Testing ordering interventions: Discussion

We found that listing the products in an online supermarket in order of sustainability did not have an effect on the proportion of sustainable choices, either when the ordering was covert or when it was accompanied by a statement informing participants about the product ordering. Participants chose a more sustainable product 59.5% of the time in the random-ordering arm, 58.8% of the time in the covert ordering arm, and 59.6% of the time in the overt-ordering arm. Nor did we find any effect of the interventions in sensitivity analyses on secondary measures, which included the selection of the most sustainable product in a category and the sustainability rank of chosen products. This may be because neither of the two mechanisms by which we surmised our interventions would work were operative. We expected that the covert ordering would work because there would be an effect of position, with participants choosing products that were higher in the lists because they were more reachable and salient. However, additional analyses suggested that, in the random-ordering arm, there was no effect of the position of a product on the product category page. We expected that overt ordering would operate via conscious decision-making processes. However, in the overt-ordering arm, only 19.5% of participants correctly identified that the products were ordered by sustainability in a post-task questionnaire, so it seems the majority did not notice the statement telling them about the ordering.

It is surprising that our covert ordering intervention did not have an effect, given results from other studies that carried out similar interventions. Our study was powered to detect a 5% difference (when there is relatively low variability of individual and product category random effects), which is the magnitude of the difference found by Koutoukidis et al (2019). Our study does differ from many of the few existing studies, which were based on behaviour in bricks-and-mortar environment using hard-copy menus when healthier items were placed at the top of lists (Dayan & Bar-Hillel, 2011; Mueller et al., 2020), including the one that ordered products by environmental impact (Langen et al., 2022). There are also differences in study design, for example we used a randomised controlled trial, whereas Schmidtke et al. (2019) used a pre-post design in actual kiosks. In comparison, participants in our study made product choices in a simplified online supermarket environment following a structured shopping task. This setting might make it easier for participants to explore and pay attention to all product options, compared to a noisier off-line environment or an online environment with more choices. However, it should also be noted that our modelling approach minimised Type-I errors and studies that used traditional approaches such as the Analysis of Variance have a higher risk of finding spurious effects (Jaeger, 2008). We used logistic mixed-effects models, which included separate error terms for participant and product category, allowing us to incorporate additional uncertainty in the estimates of intervention effects associated with variation between participants and categories. Models that do not take this variation into account are likely to underestimate the standard errors of coefficients, potentially leading to an overstatement of statistical significance.

Another possibility is that ordering interventions only have an effect in online interventions when there is the potential for a lot of scrolling, so people do not make it to the bottom of the list. In our experiment, the number of products that were shown on the screen without scrolling varied by device model and screen size, but it was designed so that four products were shown without scrolling on most devices and screens. As mentioned in Footnote 5, we conjecture ordering to have an effect based on position, which is different from scrolling, which relates more to the “above-the-fold” effect. If scrolling modified the effects, we would expect products displayed at the bottom two positions to be selected less as participants would need to scroll down to see them, however, there is no evidence of this in our data (see Table 3 and Figure 9). We cannot rule out

the possibility that if there were more products and more scrolling was required to see products at the bottom of the list, then products further down the list might be chosen less. However, in that case, it is questionable whether ordering is still the relevant mechanism (when you think of ordering as being analogous to placing products at the top of the list on a physical menu). Instead, you might think it is more similar to interventions that decrease the availability of certain products or increase their costs, by making them harder to find.

As far as we know, ours is the first study to investigate ordering effects for sustainable products in online environments; the closest comparator study aims to promote the choice of healthier products. Koutoukidis et al. (2019) also used a simulated online supermarket environment. The task was slightly different, as participants were given a 10-item shopping list and could browse categories rather than going through a forced-journey; and it was entirely hypothetical, participants did not receive the products they chose. The primary outcome measure was the saturated fat content of the whole basket, which decreased by 5%. One of the secondary outcome measures showed that there was a 10% decrease in the percentage of products with less than 1.5% saturated fat per 100g in the basket, i.e. products that can be labelled as 'low' in saturated fat content according to Department of Health guidance (Department of Health and Social Care, 2016). So on average participants given an ordered list put more products that were low saturated fat in their baskets. Some products that are low in saturated fact are obviously labelled as low fat, for instance semi-skimmed milk or lighter butter (both of which were used in the experiment). Open-ended comments left at the end of the experiment suggested that participants wanted to buy healthier food and that they noticed that the healthier products were at the top of the ordering. This is potentially quite different from sustainable products, where it is not always obvious which product is more sustainable, especially within product categories, and where consumers often do not know what choices will reduce their carbon footprints (Kause et al., 2019). Further, when making food choices, health and nutrition are more important to consumers than sustainability (Fox et al., 2021; Ghvanidze et al., 2017; Grunert et al., 2014). So it is possible that re-ordering according to nutrition content is noticed by consumers and supports their reflective decision-making, in a manner that re-ordering according to sustainability did not.

Even in our overt ordering intervention, where we had a statement at the top of the product category telling participants about the ordering (see Figure 4 and Figure 6) , participants did not notice that products were listed in order of sustainability. Other researchers have also reported that many people did not notice their disclosure statements (see Wachner et al.(2020) for an online study and Kroese et al.(2016) for a field study). We had thought that putting the statement in a box at the top of the page would be salient, but participants may have focused on the product list itself. Our interface was simplified compared to an actual online supermarket, so in real-life shopping people might be even less likely to pay attention to information about sustainability rankings of products. However, in future studies, if time and budget allow, it would be good to pre-test disclosure messages to determine how best to display them.

Future research could also investigate how to make an environmental ordering more salient. One possibility would be to use pop-ups, which have been successful in prompting people to make healthier swaps (Bunten et al., 2021; Huang et al., 2006; Jansen et al., 2021; Koutoukidis et al., 2019); whereas see Forwood et al. (2015) for a swap experiment in simulated supermarket that was not successful. However, it should be noted that a trial using pop-ups in a real online supermarket, where students were given money to place an order and actually received their products, pop-ups for healthy choices did not increase the proportion of healthy purchases despite being powered to detect a difference of 1% (Stuber et al., 2022). More field research in real supermarket environments is required to establish the external validity of the effects of pop-ups on behaviour, but they at least seem to be noticed in simulated environments.

The trial in this study was designed to simulate an online supermarket environment in real life, and to be as close as possible to a real online supermarket. The experimental interface was designed to mimic the layout of existing major online supermarkets and all the products with their

names, prices and pictures were taken from existing online supermarkets as well. A large sample from England, Wales and Northern Ireland representative of age, gender and ethnicity was recruited to complete the experiment, with real material incentives in terms of getting the chosen products and strict quality checks to ensure that final sample excluded participants who were likely to have not participated fully. Therefore, the absence of position effects and the large variation in preferences for different products should reflect to some extent consumer behaviour in real-life online grocery shopping environment. Even for products in a relatively narrow product category that were chosen to be comparable to each other, characteristics other than position and price play important roles in determining the purchasing decisions. Research has pointed out the importance of habits and tastes in food choices (Fox et al., 2021; Osman & Jenkins, 2021; Riet et al., 2011), and our results show that such habits and tastes might be too ingrained to be changed by subtle modification of situational cues.

We do not have data on popularity of the chosen products (the supermarket whose products we used did not provide that information), although the results did seem to suggest that consumer choices were driven strongly by preferences for specific products which is probably taste based. For example, in the soup category, the third product is among the most popular ones despite being expensive, while the fifth product which has the same brand as the third product is not popular. Selection of products in this experiment was subject to multiple restrictions mentioned in Section 2.3. In future studies, if time and budget allow, it would be desirable to run pre-tests to make sure the products are of closer popularity.

There are limitations to this study. Given the reported variance estimates in Table 2, our trial was able to detect an effect of 5%. It is possible that the effects of ordering interventions are smaller than 5% so cannot be detected by the current sample size. However, the point-based estimates of our optimal model did not suggest there are such effects. Secondly, it is possible that our results are specific to the products chosen, the product categories chosen, the number of products on each page, the measure of sustainability, and other details in terms of experimental design. We did try to choose products that were comparable to each other in terms of popularity and price, while still having variance in their carbon footprint.

Nevertheless, we cannot rule out the possibility that the findings are specific to the experimental design and more research using variations of the design is welcomed to generate more evidence on the effects of such interventions. Finally, despite our effort to simulate the online grocery shopping experience with high ecological validity and recruit a sample that is representative, the results still come from an online experiment completed by panellists, which potentially threatens the ability to generalise from our results to the real-life situation we are studying (external validity) and which could be better dealt with using a field trial.

It is worth pointing out that, as discussed in Section 1.2, the trial simplified many of the real-life considerations. In a real online supermarket, consumers can go straight to the lists of their favourite products or products they have bought before and choose from there, without being exposed at all to other products that might be more sustainable (Bunten et al., 2022). Consumers can also sort the products by price, popularity, or other factors that they care more about, which is likely to overwrite any default ordering that is intended to nudge consumer choices. Because our online shopping task did not have either of those features, we increased the chance that we would find an ordering effect compared to a real-life environment, and we still did not find an effect.

Many interventions around changing the choice architecture have been shown to be effective in encourage healthier and more sustainable diets. However, the particular interventions investigated in this study – ordering food products in a simulated online supermarket by their environmental impact covertly/overtly – were not found to generate the expected increase in consumers' likelihood of choosing more sustainable products. Our results suggest there are ingrained preferences for different grocery products that largely determine purchasing choices, and the difficulty of conveying information effectively to consumers in the online grocery shopping

environment. To our knowledge, this study is the first to explore the effects of ordering interventions by environmental impact in an online shopping environment. More research needs to be done to continue to fill the evidence gap, provide more solid answers to the research questions on the effectiveness of such interventions, and enrich our understanding of when changing choice architecture works.