STATISTICAL ANALYSIS

The data collected in the laboratory (e.g., foods/drink intake, body weight, etc.) will be recorded manually in duplicate (study log) and then inputted into relevant spreadsheet with independent verification on a different day (add column "checked" to confirm). Software-derived data (e.g., liking, appetite, etc.) from Qualtrics will be sense checked for expected counts; histogram (as appropriate) for visual scanning of all continuous data prior to any analyses. Step-by-step documentation (from raw to edited files for analysis; recording see study checklist) will be saved in study folder and cross-checked on a separate occasion before next steps. Statistical analysis will be conducted blinded (and unblinded post final analyses verified by statistician); all outputs will be saved for future reference. The data will be independently checked by two further researchers.

The analysis will be conducted in R adhering to a pre-agreed protocol with significance defined as $p < 0.05$ (primary) and $p < 0.01$ (secondary/additional: 0.05 divided by six sets of outcomes). Descriptive statistics (mean, standard deviation, counts, percentages) will be utilised to report participant characteristics (e.g., sex, age, BMI, ethnicity, education, employment status, habitual intake, etc.).

Primary Analysis

Energy Intake

Data: daily energy intake (kcal/d; calculated from breakfast, lunch, dinner, snacks) and amount consumed (grams; breakfast, lunch, dinner, snacks) from four-day testing period for each diet (UPF, unmatched LPF and matched LPF). For descriptive purposes we will also compare energy intake by individual meal type (breakfast, lunch, dinner, snacks).

Test: repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF) and four-days (Monday, Tuesday, Wednesday and Thursday) design. If a significant main effect of condition is present, then follow-up t-tests between trial arms will be conducted.

Significance: $p < 0.05$.

Secondary Analysis

Appetite

Data: appetite ratings (hunger, fullness, thirst from breakfast, lunch, dinner) from four-day testing period for each diet (UPF, unmatched LPF and matched LPF).

Test: repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF) and four-days (Monday, Tuesday, Wednesday and Thursday) design using AUC (calculated for hunger, fullness and thirst). If a significant main effect of condition is present, then follow-up t-tests between trial arms will be conducted.

Significance: $p < 0.01$.

Additional Analysis

Eating Rate

Data: meal eating rate (g/min; breakfast, lunch, dinner) and energy intake rate (kcal/min; breakfast, lunch, dinner) from four-day testing period for each diet (UPF, unmatched LPF and matched LPF).

Test: repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF) and four-days (Monday, Tuesday, Wednesday and Thursday) design. If a significant main effect of condition is present, then follow-up t-tests between trial arms will be conducted.

Liking + Familiarity

Data: liking and familiarity ratings (breakfast, lunch, dinner, snacks) from four-day testing period for each diet (UPF, unmatched LPF and matched LPF).

Test: repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF) and four-days (Monday, Tuesday, Wednesday and Thursday) design. If a significant main effect of condition is present, then follow-up t-tests between trial arms will be conducted.

Food Cravings + Well-Being + Gut Health + Sleep

Data: food cravings, well-being, gut health and sleep from four-day testing period for each diet (UPF, unmatched LPF and matched LPF).

Test: one-way repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF). If a significant main effect is present, then follow-up t-tests between trial arms will be conducted.

Physical Activity

Data: physical activity expenditure from four-day testing for each diet (individual-level; UPF, unmatched LPF and matched LPF).

Test: one-way repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF). If a significant main effect is present, then follow-up t-tests between trial arms will be conducted.

Body Composition

Data: body weight (kg) from four-day testing for each diet (individual-level; UPF, unmatched LPF and matched LPF).

Test: repeated measures (within-subject) ANOVA with three diet arms (UPF, unmatched LPF and matched LPF) and pre-post testing period (Monday AM and Thursday PM) design. If a significant main effect is present, then follow-up t-tests between trial arms will be conducted.

Exploratory Analysis

Relationship Between Participant-Level Energy Intake + Body Weight Change

Data: body weight (kg) (pre-post) vs average energy intake (kcal/d) for each diet (individual-level; UPF, unmatched LPF and matched LPF).

Test: multi-level modelling by using (i) change in body weight (pre vs post per diet); and (ii) person-mean centred energy intake by subtracting "mean energy intake across three diets" (UPF, unmatched LPF and matched LPF) from "energy intake within each diet" to isolate changes in energy intake across arms. This will be conducted using the 'lme4' package in R, with a random intercept of participant.

Missing Data

The primary outcome (energy intake) is generally not expected to have any missing data as we will only include participants that complete all three testing periods. However, if there is a small amount of missing data (e.g., missing record for snack box intake on one day, this will then be imputed using the 'mice' package in R (van Ginkel and Kroonenberg, 2014). This will use five imputations with predictive mean matching, and include all relevant data in the imputation. In addition, we will consider using a multi-level modelling approach if we identify significant missing data (e.g., >20%) to more appropriately address this.

Data Exclusion

Data will be excluded from analyses if non-compliance with study protocol (e.g., non-study foods exceeding 350kcal/d; Forde et al. (2026)) or if considered implausible (24-hr recall) (female: 500-3500 kcal/d and male: 800-4000 kcal/d; Willett, 2013) – this will be imputed (as described above).

Sensitivity Analysis

To understand if the results are the same based on excluding: (i) aim guessers (impact of food processing or nutritional composition on energy intake); (ii) manipulation check (identified perceived differences in processing between diets and/or correct order from ranking task); (iii) outliers (if outside threshold: 2.5 median absolute deviation from group mean) (Leys et al 2013); (iv) influential cases (data points above 1.0 based on Cook's distance) (Stevens, 1984); and (v) we will also examine results when any energy intake from non-study foods are included to calculate daily energy intake. For (i) to (v); any deviations in results from the main analyses will be reported. In addition, if we find any evidence of condition differences in the secondary and/or additional analyses then we will examine if results remain the same when they are controlled for (e.g., inclusion as covariates). The primary outcome will be reviewed for diet sequence effects by using a six-level (ABC, ACB, BAC, BCA, CAB, CBA) between-

subjects factor to understand if patterns of results are influenced by presentation order. Data will also be visually evaluated (due to lack of power to detect an effect) to cross-check differences (energy intake vs diet sequence) from main analyses.

Interim Analysis

Interim analysis approaches guide recruitment targets by balancing pre-study uncertainty over expected effect sizes and research waste by analysing the primary outcome at a pre-defined interim timepoint (in this case when N = 18 participants have completed all three conditions) (Lakens, 2014; Ciolino et al., 2023). At interim, we will use Bayes Factors to determine if there is conclusive evidence based on conventional cut-offs to inform if additional data collection would be valuable (e.g., full sample size of up to n = 30) (Rounder et al., 2012; Dienes, 2014; Dienes, 2016).

a) UPF vs. unmatched LPF diet comparison: For informed priors, based on the difference in energy density of the diets in the present study and accounting for a small degree of dietary compensation, a ~480kcal mean difference between the two conditions could be expected and an SD of the difference of ~300kcals, somewhat narrower than previous estimates (Hall et al, 2019) but accounting for the different diets offered. As such our priors will be = 480 SD = 300 with a normal distribution for the alternative hypothesis, and 0 with SD = 100 for the null. Bayes Factors (BF^{10}) will be computed using the 'brms' package, with marginal likelihoods estimated via 'bridgesampling'. We will use a stopping rule of $BF^{10} > 10$ to indicate strong evidence in favour of the null.

b) UPF vs. matched LPF diet comparison. Based on matching of energy density between diet, for the null model (i.e., no difference), we will specify a normal prior with a mean of 0 kcal and a standard deviation of 100 kcal. This prior suggests the mostly likely effect is a 0-kcal difference, with small effects plausible and larger differences becoming progressively less plausible. As the null model is based on not anticipated difference, as such we will examine the BF01 (evidence in favour of the null hypothesis). The alternative prior will be similar to that above (mean = 480, SD = 300). As above, we will use a stopping rule of $BF^{01} > 10$ to indicate strong evidence in favour of the null. Given the uncertainty around both priors we will also examine default/uninformed priors using 'ttestBF' function.

If there is strong evidence for both a) and b) based on the above cut-offs, we will end data collection and progress to the pre-registered analysis approach (described above).

If there is not convincing evidence based on the above, we will conduct inferential statistics to calculate effect sizes for the above comparisons and simulate the effect that increasing sample size to N=24 and N=30 would have on evidential certainty. If these simulations suggest that increasing sample size to N=24 or N=30 would meaningfully increase evidential certainty (e.g., BF>10), we will continue to the desired sample size. If simulations do not suggest that further data collection up to N=30 would meaningfully increase evidential certainty (e.g., minor change in BF), we will end data collection and progress to the pre-registered analysis approach. Bayes factors will also be reported for descriptive purposes for the final analyses.