many fruitful insights regarding temperature and light signaling over the last century () and has great potentially to continue to advance fundamental and applied biological inquiries in the decades to come ().

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 (). 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 and for a brief overview of how temperature and light interact to influence phenology see the Suppliment. 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 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.

2   Testing interactions in contolled environmental

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 Z controlled environment studies (), only X of them manipulated both light and temperature cues in the same experiment and only Y of those X 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.

3   Axes of environmental variation in experiments and their implications

For most environmental variables, including light and temperature, there are two axes of variation that can be manipulated in an experiment.

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

2. <u>Periodicity</u>: 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.

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, orthogonality is violated, in that 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. Ultimately, this non-orthogonality introduces covariation among the photoperiod and temperature treatments, making it statistically impossible to differentiate their independent and interactive effects.

Of the X studies in the OSPREE database that manipulated both photoperiod and temperature, we found that at least Y 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 ().

Quantifying the uncertainty

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.

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 ? 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 (??).

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 .

4 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 (**??**b.). 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 (**??**c.). 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 ().

3. **Uncouple thermoperiod and photoperiod.** By varying thermoperiod and photoperiod independently (**??**d.), 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 (**?**). 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.
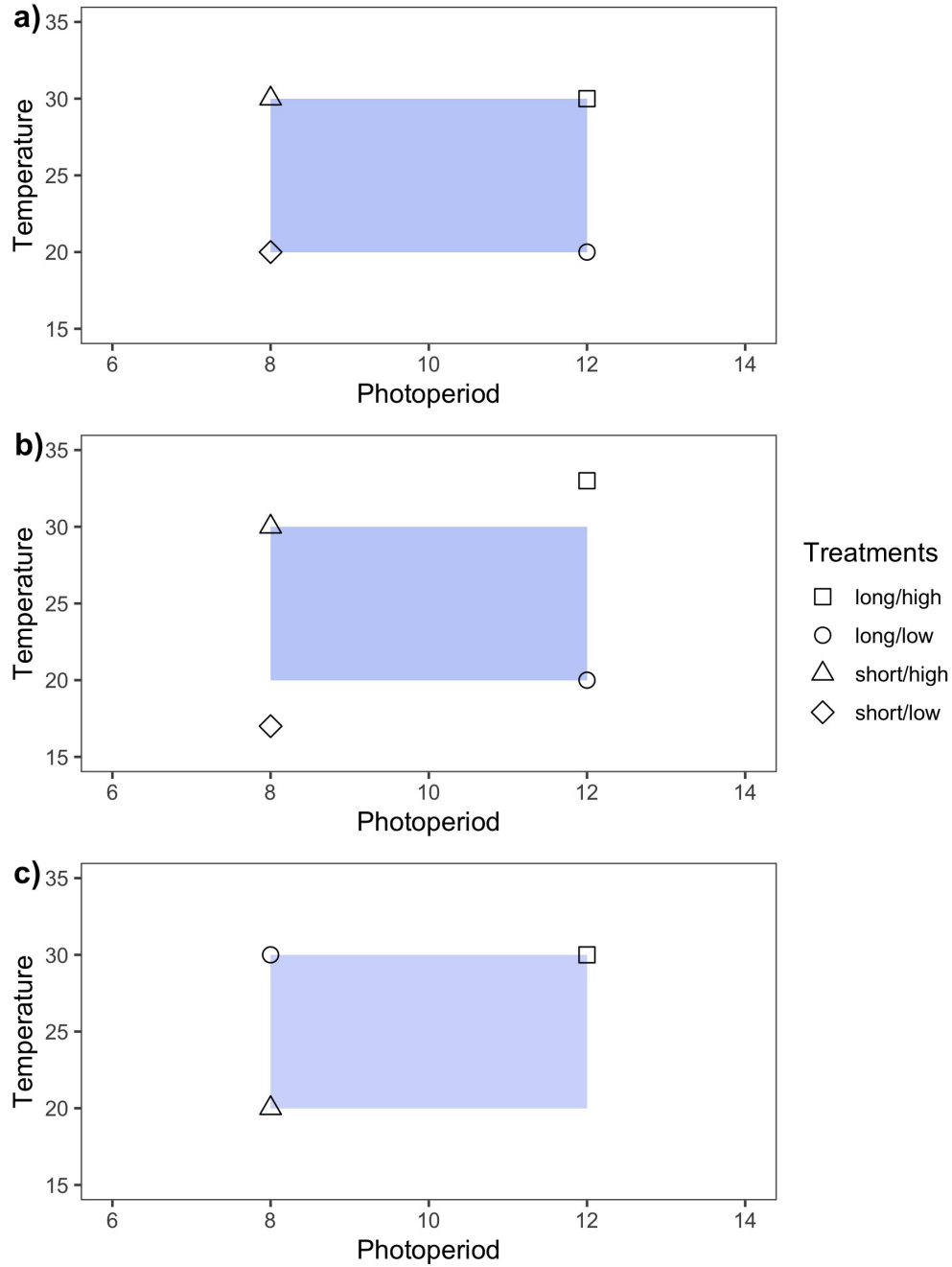
Figure 1: Idealized experimental designs demonstrate three approaches for varying temperature and light treatment level in controlled environment experiments. Design **a)** is fully factorial in that treatments levels are balanced and orthogonal. This design is appropriate for testing interactions between 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.
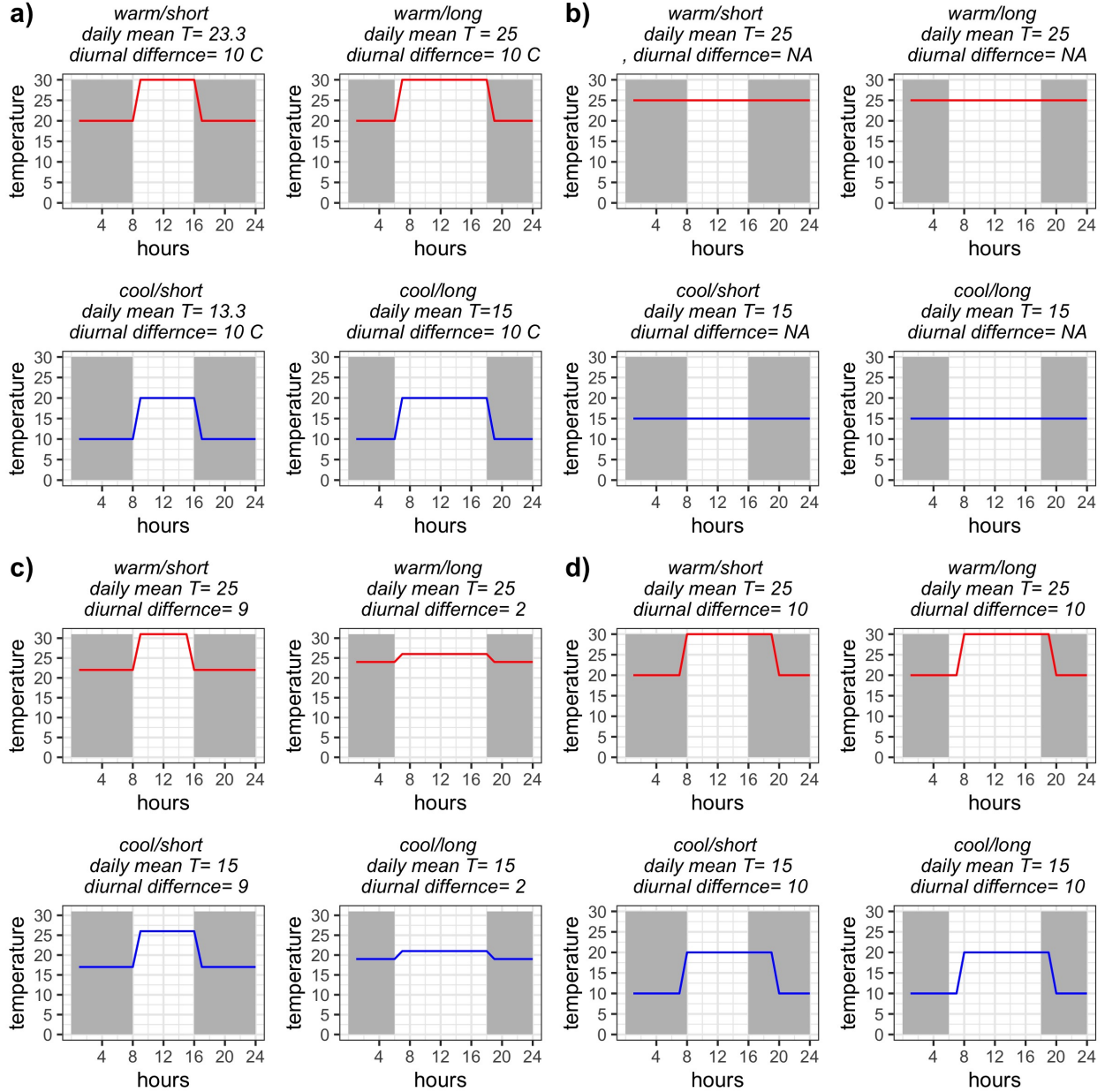
Figure 2: Conceptualized experimental designs to test temperature and daylength interactions on a biological response. In **a)** the design incorporates a standardize diurnal temperature fluctuation across all treatment. Because this thermoperiod is coupled with the photoperiod, while the same day and 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.
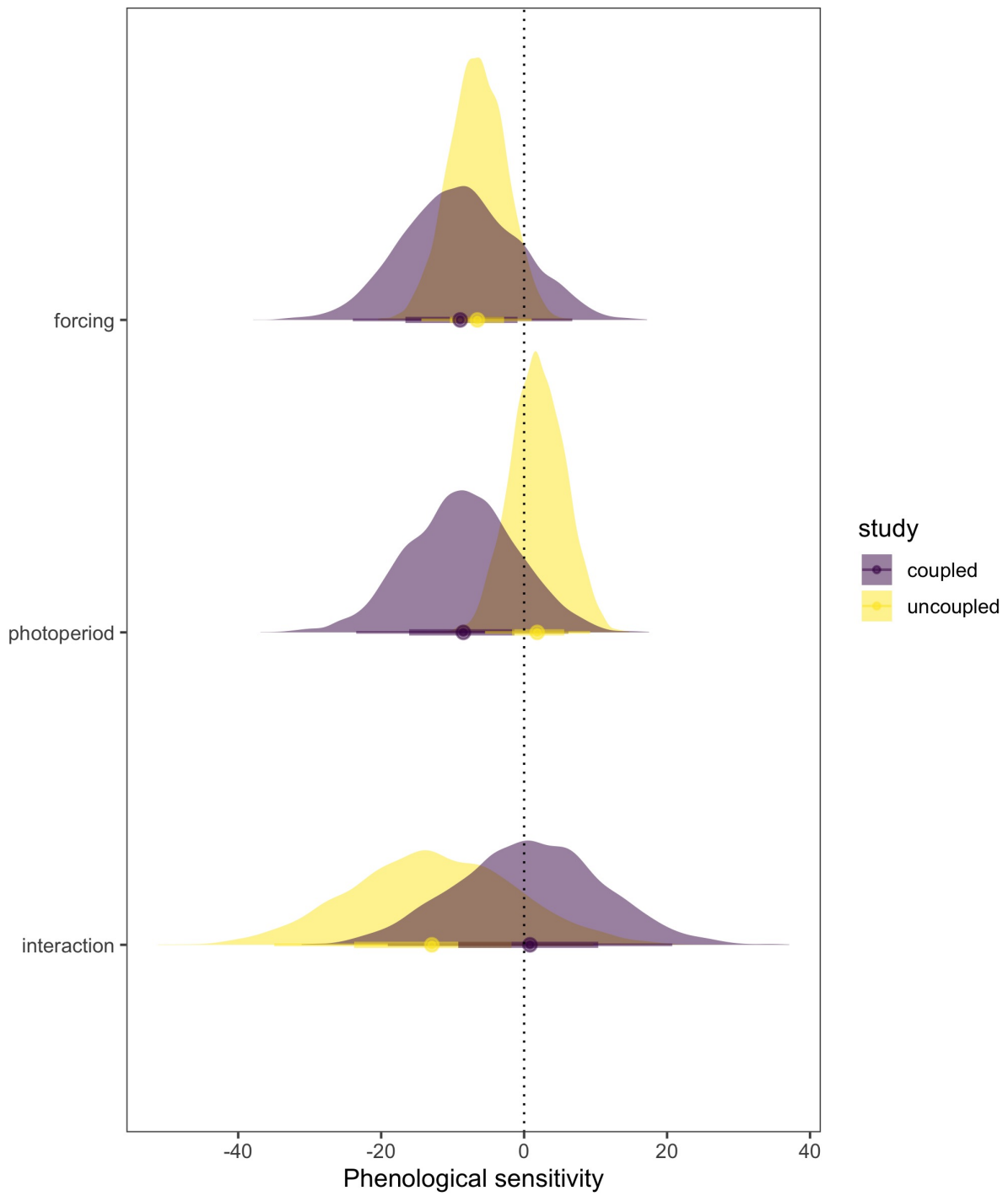
Figure 3: Need a caption but basically, estimates differs in expected ways