The Free-Linking Task: A graph-inspired method for generating non-disjoint similarity data with food products

Jacob Lahne
Katherine Phetxumphou
Marino Tejedor-Romero
David Orden

Keywords
sensometrics, free sorting, free linking, DISTATIS, graph theory, network theory, rapid methods

Abstract
"Free sorting", in which subjects are asked to sort a set of items into groups of "most similar" items, is increasingly popular as a technique for profiling sets of foods. However, free sorting implies an unrealistic model of sample similarity: that similarity is purely binary (is/is not similar) and that similarity is fully transitive (similarities {A, B} and {B, C} imply {A, C}). This paper proposes a new method of rapid similarity testing—the "free-linking" task—that solves both problems: in free linking, subjects draw a *similarity graph* in which they connect pairs of samples with a line if they are similar, according to the subject's individual criteria. This simple task provides a more realistic model of similarity which allows degrees of similarity through the *graph distance* metric and does not require transitive similarity. In two pilot studies with spice blends (10 samples, 58 subjects) and chocolate bars (10 samples, 63 subjects), free linking and free sorting are evaluated and compared using DISTATIS, RVb, and the graph parameters *degree*, *transitivity*, and *connectivity*; subjects also indicated their preferences and ease-of-use for the tasks. In both studies, the first two dimensions of the DISTATIS consensus were highly comparable across tasks; however, free linking provided more discrimination in dimensions three and four. RVb stability was equivalent for the two methods. Graph statistics indicated that free linking had greater discrimination power: on average subjects made similarity groupings with lower degree, lower transitivity, and higher connectivity for free linking in both studies. However, subjects did overall find free sorting easier and liked it more, indicating a higher cognitive difficulty of free linking. The free-linking task, therefore, provides more robust, realistic similarity maps at the cost of higher panelist effort, and should prove a valuable alternative for rapid sensory assessment of product sets.

1. Introduction

Methods for rapidly identifying similarities and differences in sets of food products have become increasingly popular in sensory evaluation (Delarue, 2015; Valentin et al., 2012; Varela & Ares, 2014). In particular, "free sorting", in which subjects are asked to sort a set of items (in this case, foods or beverages) into groups of "most similar" items is increasingly popular as a technique for profiling sets of foods (e.g., Lahne et al., 2018). Free sorting presents several advantages: it does not require that subjects be trained, it is sensitive and stable with relatively low numbers of subjects (usually as low as 25 subjects), it can accommodate relatively high numbers of samples (as many as 20), and it has been shown to give product "maps" or "configurations" (through multivariate analyses) that bear a close resemblance to those from traditional and more work-intensive methods like Descriptive Analysis. Furthermore, unlike other rapid methods like Projective Mapping or Flash Profiling (Dehlholm et al., 2012), free sorting only requires that subjects make simple, holistic decisions of similarity or difference, rather than requiring a scaled degree of difference that may induce a higher cognitive load.

However, a key disadvantage of free sorting is that the task of sorting samples makes some strong assumptions about the underlying similarities between the products that are being modeled. Groups in free sorting are *disjoint*, meaning that no element can belong to two groups. Given samples A, B, and C there is no way that the same subject can create two similarity sets such as {A, B} and {B, C} without creating a superset {A, B, C}. This simplifies the task for the subjects and reduces the time and amount of samples required (because retasting is minimized), but this restriction has two potentially undesirable consequences. The first is that the same subject cannot represent different *types* or *dimensions* of similarity in the same sort: it is easily conceivable that A and B are similar in terms of one attribute, say, "sweetness", while and A and C are similar in terms of another, say "appearance". It is quite easy to imagine real-world situations in which this occurs. The second consequence is that similarity is necessarily modeled as fully transitive: if A is similar to B, and B is similar to C, then A must be similar to C, and furthermore the data can only indicate that all three samples *are equally similar*. This is also clearly contrary to easily imagined real circumstances: perhaps A, B, and C are all "sweet", but while A and B are equally sweet, C is only half as sweet. Should a single subject be required to group these together?

Two closely related alternatives have been suggested for the simple free-sorting task that address these issues: free *multiple*-sorts (Blanchard & Banerji, 2016; Dehlholm, 2015; Dehlholm et al., 2012) and *hierarchical* free-sorts (Koenig et al., 2020, 2021). The former modification asks subjects, after they have completed a simple free-sorting task, to repeat the task until they feel they have exhausted all possible grouping configurations (Dehlholm, 2015); the latter asks subjects, once they completed a simple free-sorting task, to continue making groups *of groups* until they cannot proceed further (Koenig et al., 2021). Thus, free multiple-sorting solves the first problem highlighted above, and hierarchical free-sorting solves the second problem. However, neither approach solves *both* problems, and they both introduce problems of panelist motivation, in that they require a much more extensive data-collection procedure that will be discouraging for some subjects. This is a more major problem when a large number of samples is used, as in Koenig et al. (2020), but difficulty and motivation problems are reported with as few as 18 complex samples sorted by taste (Kessinger et al., 2020). In addition, the data collection for both methods is much more complicated and more poorly supported in practical data-management programs (based on the authors' personal communications with major sensory and survey software providers in pursuit of these methods), which appears to have limited the adoption of either approach in academia and industry in favor of the simple free-sorting task. For example, Spencer et al. (2016) had to write custom software to support hierarchical free-sorting, and authors as recent as Koenig et al. (2020, 2021) have used paper ballots because of the lack of software supporting hierarchical free-sorting, requiring extensive transcription of results.

Therefore, in this manuscript we propose an alternative task to the free-sorting task, inspired by graph theory (Gross et al., 2014), which we term the "free-linking" task. In the free-linking task, subjects are given a set of samples just as in free sorting, but rather than forming disjoint groups, subjects are asked to indicate, for each pair of samples, whether the samples are similar. This connect-the-dots interface was implemented in the SensoGraph system (Orden et al., 2019, Alcalá, ES) in order to support this task, in which subjects are asked to draw "links" between samples if they are similar (Figure 1). However, a paper-based system for free linking would be no harder to implement than a paper-based simple free-sorting task.

FIGURE 1 GOES HERE

While the free-linking task solicits binary similarity data on a pair-wise basis for samples—two samples are either similar or they are not—it does not impose the disjoint, restrictive model of similarity implied by free sorting. Given 3 samples A, B, and C it is possible for a subject to indicate, pairwise, that there are similar pairs {A, B} and {B, C} without indicating that A and C are directly similar. Put another way, the free-linking task asks each subject to draw their own similarity graph for the samples (Lahne, 2020; Orden et al., 2019, 2021). Unlike previous graph-based approaches to similarity in food products, where just the presence or absence of a connection was considered, in free linking we make use of the *graph distance* between samples as a basis for a dissimilarity matrix for further analysis (Chartrand & Zhang, 2014). In the example above, *distance*(A, B) = *distance*(B, C) = 1, while *distance*(A, C) = 2. This allows the analyst to *infer* from a single subject's data that, for the example above, there might be some shared similarity between A and C without the link {A, C} actually being drawn. This same change also addresses the second problem with simple free-sorting: subjects can now indicate pairwise whether samples are similar, but because there are not larger similarity groups (e.g., {A, B, C} in free sorting) it is not required that all samples that are connected be similar *in the same way*. This allows more flexibility for a subject's holistic similarity judgments (Figure 2; see also Figure 3 for details on the dissimilarity).

FIGURE 2 GOES HERE

The free-linking task can be analyzed by the same tools that exist for the free-sorting task: dimensionality reduction (through MDS, DISTATIS, and other approaches) and graph-based approaches like Sorting Backbone Analysis. This allows analysts used to free sorting to easily employ free linking, and for direct comparison of results.

Therefore, it is reasonable to hope that the free-linking task will provide results that are comparable to free-sorting in terms of ease of deployment and data collection, but might allow for more realistic and detailed results. In particular, the lack of forced memberships to a group should allow for easier distinction among similar but not identical samples—that is, a more multidimensional structure of similarity and difference. In order to investigate the utility of the free-linking task, we report the results of two pilot studies in which subjects used both free sorting and free linking to report their perceptions of different food products. In both pilot studies subjects completed both free-sorting and free-linking tasks for the same samples in a counterbalanced order. In the first study, subjects evaluated 10 blends of 4 dried spices (cinnamon, turmeric, pepper, and cardamom) for similarity by aroma. In the second study, subjects evaluated 10 commercial chocolate samples for similarity by taste. We hypothesized that the overall similarity configuration should be similar between the two methods, and that the results of the two methods should be equally stable, but that the free-linking results would provide more realistic, multidimensional models of similarity, which should be evident in parameters for the graphs derived from the similarity measurements as well as in visualizations from DISTATIS.

2. Materials and Methods

The two studies reported were very similar in most details besides sample type, and so the basic information distinguishing the studies is given below, followed by details on methodology and analysis that were the same for both studies.

2.1. Study 1–Spice sorting

Study 1 was conducted in November and December of 2019, and used spices and spice blends as stimuli. Sample details are given in Table 1. All spices were purchased at Kroger (Blacksburg, VA, see Table 1). Samples were presented to subjects in foil-wrapped glass vials in order to avoid visual discrimination, and evaluation was entirely orthonasal.

A total of N = 58 subjects (38 female, 20 male, average age 29 years old) participated in Study 1. Subjects were recruited from the Virginia Tech/Blacksburg community. Subjects were not trained sensory panelists (e.g., for Descriptive Analysis), but some had participated in previous untrained sensory tests at Virginia Tech. Subjects received no compensation, but were given snacks after completing Study 1.

2.2. Study 2–Chocolate sorting

Study 2 was conducted in November of 2020, and used commercial chocolate bars as stimuli. Sample details are given in Table 1. All chocolate bars were purchased at Kroger (Blacksburg, VA, see Table 1). Samples were presented in souffle cups with the bars' identifying details (e.g., logos) effaced, in natural light, and evaluation was by taste and retronasal flavor.

A total of N = 63 subjects (49 female, 14 male, average age 34 years old) participated in Study 2. Subjects were recruited from the Virginia Tech/Blacksburg community. Subjects were not trained sensory panelists (e.g., for Descriptive Analysis), but some had participated in previous untrained sensory tests at Virginia Tech, including some who had participated in Study 1. Subjects received no compensation, but were given snacks after completing Study 2.

TABLE 1 GOES HERE

2.3. Overall study design

Both studies used the same overall design. Subjects were recruited to participate in free-linking and free-sorting of the same samples. In order to obtain within-subjects data, subjects were randomly assigned one of the two tasks first, then took a short break, then completed the other of the two tasks, then completed a short survey that asked them about their perceptions of the tasks and some basic demographic details. In both free-sorting and free-linking studies, subjects were seated at tables with a 36" x 36" workspace available and allowed to organize their samples spatially prior to entering their judgments into the data-collection software.

2.4. Free-sorting task

In the free-sorting task, subjects received all 10 samples at the same time in a randomized order. Sorting data was collected using the Compusense Cloud (Guelph, ON) system. Subjects were prompted to "sort into groups based on similarities". They were informed that there was no right answer, and told that they could make any number of groups between two (2) and nine (9), with as many samples as they chose in each group.

2.5. Free-linking task

In the free-linking task, subjects received all 10 samples at the same time in a randomized order, positioned as the vertices of a regular polygon (see Figures 1 and 2). Linking data was collected using the SensoGraph (Orden et al., 2019, Alcalá, ES) system. Subjects were prompted to "join with a line those pairs of products you consider similar, dragging from one to the other with the finger or the mouse" (see Figure 1). The codes presented on the screen for the SensoGraph interface were given in random order for each subject. Subjects were able to remove lines they had previously made (in case of mistakes or revisions in judgment) before submitting their answers.

2.6. Data Analysis

Results from both free sorting and free linking were analyzed in parallel in order to compare the results of the method. This parallelism is enabled by the data structure provided by both methods: the dataset for each analysis is an $N \times K \times K$ array of (dis)similarity matrices, where N is the number of subjects and K is the number of samples. In free sorting, each $K \times K$ slice is composed by cell entries a_{ij} which are binary (either 0 or 1), representing whether, for the current subject, samples i, j were sorted together. The raw data is a *similarity* measure in which a 1 indicates similarity through group membership, and the dissimilarity matrix, which is obtained by subtracting every entry from 1, can be treated as binary distance and is analyzed via MDS or DISTATIS (Abdi et al., 2007). In free linking, the graph drawn by the current subject provides a graph distance between each pair of samples i and j, as an integer between 1 (if the connection $\{i, j\}$ is present) and ∞ (if there is no path between i and j on the graph). The raw graph distance is the number of edges comprising the shortest path between the two pairs of samples in the graph (see Figures 2 and 3). For the dissimilarity matrix actually analyzed by DISTATIS we adapt the cophenetic dissimilarity from Koenig et al. (2021, see Figures 2 and 3): the corresponding cell entry a_{ij} of the $K \times K$ slice is defined as the subtraction from 1 of the inverse of the graph distance between i and j (defining $1/\infty$ as 0, and setting a minimum of 0 for dissimilarity of a sample with itself or with samples to which it is directly linked), so that the cell entries a_{ij} are no longer binary but range in the interval [0, 1], with larger values indicating lower similarity and smaller values standing for higher similarity. The diagonal of the matrix is set to 0, indicating that all samples are identical with themselves as would be expected for a distance matrix.

FIGURE 3 GOES HERE

Data were first analyzed by DISTATIS in order to compare consensus similarity configurations for samples across methods (Abdi et al., 2007). Confidence ellipses were generated through bootstrapping (Beaton et al., 2013). A key property of any rapid sensory method is how well samples and groups of samples are distinguished: this is clearly related to (but also not identical to) discrimination ability for the method. Examination of product separation on the first four DISTATIS axes for both methods via actual observations as well as bootstrapped confidence intervals were considered as evidence. Choice of 4 axes for examination (out of a possible 10 for each sample) were motivated by examination of scree plots for the DISTATIS $\boldsymbol{S}_+$ matrices (Abdi et al., 2007, not shown) as well as by general practice in industry and the literature for "significant dimensions" for interpretation.

The stability of results for a given number of subjects—that is, the required number of subjects—for each method was evaluated through a bootstrapping approach to simulate panels of different sizes and compare these simulated results to the actual, observed results. Specifically, generalized stability, termed RVb by Blancher et al. (2012), was calculated for free sorting and free linking: bootstrapped samples of subjects, of sizes 2 to N (where N is the number of subjects in the particular study) were drawn (with i = 100 replicates at each sample size), and the average RV between the DISTATIS $\boldsymbol{F}$ (factor score) matrices from the bootstrap sample and the full dataset was calculated at each sample size. Blancher et al. (2012) recommend that stability can be considered achieved at the number of subjects for which the bootstrapped average RVb exceeds 0.95.

Graph theory was also used to evaluate whether individual subjects' free-sorting and free-linking groupings were in fact different. For sorting, each individual's $K \times K$ slice was treated as the (symmetric) adjacency-matrix representation of an undirected graph (Gross et al., 2014). For linking, the undirected graph drawn by each individual was used. In each subject's graph, the nodes represent the samples, and an edge between two nodes indicates that the subject sorted or linked two samples as similar (Lahne, 2020; Orden et al., 2019). This graph representation provides several simple parameters that give insight into the similarity structure.

The *degree* of each node indicates how many edges are incident to it (Gross et al., 2014); thus, in sorting or linking higher degree for a node means the corresponding sample was considered similar to more other samples. Comparison of average degree per subject and sample for each method gives an indication of discrimination capacity: higher average degree indicates less discrimination between samples, as subjects consider more samples similar.

The *transitivity on triads* (Arney & Horton, 2014) is the fraction indicating, for the total number of node triads A, B, C with connections {A, B}, {B, C}, how many of them also contain the connection {A, C}. In the literature this is also called the graph "clustering coefficient" (Kolaczyk & Csárdi, 2014). In terms of the sorting and linking tasks, this is a measure of the likelihood that similarities {A,B} and {B,C} imply that similarity {A,C} also exists; when transitivity is higher it may indicate a lower discrimination capability.

The *average connectivity* of a graph (Beineke et al., 2002) is a parameter that measures, in each subject's results, the average over all pairs of nodes A and B, how many independent paths connect A and B. In the context of sorting and linking, lower average connectivity will be associated with more disjoint groups, which is an indicator of less robust or realistic models of similarity.

Subjects' preferences for method were evaluated for each study using simple contingency-table measures, and their opinions of the sorting and linking tasks' ease of use and enjoyability were evaluated using repeated measures ANOVA.

Data analyses were conducted in R (version 4.0.2). Code for analyses is available from the corresponding author upon request.

2.7. Ethics statement

All research methods were reviewed and approved by the Virginia Tech Human Research Protection Program (IRB # 19-1030).

3. Results

3.1. Product configurations (via DISTATIS)

The overall DISTATIS results for both the spice samples (Study 1) and the chocolate samples (Study 2) are quite similar (Figures 4 and 5). In the first 2 dimensions of the DISTATIS solutions the configurations of samples are almost identical, although it is worth noting that the derived distances among samples in the chocolate study are larger (Figure 5). However, for both studies it is apparent that the 3rd and 4th dimensions of the solution contain more valuable discrimination information for free linking than for free sorting. In each case, more samples are clearly discriminated (as can be seen from non-overlapping confidence ellipses) by subjects using free linking than by subjects using free sorting.

FIGURE 4 GOES HERE

The same basic product differences are identified by both methods, but with better resolution through free linking. For the spices, the first dimension separates cinnamon-containing mixes from the rest of the samples, while the second dimension separates cardamom-containing mixes (in both analyses the cinnamon+cardamom mixture falls in between these groups, with a stronger attraction to the cardamom region on the second axis). The third dimension for both studies separates pepper from the remaining samples, but with free linking it is also possible to infer that pepper is being directly opposed to turmeric-containing samples (Figure 4). In the fourth dimension, two samples that both contain turmeric are opposed: cardamom+turmeric and cinnamon+cardamom, but again in the free-linking study several other samples (cinnamon+pepper, cardamom) separate clearly on this dimension).

For the chocolate, the first dimension distinctly separates premium, dark chocolates from milk chocolates, while the second axis separates mass-market dark chocolates (Hershey's and Cadbury's) from the other samples. In the third dimension, the sole premium, milk chocolate (Endangered Species) is separated from the remaining samples, but only in the free-linking study is it clear that this dimension is capturing similarities between both chocolates from this producer (Figure 5). Finally, the fourth dimension separates the dark chocolate from Endangered Species from the remaining chocolates, but, again, in the free-linking study it is clear that there is more separation on this axis, with a strong separation between the two dark chocolates from Green & Black on this axis as well as separation among the other samples.

FIGURE 5 GOES HERE

3.2. Stability (via RVb)

In order to investigate stability of the solutions as a function of the number of panelists, RVb was calculated as described in Blancher et al. (2012). Figure 6 shows the RVb results for free sorting and free linking. As is apparent, the desired level of stability (the 0.95 level) is achieved with essentially the same number of subjects for both sorting and linking—although an average of about 1 subject less is required for stability in free sorting than in free linking. Given that this level of stability is achieved at between 8-10 subjects in these studies, this difference of a single subject is unlikely to be important in practical applications. In contrast to Blancher et al. (2012), we used all 10 dimensions to calculate RVb, but results for *only* Dimensions 1 and 2 (as calculated in Blancher et al. 2012) were almost identical to the full factor bootstraps (results not shown).

This is a quite low number of subjects when compared to those calculated by Blancher et al. (2012)—it corresponds most closely to the results in that study for a similar dataset of chocolate aromas (DS1, a free sort of 11 samples). While Blancher et al. (2012) do not give details on sample-inclusion criteria, in the case of both Study 1 and Study 2 samples were chosen specifically for their potential to be grouped by subjects (i.e., blends of the same spices and chocolates from the same manufacturers, see Table 1), which may explain the high stability observed here. It is also noticeable that the number of subjects required is slightly lower in Study 2 (chocolate, solid line) than in Study 1 (spice, dashed line). This difference seems like it may be attributed to the difference in modality—taste and flavor for Study 2, and only aroma for Study 1; differences in the products themselves may also be in play. This difference is also evident in the relative size and overlap of confidence ellipses for DISTATIS results (in which the RV coefficient is a key statistic) seen in Figures 4 and 5. However, there is no evident difference in the RVb patterns between sample type, modality, and methodology (sorting vs. linking). The apparent stability of each method is equivalent.

FIGURE 6 GOES HERE

3.3. Graph parameters

Three key graph parameters were investigated for this study. In a graph, the degree of a node represents the number of incident edges; for the sorting and linking studies, for each subject the degree of each sample indicates the number of other samples to which it was judged similar. Higher degree thus indicates a potentially lower discrimination ability among subjects, as fewer distinctions are made. For both Study 1 and Study 2, the degree distribution for free linking is clearly skewed more right than the degree distribution for free sorting (see Figure 7). Wilcoxon rank-sum tests indicate that the free-sorting task produces significantly larger degrees per node than the free-linking task for both the spice ($W = 143331, p < 0.05$) and the chocolate ($W = 167328, p < 0.05$) studies. This indicates that free linking better discriminates the samples than free linking.

FIGURE 7 GOES HERE

The *transitivity* on triads of a graph indicates the likelihood, given three nodes A, B, and C and edges {A, B} and {B, C}, that there will also be an edge {A, C}. In terms of free sorting and free linking, transitivity gives another indication of discrimination ability—it is a direct measurement of the degree to which similarities among samples are forced by the method or are allowed to be indicated by the subjects, and ranges from 0 to 1. In Figure 8, transitivity is plotted on the Y-axis against degree (see above) on the X-axis. By the nature of the sorting task, transitivity is always 0 or 1; it is only 0 in the degenerate case, when subjects made only pairs of samples, which happened several times in the spice study. For free sorting there is a much broader range of transitivity values in the [0, 1] range, indicating a higher likelihood of actual discrimination by the subjects.

FIGURE 8 GOES HERE

Finally, the *connectivity* of a graph is a measure, for each subject, of the number of distinct, connected paths between all pairs of nodes. Higher connectivity indicates a less disjoint (or disconnected) graph; in terms of free sorting and free linking, lower connectivity would mean more disjoint graphs, which are likely the result of a less realistic similarity model. In Figure 9, connectivity is plotted on the Y-axis against degree on the X-axis. For both studies, free linking tended to exhibit higher connectivity values than free sorting, as expected, but the differences were in general rather smaller than the differences in connectivity or degree. Thus, while subjects did produce more connected graphs using free linking than free sorting, they did not always produce fully connected graphs.

FIGURE 9 GOES HERE

3.4. Subject preferences

Finally, it is important to consider subjects' experience of the two tasks. In a simple question of overall preference ("Did you prefer the free-sorting or free-linking task?"), panelists preferred free-sorting to free-linking narrowly but insignificantly in Study 1 ($\chi_1^2 = 1.10, ns$), and by a broad and significant margin in Study 2 ($\chi_1^2 = 19.44, p < 0.05$; see Table 2). In neither task did it matter which task the subjects completed first (Study 1: $\chi_1^2 = 0.69, ns$; Study 2: $\chi_1^2 = 1.49, ns$). This can potentially be explained by the difference in complexity of the relative tasks: in Study 1, the test was by aroma only, whereas in Study 2 the subjects had to taste the chocolate. Therefore, it is possible that Study 2 involved a more fatiguing sensory task and a more taxing memory task, and in these circumstances it would make sense that subjects would prefer the simpler free-sorting task, which involves fewer pairwise comparisons. Alternatively, it is possible that the difference may be that the set of samples evaluated in Study 1 was "designed" by blending spices, providing an "easier" similarity structure.

Subjects also answered questions about ease-of-use and rated liking for each task, both on unstructured line scales converted to 10-pt values. Results were analyzed by mixed-effects ANOVA, with the dependent variable (liking or ease) modeled as dependent on the random effect of the particular subject, with the task (free sorting or free linking) as a within-subjects variable and the order of task completion as a between-subjects variable. For all tests, there was no effect of order of task, and no interaction between order and the task itself, so these results will not be reported in detail. For Study 1, subjects indicated that they did not find any difference in ease-of-use for the two tasks (effect of task on ease-of-use: $F_{1,56} = 0.016, ns$; free sorting $M = 7.62, SD = 1.78$, free linking $M = 7.14, SD = 2.08$), but they did report a significantly higher liking for the free-sorting task (effect of task on liking: $F_{1,56} = 5.14, p < 0.05$; free sorting $M = 7.57, SD = 1.70$, free linking $M = 6.85, SD = 2.08$). For Study 2, subjects indicated significant differences in both ease-of-use (effect of task on ease-of-use: $F_{1,61} = 24.76, p < 0.05$; free sorting $M = 8.36, SD = 1.48$, free linking $M = 6.96, SD = 2.38$)

and liking (effect of task on liking: $F_{1,61} = 19.40, p < 0.05$; free sorting $M = 7.54, SD = 1.61$, free linking $M = 6.11, SD = 1.92$). These results can be explained in the same way as the preference results: possibly a significantly higher memory and sensory-fatigue loads for tasting would make free linking a more difficult and less pleasant task than free sorting, or possibly the set of samples evaluated in Study 1 was slightly "easier" than the chocolates in Study 2. In both cases, it is also possible that subjects are simply more familiar with free sorting than with free linking, and familiarity has bred comfort with and preference for that method: while subjects were not surveyed about previous experience, our lab frequently conducts free-sorting studies and some subjects were definitely previous participants.

4. Discussion

Free sorting, as a rapid method for assessing similarities among a set of samples, has become an extremely popular method in both industry and academia (Dehlholm, 2015; Koenig et al., 2020, 2021; Valentin et al., 2012). However, the basic instruction of free sorting—that subjects form disjoint groups according to similarity—implies a model of similarity among the products that is likely to be unrealistic. Specifically, sorting requires that similarities be fully transitive and essentially unidimensional. In contrast, the method of pairwise free-linking, which we have formalized and demonstrated in this paper, provides results that are comparable to free sorting, while avoiding these restrictive assumptions.

In particular, on the same product sets, free linking results in significantly lower vertex degree measurements for each product, indicating that subjects are making more discriminating similarity judgments (Figure 7). In addition, the transitivity (or "clustering coefficient" Kolaczyk & Csárdi, 2014) of the similarity graphs from free linking were significantly more diverse than those from sorting, which are in general fully transitive (Figure 8); this explicitly indicates that subjects in free-linking studies are not forced to "close the triangle" when they want to indicate that A and B are similar, as are B and C. At the same time, the connectivity of the free-linking graphs was also noticeably higher than that of the free-sorting graphs (Figure 9), indicating that individual models of similarity generated through free linking were more robust, with graph distance giving a non-binary similarity measure (Chartrand & Zhang, 2014), which should capture a more multidimensional model of similarity.

This more "multidimensional" similarity is evident in DISTATIS biplots of results of free sorting and free linking on the same samples. Although for both spices (Figure 4) and chocolate (Figure 5) gross similarities, represented by Dimensions 1 and 2 of the biplots, are almost identical, there is much better discrimination of samples in Dimensions 3 and 4 for both sample sets. This follows naturally from the two different models of similarity implied by free sorting and free linking. Free sorting emphasizes rapidly finding gross similarities; free linking, while more intensive because of the need for multiple pairwise judgments (Figure 1), focuses on multidimensional similarity. Nevertheless, both methods provide stable results, as indicated by RVb, at approximately similar numbers of subjects (Figure 6). However, it is important to note that, on the whole, subjects found free sorting less taxing and more pleasant than free linking. It will be important to take subject fatigue into account when designing future studies that employ free linking. We might imagine that free linking would also be less fatiguing for trained subjects, who are used to making frequent, analytical, sensory judgments.

4.1. Limitations and future work

A key limitation of this study was the artificial nature of the sample sets: for both the spices and the chocolates, the samples were chosen to span a product category. In a real product-development or other applied situation, it is unlikely that there would be such a structured set of products. Arguably, free linking, which relies on pairwise comparisons, should perform better in these real situations, but this could not be determined from these sample sets. It also remains to be seen whether the lower preference and liking ratings for free linking by subjects will result in lower compliance or lower quality data when the method is used in a non-comparative setting.

The free-linking task also provides some new possibilities for the design of sensory studies. For example, to this point it has not been feasible to conduct free-sorting tests (or indeed projective-mapping tests) in an incomplete-block design, because the sorting space depends simultaneously on all samples. This has restricted the number of samples that can practically be analyzed in a free-sorting study to around 25 actual samples (the number is much higher for visual or text samples). This restriction should not apply to the free-linking task, which is based on a similarity graph of *pairwise* comparisons, but provides results that are similar or arguably superior to free sorting. Therefore, a logical future study is the investigation by free-linking of similarities in a set of samples large enough to present with an incomplete block design, but small enough to also investigate in full with free sorting in order to determine the comparability of this approach. Incomplete blocks for similarity would be a significant boon to food-sensory researchers in both industry and academia. In addition, given that free sorting appears to become exponentially more fatiguing as the number and sensory complexity of samples increases (see for example Kessinger et al., 2020), it may be hoped that free linking, which requires a larger number of simpler judgments, may perform better with large sample sets, especially when implemented in incomplete blocks as described above.

4.2. Conclusions

In this paper, we present a new, rapid method for assessing similarities among a set of samples: the "free-linking task". In the free-linking task, subjects are given a set of samples and asked to indicate pairwise similarity according to their own criteria; in effect, as we have demonstrated, subjects are drawing their own individual similarity graph for the samples. The data from free linking can be treated using existing tools for analyzing similarity data, such as DISTATIS, MFA, or even MDS.

The free-linking task explicitly solves two issues with the currently popular free-sorting task: in free sorting, subjects can only indicate one degree of similarity (is/is not similar) and are forced to make fully transitive similarity groups. While previously proposed modifications of sorting like the *hierarchical* and *multiple* free-sorting tasks can solve these respective tasks with replicated or multiple passes of sorting for each sorting, free linking solves both problems at once with only a single task. As we have demonstrated, therefore, the results of free linking provide a more realistic representation of similarity and allow finer and more powerful interpretations than free sorting. However, while the results of free linking are more realistic and robust, the cost is that free linking, because it involves more pairwise comparisons, is also more demanding for the participants. The multimensionality of free-linking data is also greater, which can be considered either a cost or a benefit, depending on the sensory analyst's goals. Therefore, we believe that the free-linking task will be a significant addition to the sensory analyst's arsenal of tools for rapidly assessing similarities, and we expect to see improvements and new uses cases for the tool in the near future.

5. Acknowledgments

David Orden is partially supported by Project PID2019-104129GB-I00 / AEI / 10.13039/501100011033 of the Spanish Ministry of Science and Innovation. Marino Tejedor-Romero is funded by a predoctoral contract from Universidad de Alcalá. This research did not receive any additional external funding. We would like to thank the volunteers who participated in this study, as well the undergraduate research volunteers at Virginia Tech who helped to coordinate data collection.

6. References

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Quality and Preference*, *18*(4), 627–640. https://doi.org/10.1016/j.foodqual.2006.09.003

Arney, D., & Horton, S. (2014). Network Science for Graph Theorists. In J. Gross, J. Yellen, & P. Zhang (Eds.), *Handbook of graph theory*. CRC Press.

Beaton, D., Fatt, C. C., & Abdi, H. (2013). *DistatisR: DiSTATIS Three Way Metric Multidimensional Scaling* (R Package version 1.0) [Computer software]. http://CRAN.R-project.org/package=DistatisR

Beineke, L. W., Oellermann, O. R., & Pippert, R. E. (2002). The average connectivity of a graph. *Discrete Mathematics*, *252*(1–3), 31–45.

Blanchard, S. J., & Banerji, I. (2016). Evidence-based recommendations for designing free-sorting experiments. *Behavior Research Methods*, *48*(4), 1318–1336. https://doi.org/10.3758/s13428-015-0644-6

Blancher, G., Clavier, B., Egoroff, C., Duineveld, K., & Parcon, J. (2012). A method to investigate the stability of a sorting map. *Food Quality and Preference*, *23*(1), 36–43. https://doi.org/10.1016/j.foodqual.2011.06.010

Chartrand, G., & Zhang, P. (2014). Distance in Graphs. In J. Gross, J. Yellen, & P. Zhang (Eds.), *Handbook of graph theory*. CRC Press.

Dehlholm, C. (2015). Free multiple sorting as a sensory profiling technique. In *Rapid Sensory Profiling Techniques* (pp. 187–196). Elsevier. https://doi.org/10.1533/9781782422587.2.187

Dehlholm, C., Brockhoff, P. B., Meinert, L., Aaslyng, M. D., & Bredie, W. L. P. (2012). Rapid descriptive sensory methods – Comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference*, *26*(2), 267–277. https://doi.org/10.1016/j.foodqual.2012.02.012

Delarue, J. (2015). The use of rapid sensory methods in R&D and research: An introduction. In *Rapid Sensory Profiling Techniques* (pp. 3–25). Elsevier. https://doi.org/10.1533/9781782422587.1.3

Gross, J., Yellen, J., & Zhang, P. (Eds.). (2014). *Handbook of graph theory*. CRC Press.

Kessinger, J., Earnhart, G., Hamilton, L., Phetxumphou, K., Neill, C., Stewart, A. C., & Lahne, J. (2020). Exploring Perceptions and Categorization of Virginia Hard Ciders through the Application of Sorting Tasks. *Journal of the American Society of Brewing Chemists*, 1–14.

Koenig, L., Cariou, V., Symoneaux, R., Coulon-Leroy, C., & Vigneau, E. (2021). Additive trees for the categorization of a large number of objects, with bootstrapping strategy for stability assessment. Application to the free sorting of wine odor terms. *Food Quality and Preference, 89*, 104137. https://doi.org/10.1016/j.foodqual.2020.104137

Koenig, L., Coulon-Leroy, C., Symoneaux, R., Cariou, V., & Vigneau, E. (2020). Influence of expertise on semantic categorization of wine odors. *Food Quality and Preference, 83*, 103923. https://doi.org/10.1016/j.foodqual.2020.103923

Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical analysis of network data with R*. Springer.

Lahne, J. (2020). Sorting Backbone Analysis: A network-based method of extracting key actionable information from free-sorting task results. *Food Quality and Preference*, 103870. https://doi.org/10.1016/j.foodqual.2020.103870

Lahne, J., Abdi, H., & Heymann, H. (2018). Rapid sensory profiles with DISTATIS and Barycentric Text Projection: An example with amari, bitter herbal liqueurs. *Food Quality and Preference*, *66*, 36–43. https://doi.org/10.1016/j.foodqual.2018.01.003

Orden, D., Fernández-Fernández, E., Rodríguez-Nogales, J. M., & Vila-Crespo, J. (2019). Testing SensoGraph, a geometric approach for fast sensory evaluation. *Food Quality and Preference*, *72*, 1–9. https://doi.org/10.1016/j.foodqual.2018.09.005

Orden, D., Fernández-Fernández, E., Tejedor-Romero, M., & Martínez-Moraian, A. (2021). Geometric and statistical techniques for projective mapping of chocolate chip cookies with a large number of consumers. *Food Quality and Preference*, *87*, 104068.

Spencer, M., Sage, E., Velez, M., & Guinard, J.-X. (2016). Using Single Free Sorting and Multivariate Exploratory Methods to Design a New Coffee Taster's Flavor Wheel: Design of coffee taster's flavor wheel…. *Journal of Food Science, 81*(12), S2997–S3005. https://doi.org/10.1111/1750-3841.13555

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: A review of new descriptive methods in food science. *International Journal of Food Science & Technology*, *47*(8), 1563–1578. https://doi.org/10.1111/j.1365-2621.2012.03022.x

Varela, P., & Ares, G. (Eds.). (2014). *Novel Techniques in Sensory Characterization and Consumer Profiling*. CRC Press. https://doi.org/10.1201/b16853

Tables

Table 1. Sample information for Study 1 and Study 2.

Study 1 – Spices*		
Sample Name	**Recipe**	
Cinnamon	*1 g ground cinnamon*	
Cardamom	*1 g ground cardamom*	
Pepper	*1 g ground black pepper*	
Turmeric	*1 g ground turmeric*	
Cinnamon + cardamom	*0.5 g ground cinnamon + 0.5 g ground cardamom*	
Cinnamon + pepper	*0.5 g ground cinnamon + 0.5 g ground black pepper*	
Cinnamon + turmeric	*0.5 g ground cinnamon + 0.5 g ground turmeric*	
Cardamom + pepper	*0.5 g ground cardamom + 0.5 g ground black pepper*	
Cardamom + turmeric	*0.5 ground cardamom + 0.5 g ground turmeric*	
Pepper + turmeric	*0.5 ground black pepper + 0.5 g ground turmeric*	
Study 2 - Chocolate		
Manufacturer	**Chocolate type**	**Cocoa content**
Cadbury	*Dark*	*35%*[?]
Hershey's	*Dark*	*45%*[?]
Green & Black's	*Dark*	*70%*
Endangered Species	*Dark*	*72%*
Green & Black's	*Dark*	*85%*
Pascha	*Dark*	*85%*
Cadbury	*Milk*	*26%*[?]
Hershey's	*Milk*	*30%*[?]
Green & Black's	*Milk*	*34%*
Endangered Species	*Milk*	*48%*

*All spices are McCormick Gourmet Organic line ground spices (no whole spices were used for the purpose of blending the recipes).
[?]information gathered indirectly from manufacturer's website rather than packaging.

Table 2. Counts of preference for free-sorting or free-linking task for each study, counted by which test was completed first.

Task Completed First	**Prefer Free-Sorting**	**Prefer Free-Linking**
Study 1: Spices (by smell)		
Free-Sorting	15	15
Free-Linking	10	18
Study 2: Chocolate (by taste)		
Free-Sorting	24	10
Free-Linking	25	4

Figures

Figure 1. Interface for individual subjects' free-linking task, as rendered in SensoGraph (Orden et al., 2019). Note that sample order is randomized between subjects.

Figure 2. Schematic representation of free sorting (top) and free linking (bottom). From the same samples (presented in random order to each subject) in (1), the methods diverge. For free sorting, subjects group samples (2) and their groupings are transformed directly to binary dissimilarities (3). For free linking, subjects indicate pairwise similarity (2), which is transformed into graph distances (3), and then to [0,1]-range dissimilarity (4, with details given in Figure 3). At this point, the same analyses can be conducted on the each of the dissimilarity matrices.

Figure 3. Schematic for deriving dissimilarity from graph distance, based on Koenig et al. (2021).

Figure 4. DISTATIS biplots for free sorting (top, in purple) and free linking (bottom, in orange) of spice-study results. The left-hand column gives Dimensions 1 and 2, while the right-hand column gives Dimensions 3 and 4 of the respective spaces.

Figure 5. DISTATIS biplots for free sorting (top, in purple) and free linking (bottom, in orange) of chocolate-study results. The left-hand column gives Dimensions 1 and 2, while the right-hand column gives Dimensions 3 and 4 of the respective spaces.

Figure 6. Stability of consensus solutions as assessed by RVb for free linking (purple) and free sorting (orange) in spice (dashed) and chocolate (solid) studies.

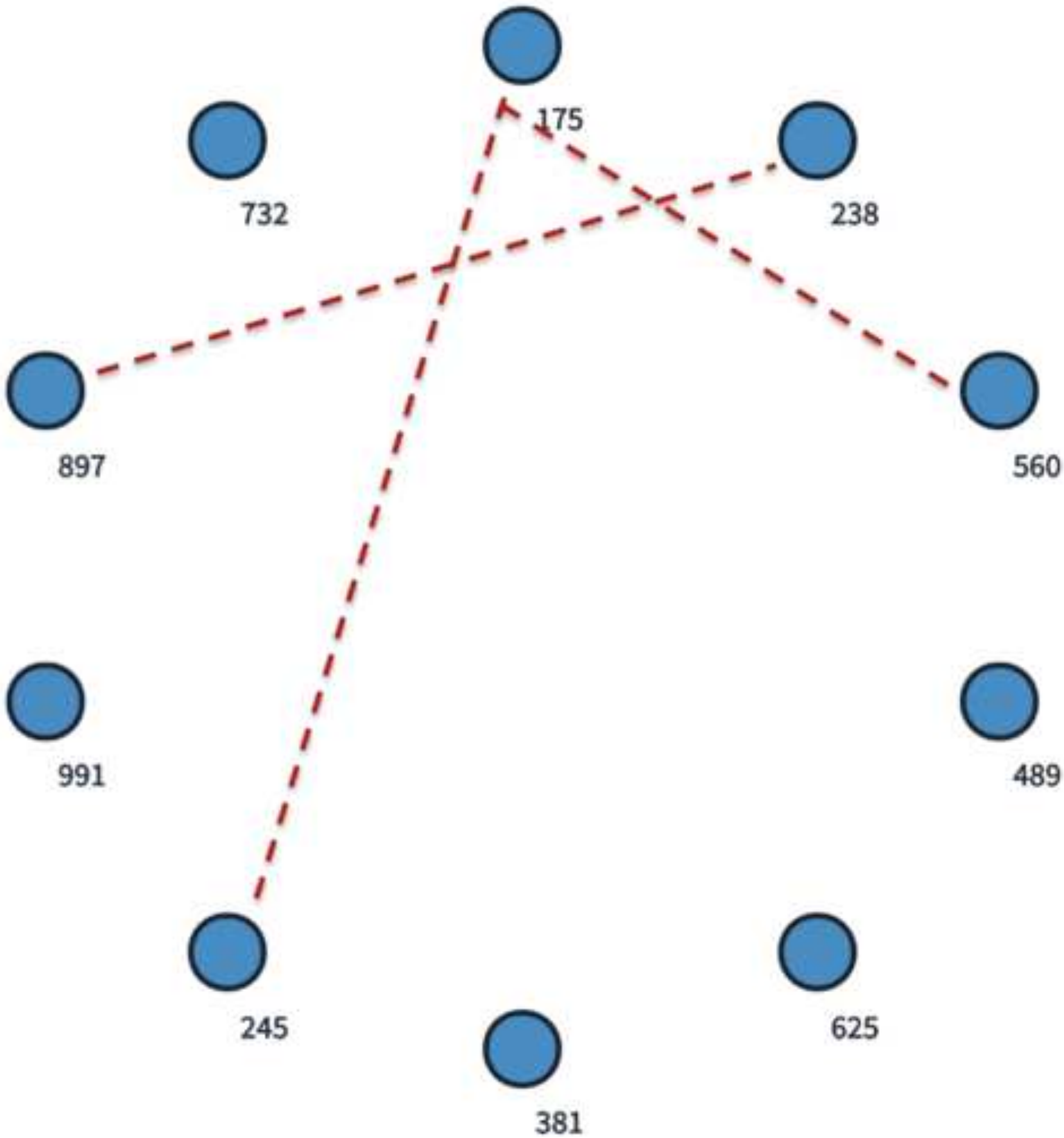
Figure 7. Degree distributions for spice (left) and chocolate (right) studies for free linking (purple) and free sorting (orange). In these studies, higher degree indicates less power to discriminate among samples.

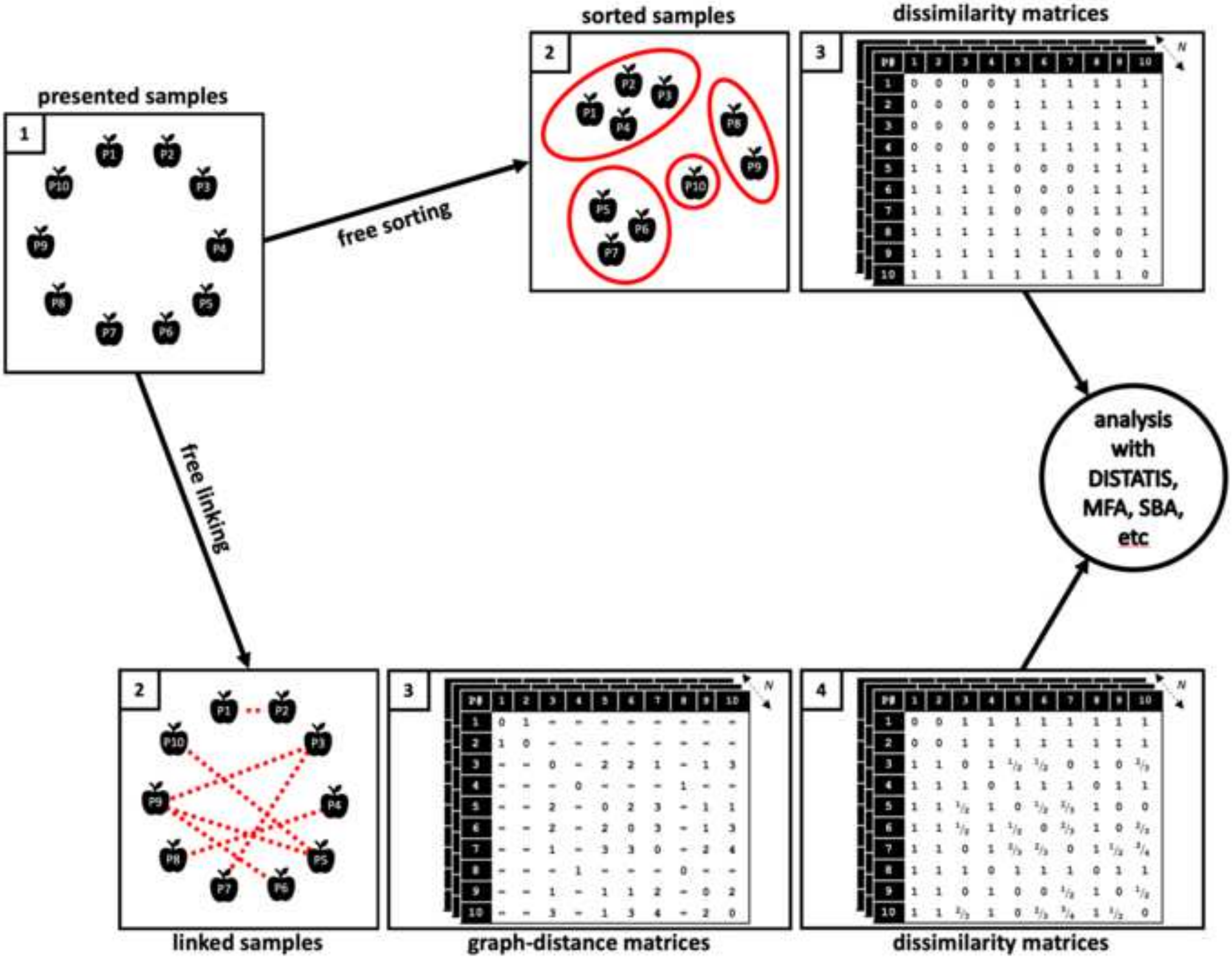
Figure 8. Scatter plots of individual subjects' degree (with lower degree indicating higher discrimination power) against transitivity (clustering coefficient, with higher values indicating forced grouping/similarity) for free linking (purple) and free sorting (orange). Note that for free sorting, transitivity is *always* equal to 1 except in the rare degenerate case in which subjects only make groups of 2 or fewer samples (bottom left).

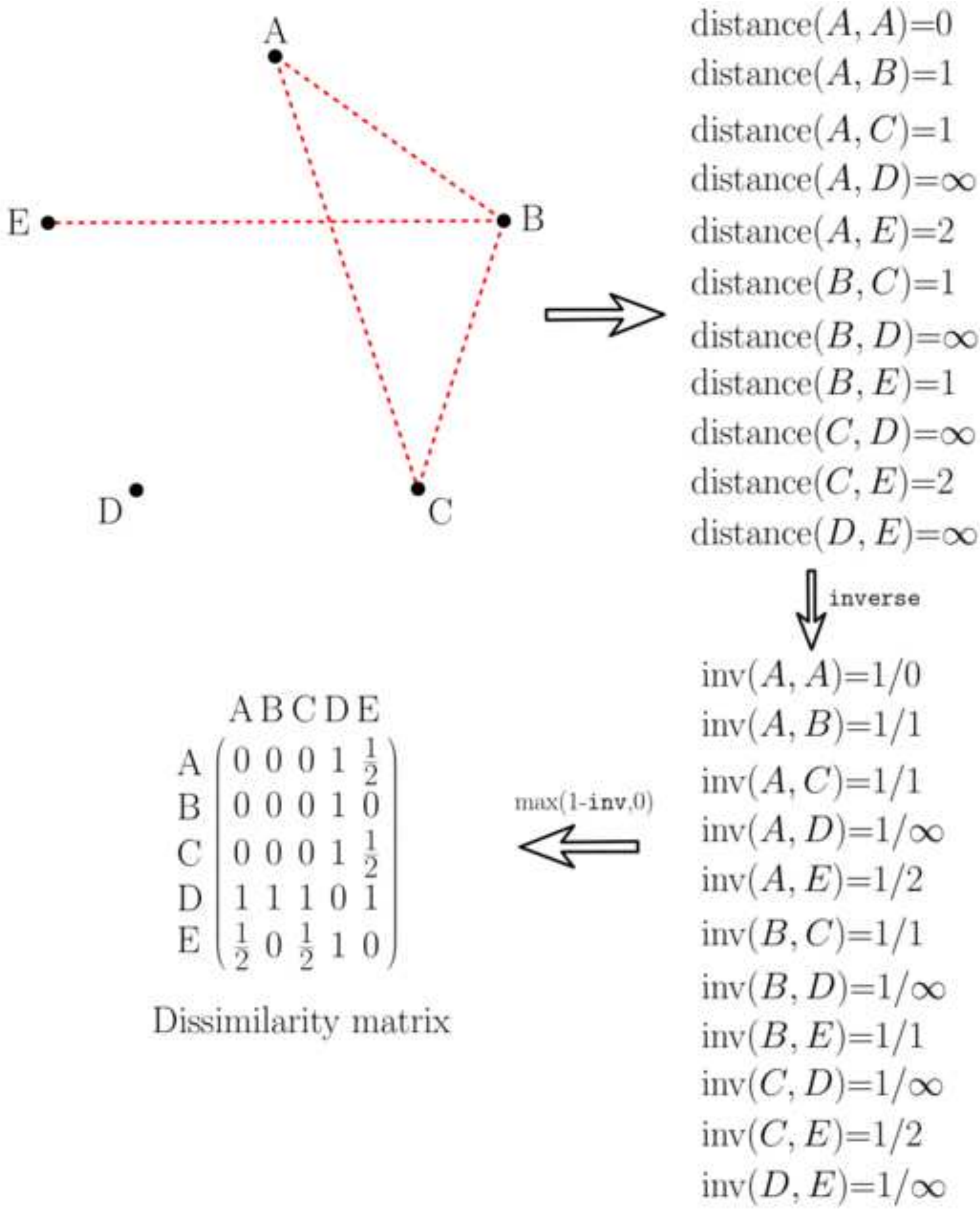
Figure 9. Scatter plots of individual subjects' degree (with lower degree indicating higher discrimination power) against connectivity (with higher values indicating ability to detect multiple levels of similarity). Note that for free sorting, only high values of degree guarantee higher connectivity, whereas in free linking higher connectivity is achieved at lower degree (with higher discrimination power).

F1

Click here to access/download;Figure;Figure 1 - Linking Tool Illustration.png

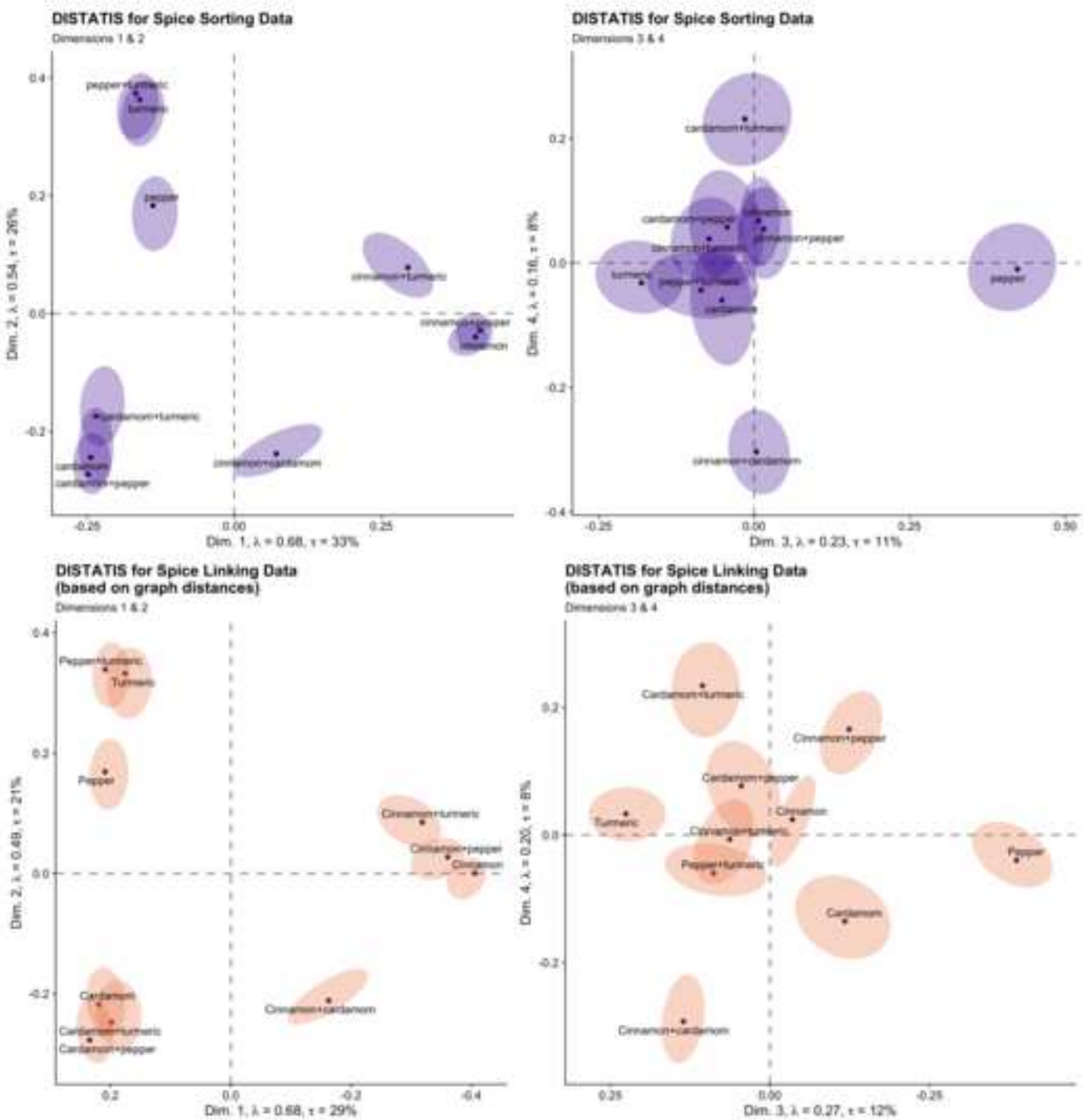



presented samples
sorted samples
dissimilarity matrices
free sorting
free linking
analysis with DISTATIS, MFA, SBA, etc
linked samples
graph-distance matrices
dissimilarity matrices


A
B
C
D
E
distance(A, A)=0
distance(A, B)=1
distance(A, C)=1
distance(A, D)=∞
distance(A, E)=2
distance(B, C)=1
distance(B, D)=∞
distance(B, E)=1
distance(C, D)=∞
distance(C, E)=2
distance(D, E)=∞
inverse
inv(A, A)=1/0
inv(A, B)=1/1
inv(A, C)=1/1
inv(A, D)=1/∞
inv(A, E)=1/2
inv(B, C)=1/1
inv(B, D)=1/∞
inv(B, E)=1/1
inv(C, D)=1/∞
inv(C, E)=1/2
inv(D, E)=1/∞
max(1-inv,0)
A B C D E
A 0 0 0 1 1/2
B 0 0 0 1 0
C 0 0 0 1 1/2
D 1 1 1 0 1
E 1/2 0 1/2 1 0
Dissimilarity matrix

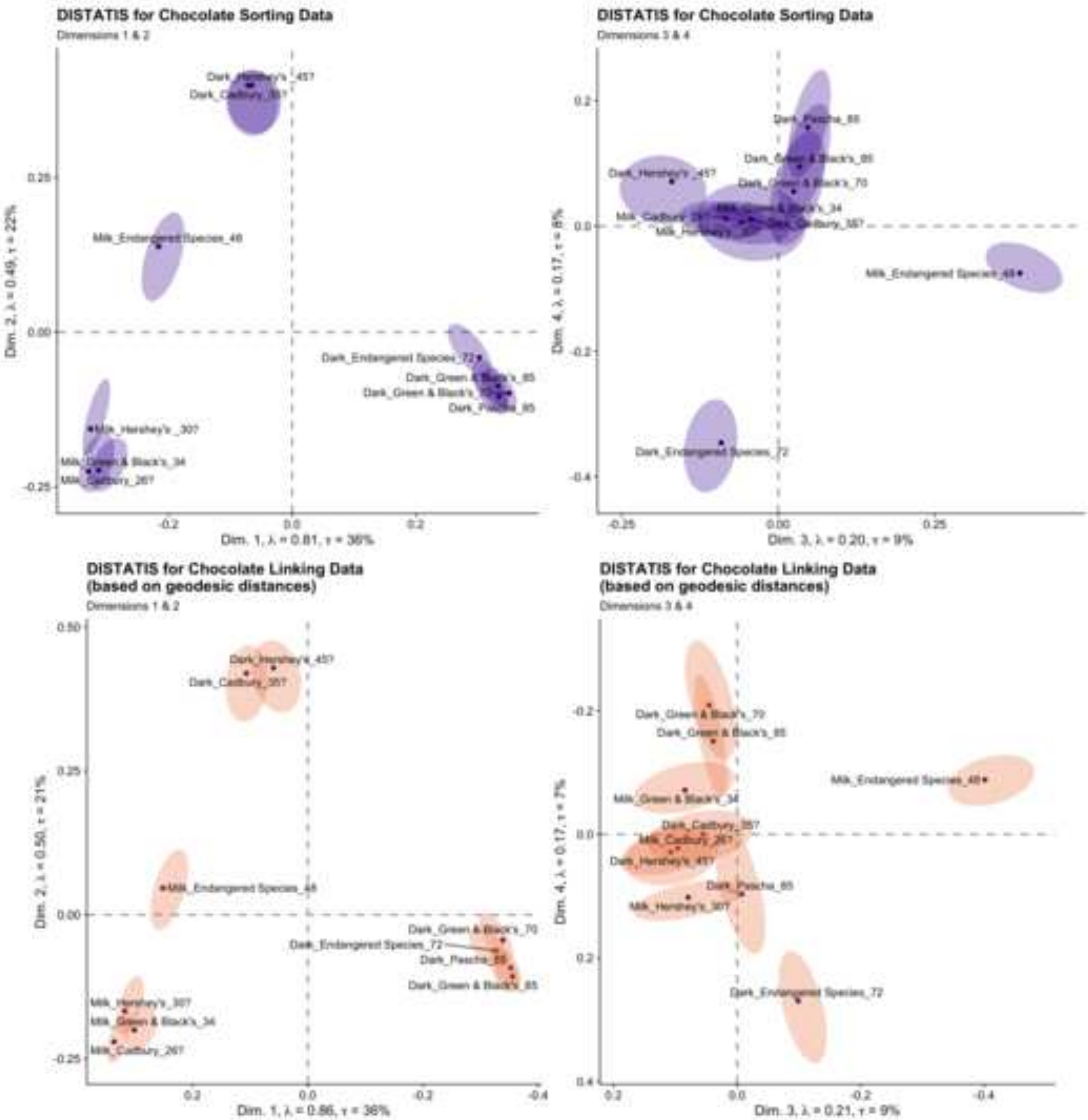
F4

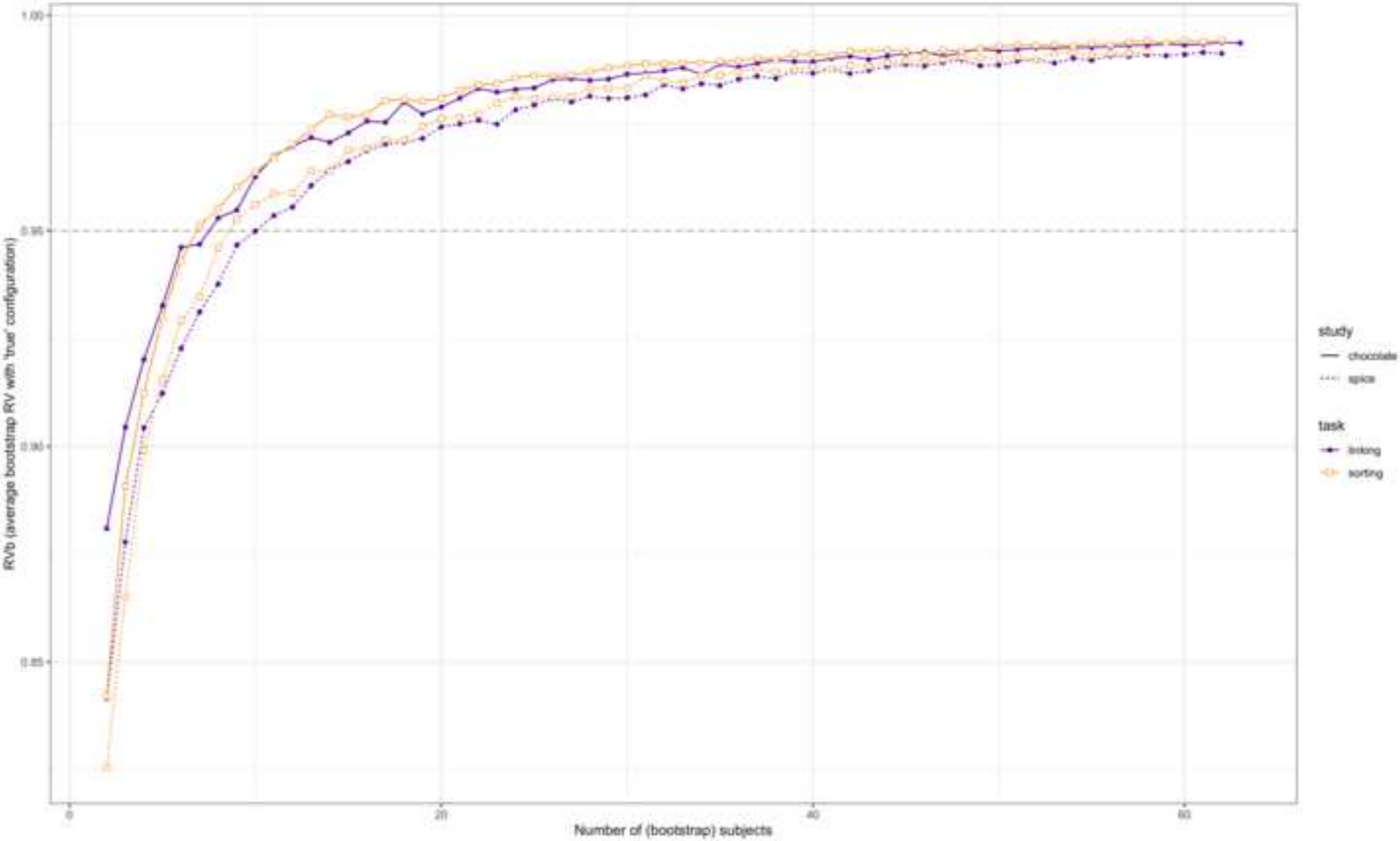
Click here to access/download;Figure;Figure 4 - Spice DISTATIS - R1.png



F5

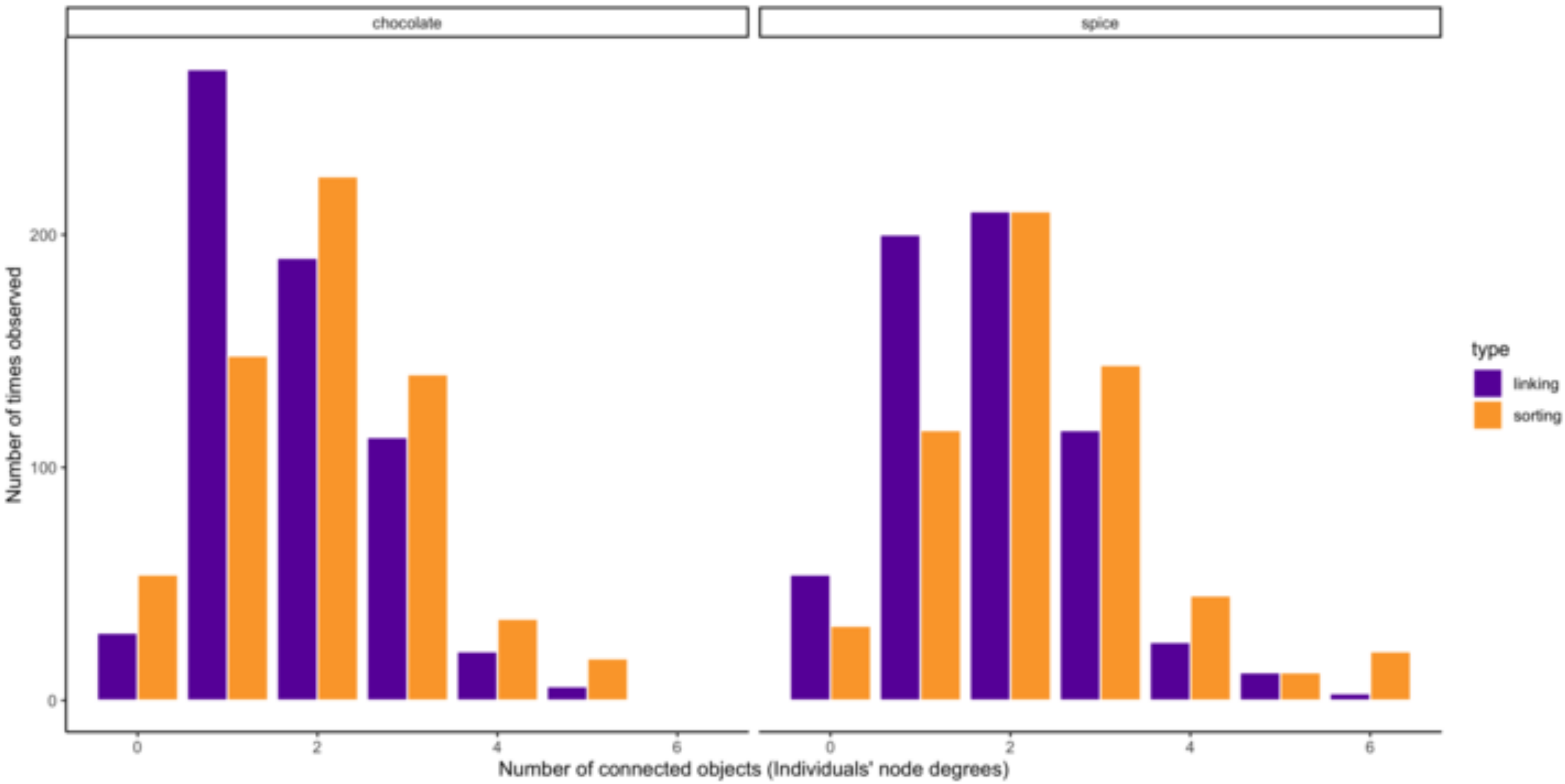
Click here to access/download;Figure;Figure 5 - Chocolate DISTATIS - R1.png




RVb (average bootstrap RV with 'true' configuration)
1.00
0.95
0.90
0.85
0
20
40
60
Number of (bootstrap) subjects
study
chocolate
spice
task
sorting

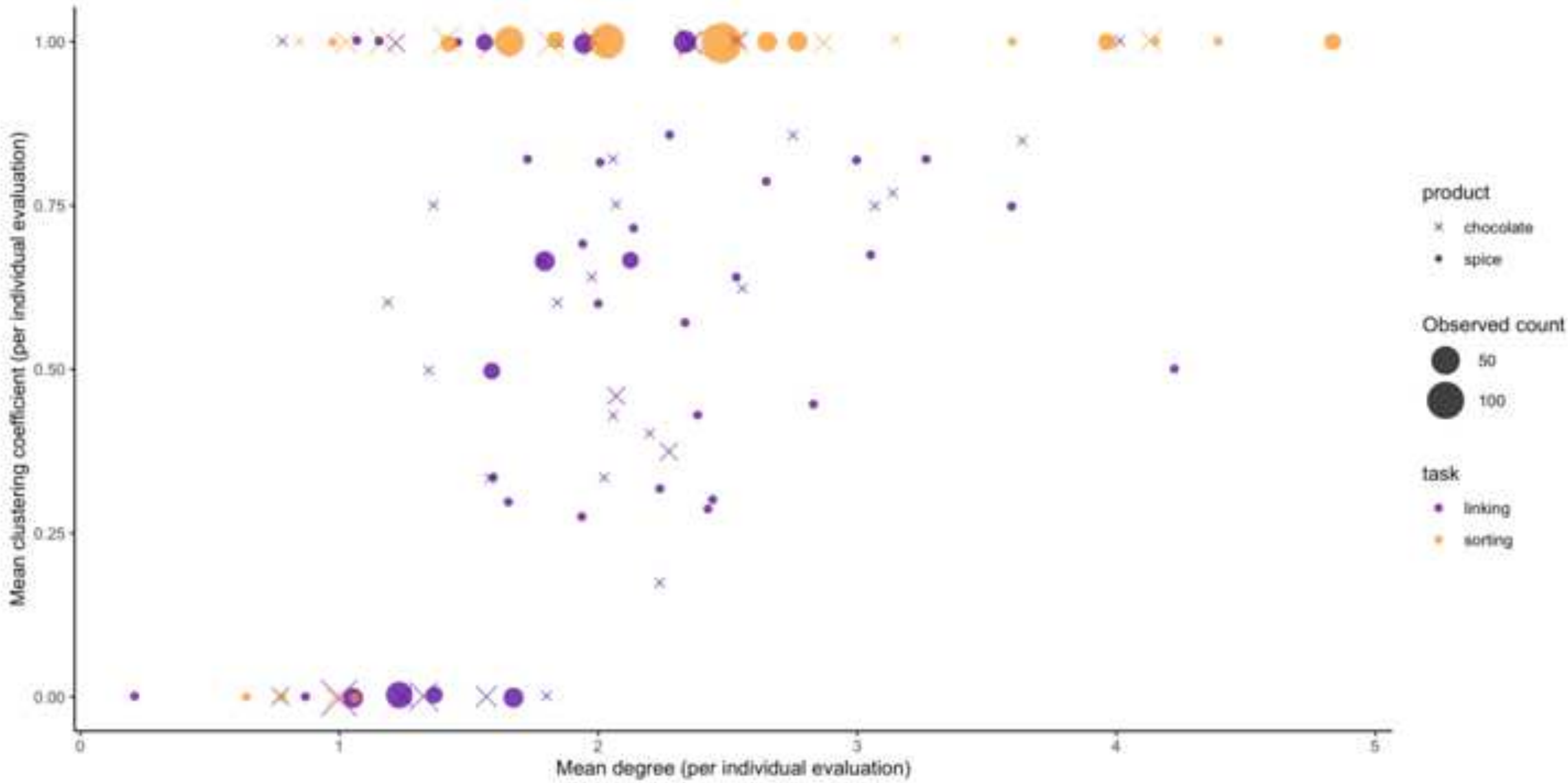
F7

Click here to access/download;Figure;Figure 7 - degree distributions.png



F8

Click here to access/download;Figure;Figure 8 - Degree vs Transitivity - R1.png



F9

Click here to access/download;Figure;Figure 9 - Degree vs Connectivity - R1.png