Title: **Sampling bias exaggerates a textbook example of a trophic cascade**

Authors: Elaine M. Brice*
Department of Wildland Resources and Ecology Center, Utah State University, Logan, UT 84322, USA
elbrice92@gmail.com
(301) 557-0521

Eric J. Larsen
Department of Geography and Geology, University of Wisconsin – Stevens Point, Stevens Point, WI 54481, USA
elarsen@uwsp.edu

Daniel R. MacNulty
Department of Wildland Resources and Ecology Center, Utah State University, Logan, UT 84322, USA
dan.macnulty@usu.edu

*Corresponding Author

Running Title: Sampling bias exaggerates trophic cascade

Keywords: aspen, carnivore, elk, predator indirect effects, nonrandom sampling, preferred browsing height, sampling bias, trophic cascade, ungulate, wolf

Article Type: Letter

Abstract Word Count: 149

Main Text Word Count: 5006

Number of References: 82

Number of Figures: 5

Number of Tables: 1

Author contributions: E.M.B., E.J.L., D.R.M. conceived and designed the study; E.J.L. collected most of the data; E.M.B. and D.R.M. analyzed the data and wrote the manuscript with input from E.J.L.

Data accessibility: should the manuscript be accepted, data will be archived in Dryad and the data DOI will be included at the end of the article.

ABSTRACT

Understanding how ecosystems respond to the loss and recovery of large predators is a major challenge because these free-living systems are difficult to sample properly. We show how an accepted practice of nonrandom sampling has confounded this understanding in a textbook system (Yellowstone National Park) where carnivore [*Canis lupus* (wolf)] recovery is often associated with a trophic cascade involving changes in herbivore [*Cervus canadensis* (elk)] behavior and density that promote plant regeneration. Long-term data indicate that a customary practice of sampling only the tallest young plants overestimated regeneration of overstory aspen (*Populus tremuloides*) by a factor of 3-8 compared to random sampling. Sampling only the tallest young plants favored plants taller than the preferred browsing height of elk and overlooked non-regenerating aspen stands. Our results demonstrate how seemingly minor departures from principled sampling can generate substantial misunderstandings about the strength of trophic cascades in response to large predator recovery.

INTRODUCTION

Knowledge about the occurrence and strength of trophic cascades (indirect effects of predators on plants and abiotic processes) is vital to understand the forces that structure food webs and to guide the practices of ecosystem conservation, restoration, and rewilding. Much of the empirical information about trophic cascades derives from tractable systems that are variously small scale, invertebrate, and captive (Ford & Goheen 2015; Piovia-Scott *et al.* 2017; Alston *et al.* 2019). Substantially less is known about trophic cascades in free-living large vertebrate systems due in part to the difficulty and cost of measuring such systems in accordance with basic sampling principles (Ford & Goheen 2015; Allen *et al.* 2017; Hayward *et al.* 2019).

Concern about sampling bias is a core issue in the longstanding debate about the cascading effects of reintroduced wolves in northern Yellowstone National Park (YNP; Bilyeu *et al.* 2008; Kauffman *et al.* 2013; Peterson *et al.* 2014; Winnie 2014; Fleming 2019; Hayward *et al.* 2019). In this system, the trophic cascade hypothesis is that wolves (*Canis lupus*) scared away and/or killed enough elk (*Cervus canadensis*) to allow woody deciduous plants (*Populus* spp., *Salix* spp.) to recover from decades of unchecked browsing. A primary support for this hypothesis is time series data showing annual decreases in browsing and annual increases in plant height following wolf reintroduction (Beschta & Ripple 2016; Beschta *et al.* 2018; Painter *et al.* 2018; Painter & Tercek 2020). A negative correlation between browsing and plant height is considered critical evidence of a trophic cascade because it demonstrates the mechanism connecting the lower two trophic levels: reduced browsing increases plant growth, leading to escape from browsing if plants grow tall (Beyer *et al.* 2007; Beschta & Ripple 2016; Beschta *et al.* 2018). The traditional, though untested, assumptions are that (*i*) plants taller than 200 cm

escape the reach of elk and recruit into the aspen overstory, and (*ii*) plants shorter than 200 cm are browsed with equal intensity.

Studies of comparable systems suggest that plant height has a nonlinear effect on elk browsing, such that browsing peaks at a preferred height of < 200 cm (Rounds 1979; Motta 2003; Renaud *et al.* 2003; Konôpka *et al.* 2018; Maxwell *et al.* 2019). If plant height beyond this preference reduces browsing, a negative browsing-plant height correlation may not reflect indicate a trophic cascade (Peterson *et al.* 2014). The more typical concern is that reports of decreased browsing and increased plant height after wolf reintroduction are based on biased sampling (Kauffman *et al.* 2013; Peterson *et al.* 2014; Winnie 2014; Fleming 2019).

Nonrandom sampling underpins nearly every published annual trend in browsing and height of young aspen (*Populus tremuloides*) that has been attributed to a cascading effect of wolves (Table 1). In this case, nonrandom sampling is the practice of measuring the three or five tallest young aspen within a stand. Five tallest data comprise most trends and they originate from one of two time series (Ripple & Beschta 2007; Painter *et al.* 2014). A third time series uses three tallest data (Halofsky *et al.* 2008). All three time series were built from a single year of sampling by retrospectively inferring the past browsing and height of sampled aspen using plant architecture methods (Keigley & Frisina 1998). Only one time series was built by annually sampling randomly selected young aspen: an unpublished dataset from E. Larsen that has received limited attention (Peterson *et al.* 2014, 2020).

Although the first study that sampled only the tallest young aspen acknowledged that such "data are only representative of the first recovering aspen…and not an estimate of aspen population response across Yellowstone's northern range" (Ripple & Beschta 2007), subsequent studies emphasized that the tallest young aspen represent a "leading edge" indicator of the future

condition of the aspen population (e.g., Ripple & Beschta 2012; Painter *et al.* 2014; Beschta *et al.* 2018). Peterson et al. (2020) explained that sampling the tallest young aspen "allows investigators to document the occurrence of any young aspen exceeding the upper browse level of elk (i.e., ~200 cm) in a given stand years before the average stem height attains this metric". However, the fate of the tallest young aspen may not represent that of the average young aspen if the former are exposed to more favorable growing conditions than the latter. For example, if the tallest young aspen exceed the preferred browsing height, and therefore experience progressively less browsing and more height growth as they grow taller, they may realize a faster growth rate that leads to a level of overstory recruitment that is unattainable for the average young aspen. Thus, sampling the tallest young aspen may misrepresent the current and future aspen population response to wolf reintroduction.

We assessed the scope of this bias with data from E. Larsen's long-term study of young aspen in northern YNP. These data reveal that the customary practice of sampling the tallest young aspen overestimated the potential for the typical young aspen to escape herbivory and reach tree height following wolf reintroduction, and thus exaggerated a popular example of a trophic cascade featured in many textbooks (e.g., Chapin *et al.* 2011; Boyce *et al.* 2012) and media (e.g., Morell 2007; Monbiot 2014). Our results highlight the importance of principled sampling for understanding trophic cascades following the loss and recovery of large predators.

MATERIALS AND METHODS

Study area

We measured young aspen in the portion of the northern Yellowstone elk winter range that lies within Yellowstone National Park (Fig. 1). This 995-km^2 area is defined by low-elevation (2000-

2600 m) grasslands and shrub steppes that fan out from the Yellowstone River and its tributaries near the Park's northern border. A variety of ungulates spend winter in the area including elk, *Bison bison* (bison), *Odocoileus hemionus* (mule deer), and *Alces alces* (moose). Elk and bison were the most abundant ungulate in the area during winter (1,200-6,000 elk; 1,400-3,200 bison), and bison numbers exceeded those of elk beginning winter 2011-2012 (Tallian *et al.* 2017). Wolves were reintroduced to YNP in 1995-1997 (Bangs & Fritts 1996) and their annual distribution was concentrated in the study area (Cassidy *et al.* 2020) where they hunted mainly elk (Metz *et al.* 2020). Other elk predators included *Puma concolor* (cougar), *Ursus arctos* (brown bear), and *Ursus americanus* (black bear) (Barber-Meyer *et al.* 2008; Ruth *et al.* 2019). Elk that moved beyond the study area into adjacent areas of Montana were also hunted by humans (Peterson *et al.* 2020).

Study population

Aspen is one of the few upland deciduous tree species in YNP and it is scattered across the study area in discrete stands on relatively moist mid-elevation benches, near streams, and along conifer forest/shrub steppe ecotones (Houston 1982). Aspen is a clonal species that mainly regenerates by root sprouting with occasional seedling establishment after disturbance. Root sprouting produces genetically identical trees from a common root system that may be substantially older than the age of the oldest tree, which rarely exceeds 150 years. Aspen requires moist soils and occurs mostly in areas with ≥ 38 cm of annual precipitation (Jones & DeByle 1985); the study area is near this lower limit (Larsen & Ripple 2003). Although aspen is a minor cover type in the arid portions of its range, it is a major source of biological diversity, providing habitat for numerous plants and animals (DeByle 1985; Mueggler 1985). Various ungulates eat the leaders and twigs of young stems, especially during winter**.** Persistent browsing of a stem's leader

106 prevents its growth to tree height, a process that has contributed to loss of overstory aspen in
107 portions of western North America, including the study area (National Research Council 2002;
108 Seager et al. 2013).

109 Historic landscape photographs suggest that overstory aspen covered ~4-6% of the study
110 area during 1880-1900 (Houston 1982; Meagher & Houston 1999), and aerial photographs
111 indicate that aspen coverage decreased to ~1% by the 1990s (Larsen & Ripple 2005). Pre-wolf
112 reintroduction planning studies did not predict any cascading effects of wolves on aspen
113 (reviewed in Yellowstone National Park 1997). Subsequent studies proposed that regeneration of
114 overstory aspen depended on wolves reducing elk browsing (White *et al.* 1998; Ripple & Larsen
115 2000). The first evidence of substantial numbers of unprotected aspen reaching tree height since
116 the early- to mid-20th century (Larsen & Ripple 2003; Kauffman *et al.* 2010) occurred about a
117 decade after wolf reintroduction (Ripple & Beschta 2007). See Appendix S1 in Supporting
118 Information for additional context.

119

120 **Data collection**

121 We measured browsing and height of young aspen in 113 plots distributed randomly across the
122 study area (Fig. 1). Each plot was a 1 m × 20 m belt transect located randomly within an aspen
123 stand that was itself randomly selected from an inventory of stands. The inventory was a list of
124 992 grid cells (240 m × 360 m) across the study area that contained at least one stand (Appendix
125 S1). A "stand" was a group of tree-size aspen in which each tree was ≤ 30 m from every other
126 tree. One hundred and thirteen grid cells were randomly selected from the inventory (~11% of
127 992 cells), one stand was randomly selected from each cell, and one plot was established in each
128 stand, with the start point and transect direction randomly determined (Ripple *et al.* 2001). We
129 verified that each plot likely represented a genetically unique stand (Appendix S1).

We measured aspen at the end of the growing season (late July to September), focusing on plants ≤ 600 cm tall, which we termed "young aspen." For each stand, we measured every young aspen rooted within a plot ('random stems'), as well as each of the five tallest young aspen rooted anywhere within the stand ('5T stems'). For all young aspen, we measured browsing (browsed or unbrowsed) and height of the leader stem (the tallest stem). A leader was 'browsed' if its growth from the previous growing season had been eaten. Most plots were measured nearly every year since 1999 (Ripple *et al.* 2001) and our analysis focused on data from 10 years (2007-2014, 2016-2017) in which sampled stands included measurements of random and 5T stems. Camera trap data indicate that elk were likely the primary ungulate species browsing young aspen in our plots during the study (Fig. S1).

Data analysis

First, we built empirical distributions of browsing levels and stem heights to visualize how these characteristics differed between 5T and random stems within and across years (Appendix S1). Second, we modeled annual changes in browsing and stem height, and tested how these changes differed between 5T and random stems. Third, we modeled the influence of stem height on browsing to identify the preferred browsing height (PBH) and browse escape height (BEH). We used these results to assess the prevalence of non-preferred stems in samples of 5T and random stems, and to test how estimates of overstory recruitment differed between random and 5T stems.

Annual changes in browsing, stem height, and overstory recruitment

We combined measurements of 5T and random stems into one dataset of all stems (N = 18,623) across all years (N = 10 years) and used generalized linear mixed models (GLMMs) to test how

153 the effect of year on browsing, height, and recruitment of stems differed by sampling method.
154 We treated the stem as the unit of analysis and used GLMMs with a Bernoulli distribution and a
155 logit link to separately analyze the probability a stem was browsed (1 = browsed; 0 = not
156 browsed) and recruited (1 = recruited; 0 = not recruited), and GLMMs with a gamma distribution
157 and a log link to analyze stem height (cm), which took only non-negative values that were
158 strongly right-skewed. A stem 'recruited' if it exceeded a presumed BEH of 200-cm or 300-cm.
159 Year was an integer that ranged from 1 (2007) to 11 (2017) and sampling method was a dummy
160 variable (1 = 5T sampling; 0 = random sampling). If 5T sampling estimates a faster annual
161 decrease in browsing and faster annual increases in height and recruitment compared to random
162 sampling, we expected (*i*) models with a year × method interaction fit the data better than models
163 with only main effects for these variables, and (*ii*) the sign of this interaction was negative in the
164 browsing model and positive in the height and recruitment models. We used likelihood ratio tests
165 to compare models.
166 All GLMMs included a random intercept for stand identity, and GLMMs of browsing and
167 stem height also included a random slope for year. We observed too few stems >200 cm or >300
168 cm to estimate a GLMM of recruitment with a random slope for year. The random intercept
169 controlled for correlation among measurements of the same stand in multiple years and
170 unmeasured stand-related effects including soil, water, and light conditions; the random slope
171 permitted stand-specific annual trends in herbivory and stem height. We estimated average
172 marginal effects (AMEs) from GLMMs to quantify and compare annual changes of 5T and
173 random stems. AMEs describe the average effect of changes in explanatory variables on the
174 change in a response variable and are useful for interpreting generalized linear models (Leeper

175 2021). We used χ^2 statistics to test how the AMEs of year on browsing, stem height, and
176 recruitment differed between 5T and random stems.
177 We calculated population-averaged fitted values from best-fit GLMMs by deriving
178 marginal expectations of the responses averaged over the random effects but conditional on the
179 observed variables. We refit year as a categorical factor and plotted the associated fitted values to
180 illustrate the distribution of the underlying data after controlling for stand-level heterogeneity,
181 and to assess the negative correlation between mean annual browsing and mean annual stem
182 height.
183 Following previous studies (Painter *et al.* 2014, 2015, 2018; Beschta *et al.* 2016), we also
184 estimated recruitment at the stand level as the percentage of sampled stands with stems taller
185 than the presumed reach of elk. We calculated this separately for 5T and random stems as the
186 annual percentage of sampled stands in which the median stem height exceeded 200 cm or 300
187 cm. Consistent with previous studies, recruitment estimates from 5T stems excluded stands that
188 produced no young aspen.
189
190 *Preferred browsing height and browse escape height*
191 We modeled the effect of stem height on browsing to estimate PBH and BEH. We estimated
192 separate GLMMs for 5T stems (N = 4,265) and random stems (N = 14,358), and included
193 crossed random intercepts for stand identity and year to account for (*i*) correlation between
194 measurements taken on the same stand in multiple years and on multiple stands in the same year,
195 and (*ii*) unmeasured stand- and year-related effects.
196

We used piecewise linear splines to identify the PBH, which we defined as the height threshold beyond which browsing probability decreased with further height increase. We compared models with a single height threshold placed from 10-200 cm (first at 10 cm and then 1 cm intervals), a model with no height threshold, and an intercept-only model. We selected thresholds *a priori* based on the traditional assumption that stems taller than 200 cm escape browsing (Kay 1990). We constructed variables containing a linear spline for stem height so that the estimated coefficients measure the slopes of the segments before and after the threshold. We evaluated competing models using marginal likelihoods and information-theoretic statistics (AIC_c; Burnham & Anderson 2002), and used AMEs of the best models to estimate how browsing probability changed with increasing stem height. We also used the best models to identify the BHE, which was the stem height at which browse probability was near zero.

RESULTS

Empirical distributions of browsing and stem height

Probability densities indicate that low browsing (≤ 20% stems browsed) and tall height (>100-cm) were more characteristic of 5T stems than of random stems, and that these differences increased from 2007 to 2017 (Fig. 2; Fig. S2). During this period, median height of 5T stems tracked increases in the 85-90th percentile height of random stems, which increased 4-5 times faster than did the median height of random stems (Fig. S3).

Annual changes in browsing and stem height

GLMMs of browsing and stem height with a year × method interaction fit the data better than did those with only main effects for these variables (browsing: χ_1^2= 134.6, $P < 0.001$; height: χ_1^2=

47.2, $P < 0.001$). The interaction was negative in the browsing model (β = -0.18; 95% CI = -0.22, -0.15; $P < 0.001$) and positive in the height model (β = 0.016; 95% CI = 0.012, 0.021; $P < 0.001$), which indicates that 5T sampling estimated a faster decrease in browsing and faster increase in height compared to random sampling (Fig. 3a, 3b). Specifically, browsing decreased 3.9 percentage points·year^{-1} (95% CI = 3.2, 4.6) for 5T stems versus 1.8 percentage points·year^{-1} (95% CI = 1.2, 2.5) for random stems (χ_1^2= 73.3, $P < 0.001$). Height increased 19.8 cm·year^{-1} (95% CI = 15.8, 23.9) for 5T stems versus 7.7 cm·year^{-1} (95% CI = 6.0, 9.4) for random stems (χ_1^2= 93.1, $P < 0.001$). The negative correlation between estimates of mean annual browsing (points in Fig. 3a) and mean annual height (points in Fig. 3b) was 47% stronger for 5T stems (Pearson's correlation coefficient, r = -0.97; $P < 0.001$) compared to random stems (r = -0.60; P = 0.07; Fig. 3c).

Preferred browsing height and browse escape height

The best-fit GLMMs ($\Delta AIC_c = 0$) indicate that the PBH was 132 cm for 5T stems and 122 cm for random stems (Fig. 4a; Tables S1, S2). Below these heights, each 10 cm increase in height increased browsing by 0.4 percentage points for 5T stems (95% CI = -0.3, 1.0; P = 0.26) and 0.5 percentage points for random stems (95% CI = 0.1, 0.8; P = 0.005). Above these heights, each 10 cm increase in height decreased browsing by 3.1 percentage points for 5T stems (95% CI = 2.9, 3.2; $P < 0.001$) and 3.5 percentage points for random stems (95% CI = 3.2, 3.7; $P < 0.001$). Stems exceeding the PBH were 1.6-5.1 times (mean ± SE = 2.8 ± 0.40) more prevalent in the sample of 5T stems than in the sample of random stems (Fig. 4b). The best-fit GLMMs indicate that browsing of stems >200 cm was as high as 0.45 (95% CI = 0.40, 0.51) for 5T stems and 0.35 (95% CI = 0.28, 0.42) for random stems. Browsing was negligible (< 0.07) only after stems exceeded ~300 cm (Fig. 4a).

244

245 **Overstory recruitment**

246 GLMMs of the probability that a stem exceeded 200 cm or 300 cm with a year × method
247 interaction fit the data better than did those with only main effects for these variables (200 cm:
248 χ_1^2= 50.1, $P < 0.001$; 300 cm: χ_1^2= 72.9, $P < 0.001$). The positive sign of the interaction (200 cm:
249 β = 0.23; 95% CI = 0.17, 0.30; $P < 0.001$; 300 cm: β = 0.40; 95% CI = 0.30, 0.49; $P < 0.001$)
250 indicates that 5T sampling estimated a faster increase in recruitment compared to random
251 sampling (Fig. 5a, 5b). Recruitment of stems >200 cm increased 4.6 percentage points·year^{-1}
252 (95% CI = 3.9, 4.2) for 5T stems versus 1.5 percentage points·year^{-1} (95% CI = 1.1, 1.9) for
253 random stems (χ_1^2= 41.3, P<0.001), and recruitment of stems >300 cm increased 3.1 percentage
254 points·year^{-1} (95% CI = 2.5, 3.6) for 5T stems versus 0.49 percentage points·year^{-1} (95% CI =
255 0.32, 0.67) for random stems (χ_1^2= 121.0, P<0.001).

256 Stand-level assessments indicate that 5T sampling estimated a total increase in overstory
257 recruitment from 2007 to 2017 that was 4-8 times greater than that estimated by random
258 sampling. The annual percentage of stands with a median stem height greater than 200 cm (300
259 cm) increased from 13-68% (5-51%) for 5T sampling versus 1-13% (0-6%) for random
260 sampling, and in any given year, the percentage estimated from 5T sampling was 4-26 times (5-
261 31 times) greater than that estimated by random sampling (Fig. 5c, 5d). Random sampling
262 revealed that young aspen were annually absent in 2-9% of stands and 11-19% of plots within
263 stands (Fig. S4, S5). Every stand produced young aspen during at least one year of the study,
264 whereas 7% of plots consistently produced no young aspen.

265

266 **DISCUSSION**

267 To the extent that annual decreases in browsing and annual increases in height of woody
268 deciduous plants in northern Yellowstone National Park reflect the cascading effects of
269 reintroduced wolves (Beyer *et al.* 2007; Beschta & Ripple 2016; Beschta *et al.* 2018), our results
270 show that an accepted practice of nonrandom sampling – one featured in a dozen peer-reviewed
271 publications over more than a decade (Table 1) – exaggerated the aspen population response to
272 wolf reintroduction. Sampling only the five tallest young aspen within a stand estimated annual
273 changes in browsing, stem height, and overstory recruitment that were significantly faster than
274 those estimated by random sampling of all young aspen within a stand (Fig. 3a-b, Fig. 5). We
275 suggest that 5T sampling exaggerated the aspen response for the following three reasons.
276 First, 5T sampling favored stems that were taller than the preferred browsing height
277 (PBH) of elk (Fig. 4). Stems taller than the PBH likely grew faster than stems shorter than the
278 PBH because the former were browsed at a decreasingly low rate as they grew taller, whereas the
279 latter were browsed at an increasingly high rate as they grew taller (Fig. 4a). Stems taller than the
280 PBH may have also allocated relatively fewer resources to defense chemistry expression as
281 browsing decreased (Rhodes *et al.* 2017), which would have further accelerated their height
282 growth (Lindroth & St. Clair 2013).
283 The PBH that we estimated (122-132 cm) is similar to the PBH of elk and comparable
284 cervids (red deer, *Cervus elaphus*) in other systems (Rounds 1979; Motta 2003; Renaud *et al.*
285 2003; Konôpka *et al.* 2018). Experiments indicate that the PBH equals ungulate shoulder height,
286 and that ungulates prefer to browse at this height because it minimizes neck angle which in turn
287 maximizes foraging efficiency (Renaud *et al.* 2003). Although published shoulder height
288 estimates for Yellowstone elk are lacking, the literature suggests a range of 119-168 cm (Murie
289 1951; Blood & Lovaas 1966; Blood & Smith 1984; Hudson & Haigh 2002). Camera trap photos

290 of elk browsing aspen in our plots is consistent with the hypothesis that elk prefer to browse at
291 the level of their shoulder height (Fig. S6).

292 Our PBH estimate is also similar to (or less than) previous estimates of mean height of
293 the tallest young aspen (Table 1), which implies that these studies also measured a large number
294 of stems taller than the PBH. These estimates refer to the 1-2 years in which field sampling
295 occurred as well as to preceding years that lacked field sampling. Estimates for the latter were
296 retrospectively inferred using plant architecture methods. Although such methods might not be
297 useful when herbivory is intense (Ripple & Beschta 2007), the prevailing assumption is that such
298 estimates are accurate indicators of browsing and stem height across decadal time scales. To the
299 extent that this assumption is valid, our results suggest that reported trends in browsing and stem
300 height based on the tallest young aspen (Table 1) do not represent trends in the aspen population
301 at large because they too were biased toward stems that were taller than the PBH.

302 A challenge highlighted by the PBH is that stem height is both a cause and an effect of
303 reduced browsing. Thus, the negative correlation between browsing and stem height (e.g., Fig.
304 3c) that some consider critical evidence of a trophic cascade (Beyer *et al.* 2007; Beschta &
305 Ripple 2016; Beschta *et al.* 2018) is not an exclusive indicator of browsing suppressing stem
306 height. Rather, the negative correlation also indicates that stem height suppresses browsing,
307 which reflects previous findings that factors besides browsing control stem height (Romme *et al.*
308 1995). This relationship helps explain why the negative correlation between browsing and stem
309 height was stronger for 5T stems compared to random stems (Fig. 3c): the stronger negative
310 correlation was consistent with the strong negative effect of stem height on browsing that most
311 5T stems experienced given most were taller than the PBH. Similarly, the weak negative
312 correlation for random stems was at least partially due the countervailing positive effect of stem

313 height on browsing that most of these stems experienced since most were shorter than the PBH
314 (Fig. 4). These results suggest that a negative correlation between browsing and stem height is
315 not reliable evidence of a trophic cascade because it does not represent an unambiguous causal
316 link between reduced browsing and increased stem height.

317 Second, the tallest young aspen were most likely exposed to the best growing conditions.
318 Tall stature is itself an indicator of a locally productive resource environment because only stems
319 exposed to sufficient sunlight, moisture, and soil quality have the capacity to grow tall (Brown *et*
320 *al.* 2006; Hansen *et al.* 2016). Thus, more productive resource conditions may have contributed
321 to the faster annual height growth and overstory recruitment of 5T stems (Fig. 3b, Fig. 5). Faster
322 height growth could have in turn contributed to the faster decrease in herbivory (Fig. 3a). It is
323 also possible that better resource conditions permitted 5T stems to respond more rapidly to a
324 reduction in herbivory (Fig. 3c).

325 Third, 5T sampling did not detect the absence of aspen regeneration. By definition, 5T
326 sampling measures only stands and locations within stands that produce young aspen. Stands and
327 locations without young aspen are not sampled and therefore excluded from stand-level estimates
328 of overstory recruitment calculated as the percentage of sampled stands with stems taller than the
329 presumed reach of elk (e.g., Fig. 7b in Painter *et al.* 2014, Fig. 2b in 2015; Fig. 5d in Beschta *et*
330 *al.* 2016). This omission likely inflates stand-level estimates of overstory recruitment given our
331 finding that young aspen were annually absent in 2-9% of stands and 11-19% of plots within
332 stands (Fig. S4). Overestimating recruitment because the absence of regeneration is
333 undocumented in poorly recruiting stands represents a form of visibility bias that has received
334 little attention in discussions about the cascading effects of wolves on aspen.

A separate source of bias that further inflates stand- and stem-level estimates of overstory recruitment is the traditional assumption that stems taller than 200 cm escape herbivory. Our analysis indicates that escape from herbivory is not certain until stems exceed approximately 400 cm (Fig. 4a). This suggests that previous estimates of overstory recruitment that assume a browse escape height of 200 cm (Painter *et al.* 2014, 2015, 2018; Beschta *et al.* 2016) are best interpreted as maximum estimates.

Our results do not support the hypothesis that the tallest young aspen represent a "leading edge" indicator of a "broader shift in plant community dynamics for northern [YNP] aspen stands" (Beschta *et al.* 2018). Contrary to the expectation that characteristics of the average young aspen should resemble those of the tallest young aspen over time (Peterson *et al.* 2020), we found that mean levels of herbivory, height, and recruitment of random stems increasingly differed from those of 5T stems (Fig. 3a-b; Fig. 5). Trends in the height distributions of 5T stems relative to those of random stems indicate that this widening gap was because 5T stems represented a minority of young aspen that substantially outperformed the majority of young aspen. Specifically, height increase of 5T stems tracked the height increase of the 85-90th percentile of random stems, which was substantially faster than the height increase of more than half of random stems (Fig. S3). Assuming our sample of random stems was representative of the population at large, our results suggest that the dynamics of 5T stems represent those of the top-performing ~10-15% of young aspen.

The general lesson from our study is that apparently minor deviations from principled sampling can lead to major misunderstandings about the strength of trophic cascades in response to large predator recovery. Had we sampled only the tallest young aspen, we would have concluded, incorrectly, that wolf reintroduction was associated with a strong regenerative

358 response of aspen. Instead, our random sampling design, which included a more representative
359 sample of aspen stands and stems, indicated that wolf reintroduction was associated with a more
360 limited regenerative response that did not reverse the deterioration of all aspen stands (e.g., Fig.
361 S5). A limited aspen response is consistent with (*i*) documented losses of other aspen stands in
362 the study area since wolf reintroduction (Fig. S7; Beschta *et al.* 2020), (*ii*) evidence that wolves
363 had weak and/or inconsistent effects on elk foraging behavior (Laundré *et al.* 2001; Childress &
364 Lung 2003; Wolff & Van Horn 2003) and habitat selection (Kohl *et al.* 2018, 2019; Cusack *et al.*
365 2019), and weak to moderate effects on elk population density (Vucetich *et al.* 2005; Peterson *et*
366 *al.* 2014; Macnulty *et al.* 2020), and (*iii*) theory that predicts weak cascading effects in food
367 webs like Yellowstone that are resource-limited, spatially heterogeneous, and reticulated (Ford &
368 Goheen 2015). We also note that climate change may have contributed to the limited aspen
369 response that we documented given that many of our plots occurred in or adjacent to areas
370 projected to become unsuitable for aspen due to anthropogenic climate forcing (Piekielek *et al.*
371 2015; Fig. S8).

372 Understanding how ecosystems respond to the loss and/or addition of large predators is
373 vital to resolving broader debates about the forces that structure food webs, determine species
374 abundance, and deliver ecosystem services (Estes *et al.* 2011; Dobson 2014; Ripple *et al.* 2014;
375 Atkins *et al.* 2019). Our study, which focused on a textbook example of large predator
376 extirpation and reintroduction, demonstrates how apparently harmless deviations from principled
377 sampling can distort this understanding. Nonrandom sampling overestimated the strength of a
378 hypothesized trophic cascade in the system we studied, but it may underestimate cascading
379 effects in other systems. In observational studies that lack control, randomization is one of the
380 few available protections against unreliable inferences and the misguided policy and

management decisions they may inspire. Growing concerns about the reliability of research findings in ecology (e.g., Fraser *et al.* 2002; Fidler *et al.* 2017; O'Grady 2020) emphasize the importance of principled sampling in studies of trophic cascades and other ecological phenomena.

ACKNOWLEDGEMENTS

We thank R. Renkin and J. Klaptosky for administrative and field assistance. We also thank J. Brodie and N. Piekielek for sharing important supplemental data. We acknowledge support from the U.S. National Science Foundation (DGE-1633756), University of Wyoming-National Park Service Small Grant Program (1003867-USU), Yellowstone National Park, Utah State University Ecology Center, and the S.J. and Jessie E. Quinney Doctoral Fellowship. We thank Dr. William Ripple for serving as principal investigator on the scientific research and collecting permits issued by Yellowstone National Park that authorized aspen data collection during 1999-2009.

Authorship: E.M.B., E.J.L., D.R.M. conceived and designed the study; E.J.L. collected most of the data; E.M.B. and D.R.M. analyzed the data and wrote the manuscript with input from E.J.L.

Data availability: Data are archived in Dryad (##).

REFERENCES

Allen, B.L., Allen, L.R., Andrén, H., Ballard, G., Boitani, L., Engeman, R.M., *et al.* (2017). Can we save large carnivores without losing large carnivore science? *Food Webs*, 12, 64–75.

Alston, J.M., Maitland, B.M., Brito, B.T., Esmaeili, S., Ford, A.T., Hays, B., *et al.* (2019). Reciprocity in restoration ecology: When might large carnivore reintroduction restore ecosystems? *Biol. Conserv.*, 234, 82–89.

Atkins, J.L., Long, R.A., Pansu, J., Daskin, J.H., Potter, A.B., Stalmans, M.E., *et al.* (2019). Cascading impacts of large-carnivore extirpation in an African ecosystem. *Science*, 364, 173–177.

Bangs, E.E. & Fritts, S.H. (1996). Reintroducing the gray wolf to central Idaho and Yellowstone National Park. *Wildl. Soc. Bull.*, 24, 402–413.

Barber-Meyer, S.M., Mech, L.D. & White, P.J. (2008). Elk calf survival and mortality following wolf restoration to Yellowstone National Park. *Wildl. Monogr.*, 169, 1–30.

Beschta, R.L., Painter, L.E., Levi, T. & Ripple, W.J. (2016). Long-term aspen dynamics, trophic cascades, and climate in northern Yellowstone National Park. *Can. J. For. Res.*, 46, 548–556.

Beschta, R.L., Painter, L.E. & Ripple, W.J. (2018). Trophic cascades at multiple spatial scales shape recovery of young aspen in Yellowstone. *For. Ecol. Manag.*, 413, 62–69.

Beschta, R.L. & Ripple, W.J. (2009). Large predators and trophic cascades in terrestrial ecosystems of the western United States. *Biol. Conserv.*, 142, 2401–2414.

Beschta, R.L. & Ripple, W.J. (2016). Riparian vegetation recovery in Yellowstone: The first two decades after wolf reintroduction. *Biol. Conserv.*, 198, 93–103.

Beschta, R.L., Ripple, W.J., Kauffman, J.B. & Painter, L.E. (2020). Bison limit ecosystem recovery in northern Yellowstone. *Food Webs*, e00142.

Beyer, H.L., Merrill, E.H., Varley, N. & Boyce, M.S. (2007). Willow on Yellowstone's northern range: Evidence for a trophic cascade? *Ecol. Appl.*, 17, 1563–1571.

Bilyeu, D.M., Cooper, D.J. & Hobbs, N.T. (2008). Water tables constrain height recovery of willow on Yellowstone's northern range. *Ecol. Appl.*, 18, 80–92.

Blood, D.A. & Lovaas, A.L. (1966). Measurements and weight relationships in Manitoba elk. *J. Wildl. Manag.*, 30, 135–140.

Blood, D.A. & Smith, G.W. (1984). Weights and measurements of Roosevelt elk on Vancouver Island. *The Murrelet*, 65, 41–44.

Boyce, M.S., Merrill, E.H. & Sinclair, A.R.E. (2012). Wildlife management. In: *The Princeton guide to ecology* (eds. Levin, S.A., Carpenter, S.R. & Godfray, H.C.J.). Princeton University Press, Princeton, NJ, USA.

Brown, K., Hansen, A.J., Keane, R.E. & Graumlich, L.J. (2006). Complex interactions shaping aspen dynamics in the Greater Yellowstone Ecosystem. *Landsc. Ecol.*, 21, 933–951.

Burnham, K.P. & Anderson, D.R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. 2nd ed. Springer, New York, NY, USA.

Cassidy, K.A., Smith, D.W., Stahler, D.R., MacNulty, D.R., Stahler, E.E. & Metz, M.C. (2020). Territoriality and competition between wolf packs. In: *Yellowstone wolves: Science and discovery in the world's first national park* (eds. Smith, D.W., Stahler, D.R. & MacNulty, D.R.). University of Chicago Press, Chicago, IL, USA, pp. 61–76.

Chapin, F.S., Matson, P.A. & Vitousek, P.M. (2011). *Principles of terrestrial ecosystem ecology*. 2nd ed. Springer, New York, NY, USA.

Childress, M.J. & Lung, M.A. (2003). Predation risk, gender and the group size effect: Does elk vigilance depend upon the behaviour of conspecifics? *Anim. Behav.*, 66, 389–398.

Cusack, J.J., Kohl, M.T., Metz, M.C., Coulson, T., Stahler, D.R., Smith, D.W., *et al.* (2019). Weak spatiotemporal response of prey to predation risk in a freely interacting system. *J. Anim. Ecol.*, 89, 120–131.

DeByle, N.V. (1985). Wildlife. In: *Aspen: Ecology and management in the western United States. USDA Forest Service general technical report RM-119* (eds. DeByle, N.V. & Winokur, R.P.). Rocky Mountain Forest Range Experiment Station, Ft. Collins, CO, USA, p. 294.

Dobson, A.P. (2014). Yellowstone wolves and the forces that structure natural systems. *PLoS Biol.*, 12, e1002025.

Estes, J.A., Terborgh, J., Brashares, J.S., Power, M.E., Berger, J., Bond, W.J., *et al.* (2011). Trophic downgrading of planet Earth. *Science*, 333, 301–306.

Fidler, F., Chee, Y.E., Wintle, B.C., Burgman, M.A., McCarthy, M.A. & Gordon, A. (2017). Metaresearch for evaluating reproducibility in ecology and evolution. *Bioscience*, 67, 282–289.

Fleming, P.J.S. (2019). They might be right, but Beschta et al. (2018) give no strong evidence that "trophic cascades shape recovery of young aspen in Yellowstone National Park": A fundamental critique of methods. *For. Ecol. Manag.*, 454, 117283.

Ford, A.T. & Goheen, J.R. (2015). Trophic cascades by large carnivores: A case for strong inference and mechanism. *Trends Ecol. Evol.*, 30, 725–735.

Fraser, E.C., Lieffers, V.J., Landhäusser, S.M. & Frey, B.R. (2002). Soil nutrition and temperature as drivers of root suckering in trembling aspen. *Can. J. For. Res.*, 32, 1685–1691.

Halofsky, J.S., Ripple, W.J. & Beschta, R.L. (2008). Recoupling fire and aspen recruitment after wolf reintroduction in Yellowstone National Park, USA. *For. Ecol. Manag.*, 256, 1004–1008.

Hansen, W.D., Romme, W.H., Ba, A. & Turner, M.G. (2016). Shifting ecological filters mediate postfire expansion of seedling aspen (Populus tremuloides) in Yellowstone. *For. Ecol. Manag.*, 362, 218–230.

Hayward, M.W., Edwards, S., Fancourt, B.A., Linnell, J.D.C. & Nilsen, E.B. (2019). Top-down control of ecosystems and the case for rewilding: Does it all add up? In: *Rewilding* (eds. Pettorelli, N., Durant, S.M. & du Toit, J.T.). Cambridge University Press, Cambridge, UK, pp. 325–354.

Houston, D.G. (1982). *The northern Yellowstone elk: Ecology and management*. Macmillan, New York, NY, USA.

Hudson, R.J.H. & Haigh, J.C. (2002). Physical and physiological adaptations. In: *North American elk: Ecology and management* (eds. Toweill, D.E. & Thomas, J.W.). Smithsonian Institution Press, Washington, D.C., USA, pp. 199–257.

Jones, J.R. & DeByle, N.V. (1985). Climates. In: *Aspen: Ecology and management in the western United States. USDA Forest Service general technical report RM-119* (eds. DeByle, N.V. & Winokur, R.P.). Rocky Mountain Forest Range Experiment Station, Ft. Collins, CO, USA, p. 294.

Kauffman, M.J., Brodie, J.F. & Jules, E.S. (2010). Are wolves saving Yellowstone's aspen? A landscape-level test of a behaviorally mediated trophic cascade. *Ecology*, 91, 2742–2755.

Kauffman, M.J., Brodie, J.F. & Jules, E.S. (2013). Are wolves saving Yellowstone's aspen? A landscape-level test of a behaviorally mediated trophic cascade: reply. *Ecology*, 94, 1425–1431.

Kay, C. (1990). Yellowstone's Northern elk herd: a critical evaluation of the "natural regulation" paradigm. Utah State University, Logan, UT, USA.

Keigley, R.B. & Frisina, M.R. (1998). *Browse evaluation by analysis of growth form*. Montana Fish, Wildlife, & Parks, Helena, MT, USA.

Kohl, M.T., Ruth, T.K., Metz, M.C., Stahler, D.R., Smith, D.W., White, P.J., *et al.* (2019). Do prey select for vacant hunting domains to minimize a multi-predator threat? *Ecol. Lett.*, 22, 1724–1733.

Kohl, M.T., Stahler, D.R., Metz, M.C., Forester, J.D., Kauffman, M.J., Varley, N., *et al.* (2018). Diel predator activity drives a dynamic landscape of fear. *Ecol. Monogr.*, 88, 638–652.

Konôpka, B., Pajtík, J. & Shipley, L.A. (2018). Intensity of red deer browsing on young rowans differs between freshly-felled and standing individuals. *For. Ecol. Manag.*, 429, 511–519.

Larsen, E.J. & Ripple, W.J. (2003). Aspen age structure in the northern Yellowstone ecosystem: USA. *For. Ecol. Manag.*, 179, 469–482.

Larsen, E.J. & Ripple, W.J. (2005). Aspen stand conditions on elk winter ranges in the northern Yellowstone ecosystem, USA. *Nat. Areas J.*, 25, 326.

Laundré, J.W., Hernández, L. & Altendorf, K.B. (2001). Wolves, elk, and bison: Reestablishing the "landscape of fear" in Yellowstone National Park, U.S.A. *Can. J. Zool.*, 79, 1401–1409.

Leeper, T.J. (2021). *Interpreting regression results using average marginal effects with R's margins*. Available at: https://CRAN.R-project.org/package=margins. Last accessed 9 June 2021.

Lindroth, R.L. & St. Clair, S.B. (2013). Adaptations of quaking aspen (Populus tremuloides Michx.) for defense against herbivores. *For. Ecol. Manag.*, 299, 14–21.

Macnulty, D.R., Stahler, D.R., Wyman, T., Ruprecht, J., Smith, L.M., Kohl, M.T., *et al.* (2020). Population dynamics of northern Yellowstone elk after wolf reintroduction. In: *Yellowstone wolves: Science and discovery in the world's first national park* (eds. Smith, D.W., Stahler, D.R. & MacNulty, D.R.). University of Chicago Press, Chicago, IL, USA, pp. 184–199.

Maxwell, J.D., Rhodes, A.C. & St. Clair, S.B. (2019). Human altered disturbance patterns and forest succession: Impacts of competition and ungulate herbivory. *Oecologia*, 189, 1061–1070.

Meagher, M. & Houston, D.B. (1999). *Yellowstone and the biology of time: Photographs across a century*. University of Oklahoma Press, Norman, OK, USA.

Metz, M.C., Hebblewhite, M., Smith, D.W., Stahler, D.R., MacNulty, D.R., Tallian, A., *et al.* (2020). What wolves eat and why. In: *Yellowstone wolves: Science and discovery in the world's first national park* (eds. Smith, D.W., Stahler, D.R. & Macnulty, D.R.). University of Chicago Press, Chicago, IL, USA, pp. 156–168.

Monbiot, G. (2014). *How wolves change rivers*. Available at: https://www.youtube.com/watch?v=ysa5OBhXz-Q. Last accessed 15 July 2019.

Morell, V. (2007). Aspens return to Yellowstone, with help from some wolves. *Science*, 317, 438–439.

Motta, R. (2003). Ungulate impact on rowan (Sorbus aucuparia L.) and Norway spruce (Picea abies (L.) Karst.) height structure in mountain forests in the eastern Italian Alps. *For. Ecol. Manag.*, 181, 139–150.

Mueggler, W.F. (1985). Vegetation associations. In: *Aspen: Ecology and management in the western United States. USDA Forest Service general technical report RM-119* (eds. DeByle, N.V. & Winokur, R.P.). Rocky Mountain Forest Range Experiment Station, Ft. Collins, CO, USA, p. 294.

Murie, O. (1951). *The elk of North America*. The Stackpole Co, Harrisburg, PA.

National Research Council. (2002). *Ecological dynamics on Yellowstone's Northern Range*. The National Academies Press, Washington, D.C., USA.

O'Grady, C. (2020). Ecologists push for more reliable research. *Science*, 370, 1260–1261.

Painter, L.E., Beschta, R.L., Larsen, E.J. & Ripple, W.J. (2014). After long-term decline, are aspen recovering in northern Yellowstone? *For. Ecol. Manag.*, 329, 108–117.

Painter, L.E., Beschta, R.L., Larsen, E.J. & Ripple, W.J. (2015). Recovering aspen follow changing elk dynamics in Yellowstone: Evidence of a trophic cascade? *Ecology*, 96, 252–263.

Painter, L.E., Beschta, R.L., Larsen, E.J. & Ripple, W.J. (2018). Aspen recruitment in the Yellowstone region linked to reduced herbivory after large carnivore restoration. *Ecosphere*, 9, e02376.

Painter, L.E. & Tercek, M.T. (2020). Tall willow thickets return to northern Yellowstone. *Ecosphere*, 11, e03115.

Peterson, R.O., Beschta, R.L., Cooper, D.J., Hobbs, N.T., Johnston, D.B., Larsen, E.J., *et al.* (2020). Indirect effects of carnivore restoration on vegetation. In: *Yellowstone wolves: Science and discovery in the world's first national park* (eds. Smith, D.W., Stahler, D.R. & Macnulty, D.R.). University of Chicago Press, Chicago, IL, USA, pp. 205–222.

Peterson, R.O., Vucetich, J.A., Bump, J.M. & Smith, D.W. (2014). Trophic cascades in a multicausal world: Isle Royale and Yellowstone. *Annu. Rev. Ecol. Evol. Syst.*, 45, 325–345.

Piekielek, N.B., Hansen, A.J. & Chang, T. (2015). Using custom scientific workflow software and GIS to inform protected area climate adaptation planning in the Greater Yellowstone Ecosystem. *Ecol. Inform.*, 30, 40–48.

Piovia-Scott, J., Yang, L.H. & Wright, A.N. (2017). Temporal variation in trophic cascades. *Annu. Rev. Ecol. Evol. Syst.*, 48, 281–300.

Renaud, P.C., Verheyden-Tixier, H. & Dumont, B. (2003). Damage to saplings by red deer (Cervus elaphus): Effect of foliage height and structure. *For. Ecol. Manag.*, 181, 31–37.

Rhodes, A.C., Anderson, V. & St Clair, S.B. (2017). Ungulate herbivory alters leaf functional traits and recruitment of regenerating aspen. *Tree Physiol.*, 37, 402–413.

Ripple, W.J. & Beschta, R.L. (2007). Restoring Yellowstone's aspen with wolves. *Biol. Conserv.*, 138, 514–519.

Ripple, W.J. & Beschta, R.L. (2012). Trophic cascades in Yellowstone: The first 15 years after wolf reintroduction. *Biol. Conserv.*, 145, 205–213.

Ripple, W.J., Estes, J.A., Beschta, R.L., Wilmers, C.C., Ritchie, E.G., Hebblewhite, M., *et al.* (2014). Status and ecological effects of the world's largest carnivores. *Science*, 343, 1241484.

Ripple, W.J. & Larsen, E.J. (2000). Historic aspen recruitment, elk, and wolves in northern Yellowstone National Park, USA. *Biol. Conserv.*, 95, 361–370.

Ripple, W.J., Larsen, E.J., Renkin, R.A. & Smith, D.W. (2001). Trophic cascades among wolves, elk and aspen on Yellowstone National Park's northern range. *Biol. Conserv.*, 102, 227–234.

Romme, W.H., Turner, M.G., Wallace, L.L. & Walker, J.S. (1995). Aspen, elk, and fire in northern Yellowstone Park. *Ecology*, 76, 2097–2106.

Rounds, R.C. (1979). Height and species as factors determining browsing of shrubs by wapiti. *J. Appl. Ecol.*, 16, 227.

Ruth, T., Buotte, P. & Hornocker, M. (2019). *Yellowstone cougars: Ecology before and during wolf restoration*. University Press of Colorado, Louisville, CO, USA.

Seager, S.T., Eisenberg, C. & St. Clair, S.B. (2013). Patterns and consequences of ungulate herbivory on aspen in western North America. *For. Ecol. Manag.*, 299, 81–90.

Tallian, A., Smith, D.W., Stahler, D.R., Metz, M.C., Wallen, R.L., Geremia, C., *et al.* (2017). Predator foraging response to a resurgent dangerous prey. *Funct. Ecol.*, 31, 1418–1429.

Vucetich, J.A., Smith, D.W. & Stahler, D.R. (2005). Influence of harvest, climate and wolf predation on Yellowstone elk, 1961-2004. *Oikos*, 111, 259–270.

White, C.A., Olmsted, C.E. & Kay, C.E. (1998). Aspen, elk, and fire in the Rocky Mountain national parks of North America. *Wildl. Soc. Bull.*, 26, 449–462.

Winnie, J. (2014). Predation risk, elk, and aspen: reply. *Ecology*, 95, 2671–2674.

Wolff, J.O. & Van Horn, T. (2003). Vigilance and foraging patterns of American elk during the rut in habitats with and without predators. *Can. J. Zool.*

Yellowstone National Park. (1997). *Yellowstone's northern range: Complexity and change in a wildland ecosystem*. National Park Service, Mammoth Hot Springs, WY, USA.

605 **Table 1**. Peer-reviewed publications showing annual trends in height and (or) browsing of young aspen in Yellowstone National Park
606 linked to the cascading effects of wolves. Listed are the authors and publication year (Authors [Year]), number of the relevant data
607 figure in the article (Fig.), source of data shown in relevant data figure (Data Source), timespan covered by the data (Data Years), type
608 of data collected (Height, Herbivory), and method of data collection (selective sampling of the tallest young aspen [Tallest] or random
609 sampling of all young aspen [Random]). Checkmarks indicate which data were collected and with which sampling method. Shaded
610 cells indicate articles that reproduced or extended data originating in Ripple and Beschta (2007), and dashed-outlined cells indicate
611 articles that reproduced or extended data originating in Painter et al. (2014). Underlined data in Peterson et al. (2014, 2020) were
612 unpublished data from E. Larsen, and are the subject of this article.

				Data Type		Sampling Method	
Authors (Year)	Fig.	Data Source	Data Years	Height	Browsing	Tallest*	Random
Ripple & Beschta (2007)	1c-d	Original data (98 stands)	1998-2006	✓	✓	✓	
Beschta & Ripple (2009)	8b	Ripple & Beschta (2007)	1999-2005	✓	✓	✓	
Ripple & Beschta (2012)	1c-d	Ripple & Bestchta (2007) Original data (resampled 98 stands)	1998-2006 2010	✓	✓	✓	
Beschta & Ripple (2016)	3g	Ripple & Bestchta (2007, 2012)	1998-2006, 2010	✓		✓	
Beschta et al. (2018)	2b-c	Ripple & Bestchta (2007, 2012) Original data (60 additional stands)	1999-2006, 2010 2005-2015	✓	✓	✓	

Table 1 continued

Authors (Year)	Fig.	Data Source	Data Years	Data Type		Sampling Method	
				Height	Browsing	Tallest*	Random
Painter et al. (2014)†	7a	Original data (87 stands)	2003-2012	✓	✓	✓	
Peterson et al. (2014)	1c	Painter et al. (2014) Original data (113 stands)	2003-2012 1999-2013‡	✓		✓	✓
Painter et al. (2015)†	2a	Painter et al. (2014)	2003-2012	✓		✓	
Beschta et al. (2016)	5c	Painter et al. (2014)	2002§-2012		✓	✓	
Painter et al. (2018)†	3a	Ripple & Bestchta (2007, 2012) Painter et al. (2014)	1999-2006, 2010 2005-2012		✓	✓	
Halofsky et al. (2008)	3a-b	Original data (44 stands)	1995-2004	✓	✓	✓	
Peterson et al. (2020)	15.4a	Painter et al. (2014) Original data (113 stands)	2003-2012 2001-2016¶	✓	✓	✓	✓

* Authors sampled the five tallest young aspen within a stand, except for Halofsky et al. (2008) who sampled the three tallest young aspen.

†Included data from a random sample that provided no information on annual trends in height, and little or no information on annual trends in herbivory; the latter was limited to changes in browsing during 1997-1998 and 2011-2012 reported in Painter et al. (2014, 2015).

‡Did not include data for 2000 and 2003.

§Did not specify the data source for 2002.

¶ Did not include data for 2015.

619 **FIGURE LEGENDS**

620 **Figure 1.** Locations of randomly sampled aspen stands in northern Yellowstone National Park.
621 The northern Yellowstone elk winter range is the maximum distribution of the northern
622 Yellowstone elk population during winter when elk often browse young aspen.

623

624 **Figure 2.** Probability densities of browsing (a, c) and height (b, d) of the five tallest young aspen
625 and randomly sampled young aspen in northern Yellowstone National Park during the first and
626 last years of the study (2007, 2017). Low browsing levels and tall heights were more
627 characteristic of the five tallest young aspen throughout the study from 2007 (a, b) to 2017 (c, d).
628 Probability densities for each year of the study are provided in Fig. S2.

629

630 **Figure 3.** Effects of sampling method on estimated annual trends in browsing and height of
631 young aspen in northern Yellowstone National Park, 2007-2017. Relative to a random sample, a
632 nonrandom sample of the five tallest stems estimated a faster annual decrease in browsing (a),
633 faster annual increase in stem height (b), and stronger negative correlation between browsing and
634 stem height (c). Results in (a) and (b) are population-averaged fitted values and associated 95%
635 confidence intervals from best-fit GLMMs of the interactive effect of year and sampling method
636 on browse probability and stem height with year modeled as a continuous (lines) or categorical
637 (points) effect. Results in (c) are the relationships between the categorical fitted values in (a) and
638 (b), with lines estimated from simple linear regressions. No data were collected in 2015.

639

640 **Figure 4**. Effects of stem height on the probability a young aspen stem was browsed (a), and the
641 annual percentage of stems in the sample of young aspen that exceeded the preferred browsing

height of 132 cm (five tallest stems) or 122 cm (random stems) (b) in northern Yellowstone National Park, 2007-2017. Lines in (a) are population-averaged fitted values and associated 95% confidence intervals from best-fit GLMMs estimated separately for five tallest stems (Table S1) and random stems (Table S2). Bars in (b) are percentages of the total annual sample size – pooled across plots – of five tallest stems (N=317-518 stems·year^{-1}) and random stems (N=1027-1748 stems·year^{-1}). No data were collected in 2015.

Figure 5. Effects of sampling method and browse-escape height assumption on stem-level (a, b) and stand-level (c, d) estimates of annual trends in overstory aspen recruitment in northern Yellowstone National Park, 2007-2017. Sampling the five tallest aspen under the assumption that stems taller than 200 cm escaped browsing and joined the overstory estimated rapid annual increases in overstory recruitment (a, c), whereas random sampling all young aspen under the assumption that stems taller than 300 cm escaped browsing and joined the overstory estimated relatively slow annual increases in overstory recruitment (b, d). Results in (a) and (b) are population-averaged fitted values and associated 95% confidence intervals from best-fit GLMMs of the interactive effect of year and sampling method on the probability that a stem exceeded 200 cm (a) or 300 cm (b) with year modeled as a continuous (lines) or categorical (points) effect. Bars in (c) and (d) are percentages of the total annual sample of aspen stands in which the median height of the five tallest stems and randomly sampled stems exceeded a presumed browse-escape height of 200 cm or 300 cm. No stands had a median height of random stems >300 cm in 2007 and no data were collected in 2015.

Figure 1.

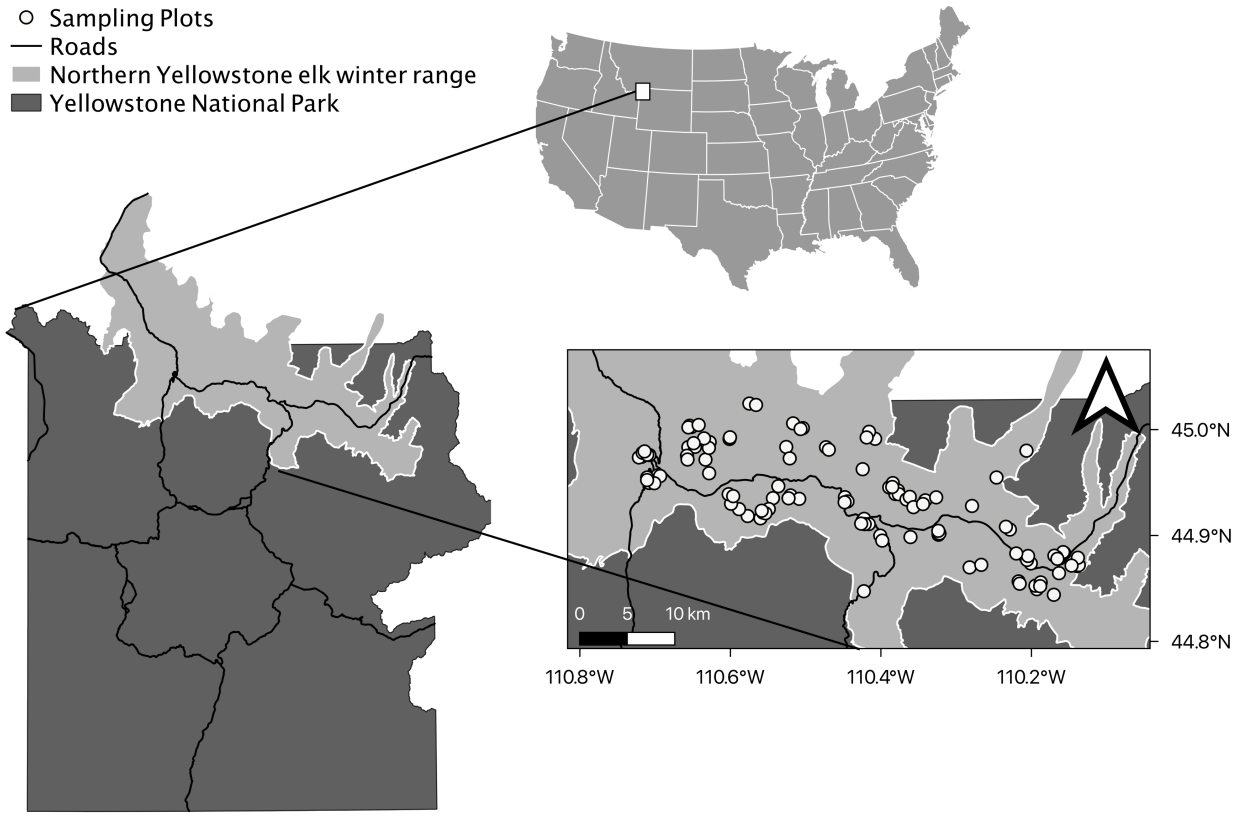

Sampling Plots
Roads
Northern Yellowstone elk winter range
Yellowstone National Park
0 5 10 km
45.0°N
44.9°N
44.8°N
110.8°W
110.6°W
110.4°W
110.2°W

Figure 2.

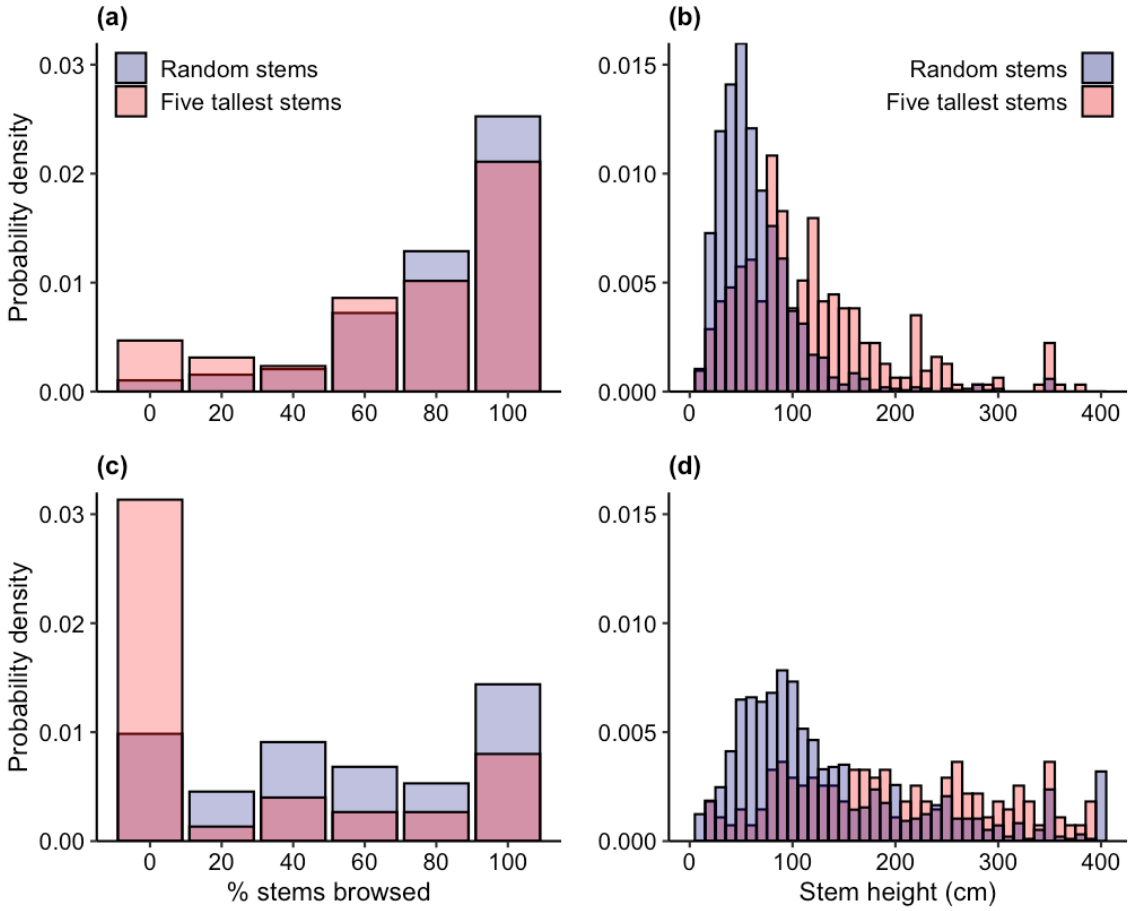

(a)
(b)
(c)
(d)
Random stems
Five tallest stems
Probability density
% stems browsed
Stem height (cm)
0.00
0.01
0.02
0.03
0.000
0.005
0.010
0.015
0
20
40
60
80
100
200
300
400

Figure 3.

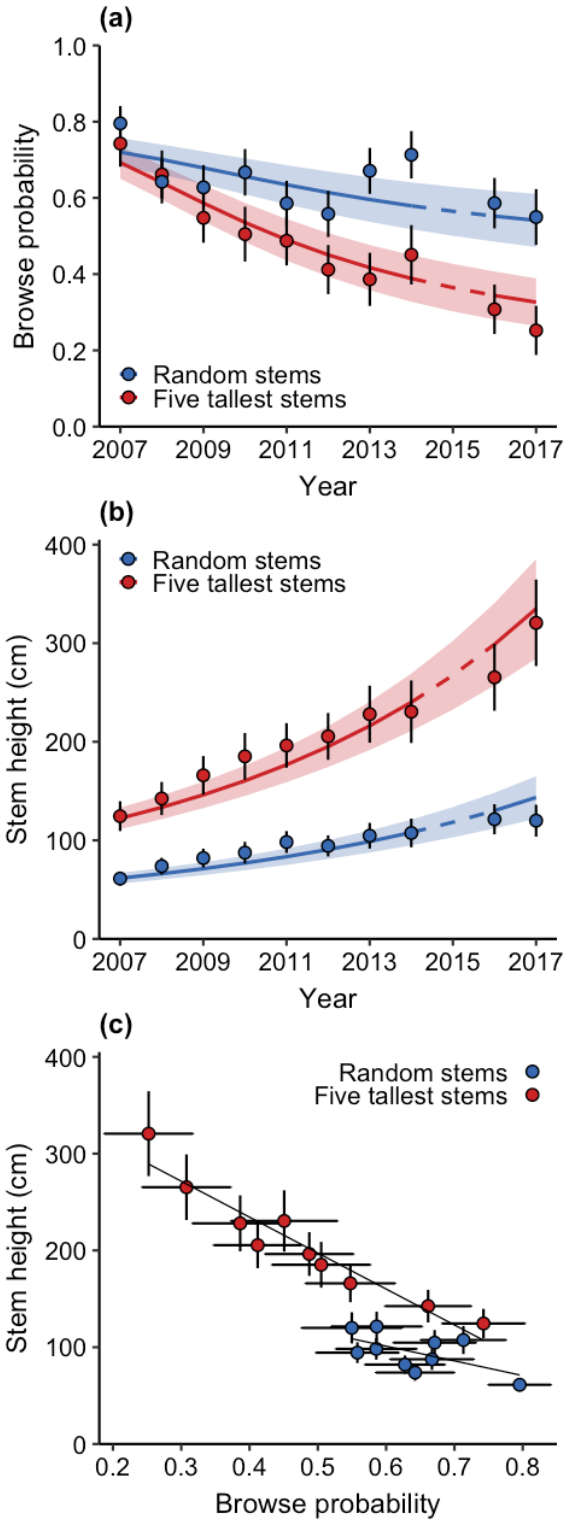

(a)
Browse probability
1.0
0.8
0.6
0.4
0.2
0.0
Random stems
Five tallest stems
2007
2009
2011
2013
2015
2017
Year
(b)
Stem height (cm)
400
300
200
100
0
Random stems
Five tallest stems
2007
2009
2011
2013
2015
2017
Year
(c)
Stem height (cm)
400
300
200
100
0
Random stems
Five tallest stems
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Browse probability

Figure 4.

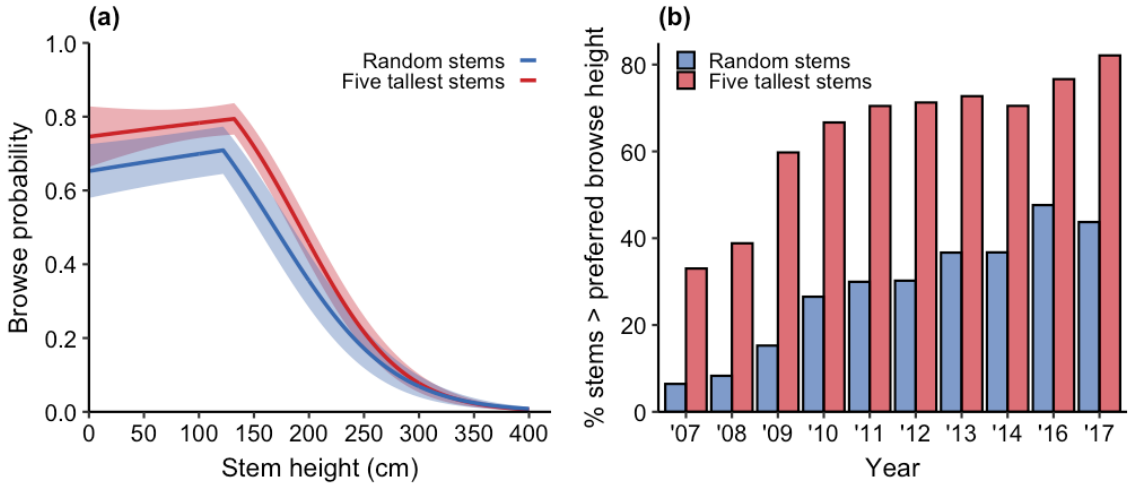
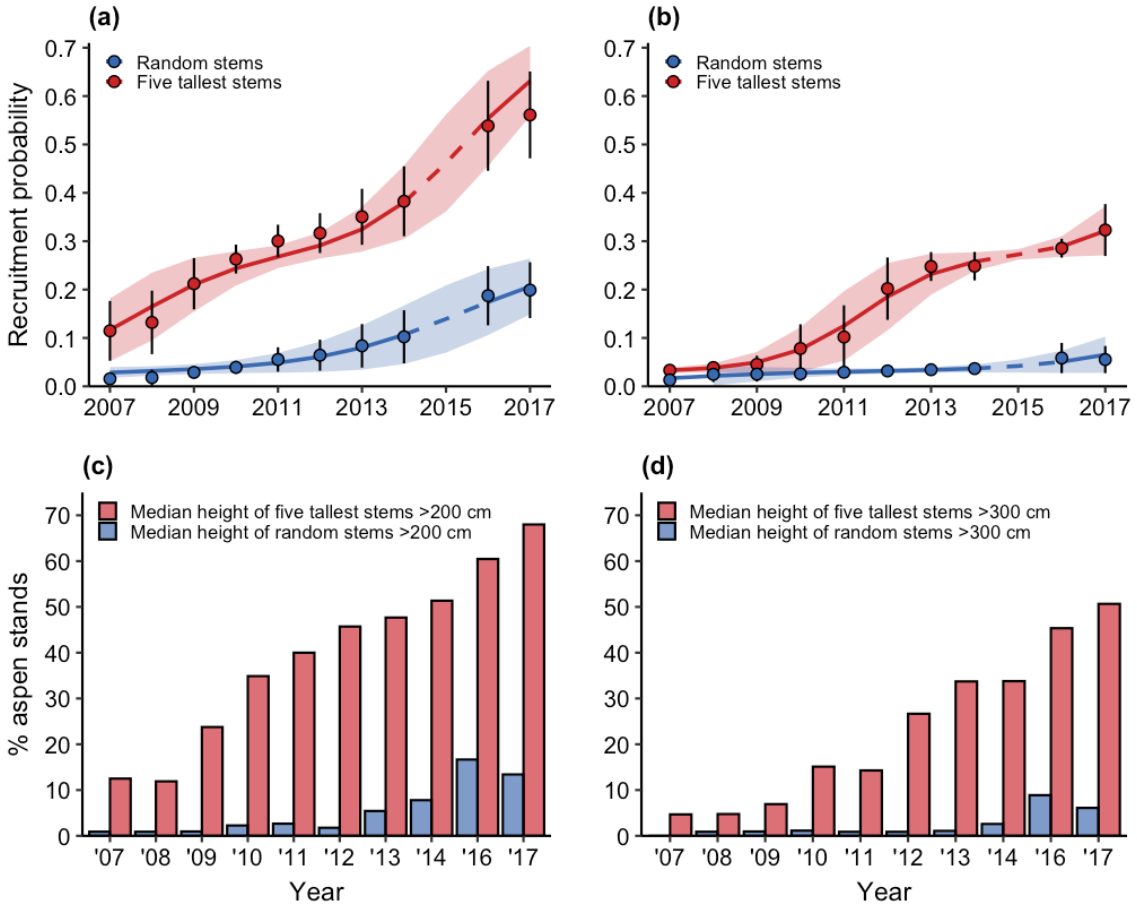

(a)
Random stems
Five tallest stems
Browse probability
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
350
400
Stem height (cm)
(b)
Random stems
Five tallest stems
% stems > preferred browse height
80
60
40
20
0
'07
'08
'09
'10
'11
'12
'13
'14
'16
'17
Year

Figure 5.


(a)
(b)
Random stems
Five tallest stems
Recruitment probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2007
2009
2011
2013
2015
2017
(c)
(d)
Median height of five tallest stems >200 cm
Median height of random stems >200 cm
Median height of five tallest stems >300 cm
Median height of random stems >300 cm
% aspen stands
0
10
20
30
40
50
60
70
'07
'08
'09
'10
'11
'12
'13
'14
'16
'17
Year