Major points :

Conceptually, I have major issues with the RDR score that indicates higher 'degree of realism'. If looking at the effect of CO2 on diatom physiology, I would rather trust small volume experiments where most conditions are controlled, and with the least amount of other species and grazers and which may show a direct effect on either photosynthesis, growth rates or silicification. The larger the volume, the larger the amount of uncontrolled factors that may impact diatom physiology and community shifts. Selective grazer selection of diatoms (from virus, bacteria, parasites to larger grazers) is probably the largest problem. In a large but confined mesocosm, grazer activity is in my opinion the factor that will control diatom community structure rather than CO2. The larger the volume, the larger are the probabilities of getting different organisms in the control and the various treatments that may confound interpretation. The second other confounding parameter would be competition between primary producers. In natural water experiments, the nutrient ratio for instance, together with absolute concentrations may change rapidly and induce indirect changes in community and abundances without any role of CO2. How these confounding factors are deconvoluted in large mesocosms is not clear at all (and in my view impossible). I agree that mesocosms are interesting because they take into account trophic relationships and give an insight into net effects of acidification on the total community, but I don't think their RDR scores should be considered more reliable if you're looking at direct impact of CO2 on diatoms only. What you mention on line 297 « Among these 3, only Pseudo-nitzschia was fairly consistently identified as a "loser" within the investigated natural diatom communities. The relatively weak performance of Pseudo-nitzschia spp. was somewhat surprising because previous monoclonal experiments with this genus often reported a sometimes pronounced positive (Sun et al., 2011; Tatters et al., 2012), or no influence of high CO2 on their growth rate (Sugie and Yoshimura, 2013; Trimborn et al., 2013) but more rarely a negative one (Tatters et al., 2013) » would indeed show that CO2 effects on unique strains show different results than mesocosms experiments... I would place the highest reliability of the effect of OA on diatoms on small volume experiments, which have the least amount of trophic levels, with monospecific experiments being the best, and why not competition experiment between 2 diatoms strains.

This being said, such a literature is much needed and would give valuable insights, but I don't agree at all with the data treatment method presented here, I further think this over-simplification is quite dangerous, and will likely motivate other scientists to run the same kind of "easy" methods to do literature reviews such as this one. This kind of review needs to be done carefully and with a lot of scrutiny. What we need are standardized protocols to look at OA's role on diverse organisms, and push scientists to better conceive their experimental designs, but not to make people think that we can achieve clear conclusions regarding a group as diverse as diatoms from all kinds of mixed experiments, most of which were probably not designed to investigate diatoms specifically.

I would welcome this review of papers, but the data needs to be presented more objectively. It is a good idea to regroup experiments by incubation volumes - small volume (1-4 L), middle volume (5-20 L), minicosms, mesocosms- and presence or not of grazers (through mesh size), but not with this weighing method.

Minor points :

line 24: what exactly do you lean by positively? higher growth rates, abundance, biomass ?

line 53 : since the title is « ocean acidification » I would expect an estimate of the number of marine diatom species here, not total. Sournia et al 91 recognized between 1,400 and 1,800 marine species, more recently the Tara-Oceans metabarcode data revealed up to 4,748 operational taxonomic units (OTU) (Malviya et al 2016), so 30,000 species seem a bit high and probably includes freshwater species ?

line 55 : you may write <2 µm, (e.g. *Minutocellus* and *Minidiscus* genera)

line 199/203 : I don't understand why you would assume spherical shape for large volume incubators, most are cylinder shaped, with a cone at the base or not depending on the model. I have not yet seen one spherical large volume incubator.

line 263 : please correct « ,. »

line 263 : I believe bibliographical references are not correctly annotated in the text, example : « microscopy except for (Endo et al., 2015) who used molecular tools »

Table 1 : I find the text too small to read, in particular if printed.