



Card Sorting with SynCaps (All Versions)

William Hudson
Syntagm Ltd
7 July 2016

Contents

1	Introduction.....	3
2	Decide on the type of sort	6
3	Print Cards	6
4	Sort Cards	12
5	Capture Data	13
6	Online Sorting	17
7	Prepare and Analyze Data.....	17
8	Analyze Additional Data.....	28
9	Advanced Concepts.....	38

1 Introduction

This user guide is a step-by-step description of how to prepare and process card sorting activities, both face-to-face and online. It covers all versions of the SynCaps¹, including the free edition (V1).

In addition to the user guide we have also produced a Release Notes document for SynCaps V3. If you are an experienced user of earlier versions of SynCaps you may find it quicker to start with that document.

1.1 Resources

In addition to the two documents described above, there are a number of free resources that can help you to understand and apply card sorting as a user-centred design tool:

- Card Sorting entry on the Interaction Design Encyclopedia web site (bit.ly/ixd-card-sorting)
- Webinars, videos and slide shows. We have three one-hour webinars plus a variety of shorter how-to videos and slide shows. The webinars and videos can be found on YouTube – just search for ‘caps card sorting’. Or use the links below:
 - [Part 1: Preparing Paper Sorts](#)
 - [Part 2: Basic Analysis \(Paper and Online Sorts\)](#)
 - [Part 3: Advanced Analysis \(SynCapsV3\)](#)
- More resources can be found at www.syncaps.com and at www.syntagm.co.uk/design/articles (including the Interactions article, *Playing Your Cards Right: Getting the Most from Card Sorting for Navigation Design*).

1.2 What's different between SynCaps versions 1, 2 and 3

- ▶ SynCaps version 1 produces a dendrogram and a set of exported spreadsheet files. V1 takes no note of group names (an Items x Groups chart is provided in SynCaps V3). V1 requires that all items (but not groups) be known in advance. Also, participants can only put the same item into more than one group if multiple sets of cards are available for each participant.
- ▶ Version 2 added a Pairs Map and an Items x Groups chart (plus a number of other features described below). It also allowed ‘ad hoc’ items to be made up by participants during the sorting exercise. These ad hoc items are simply keyed in during data collection. Also, in V2 and later, an item can appear in multiple groups simply by replicating the card in question (using either its item number

¹ Caps stands for Computer-Aided Paper Sorting. SynCaps (which also handles online sort data) is our desktop analysis application for Windows.

or the full item name). **NOTE that V2 is no longer available or supported and has been replaced by SynCaps V3.**

- ▶ Version 3 introduces the concept of nested groups which affects how card sorts are prepared. It also introduces a number of innovative features participant analysis and participant filtering. These are summarized in sections 7.6-7.7 and 9.6-9.7. They are described in detail in the Release Notes for SynCaps V3 document.

1.3 Licensing SynCaps

SynCaps V1 is free; just download it and run it from your desktop. All versions of SynCaps are compatible with most versions of Windows and will run very effectively in Windows emulators for Mac and Linux hosts. (Due to limitations in font licensing between Microsoft and Apple, the bar code fonts embedded in the templates for printing cards will not appear in the Mac versions of Microsoft Word. You need to run a Windows version to use the bar code fonts.)

SynCaps V3 requires a machine-specific license. It will run in trial mode for 14 days from installation, with a limit of 10 participants. After the trial is finished you may still use SynCaps V3 to view caps3 files, but you will not be able to perform any new analyses. To buy a license, download and install the software, then click the 'Buy Now' button in the licensing dialog. This will send the generated serial number to our web site and start the checkout process. The current cost of a license will be displayed in USD, GBP and EUR. You can choose the most suitable currency for billing. (Academic licences are free, but limited to one per academic email address. If your academic address is not recognized by the checkout process, please write to us at support@syntagm.co.uk. You still have to go through checkout, but the cost is nil. Choice of currency and VAT registration details are not relevant to academic licenses.)

[Existing SynCaps V2 users can upgrade to V3 at no cost. You must have a licensed copy of V2 installed on the machine you want to use. Then download and install V3. Click the 'Upgrade' button in the licensing dialog and you will be taken through a short (but free) checkout process. You will need to paste the SynCaps V2 serial number into the upgrade page. If an upgrade button is not displayed, the machine in question does not have a valid V2 license. Send us your V2 purchase details and the serial number from the V3 licensing dialog and we will reply with a V3 license key (email address as above).]

In all cases, the license is per-machine so can be used by anyone who logs in. However, V3 will request a license key for each new user. The key is the same for all users so just copy and paste it from your purchase confirmation email.

1.4 Overview

Card sorting is an easy, quick technique that helps you to find out how people think about and group things. The 'things' could be menu items, pieces of your information architecture, product categories of anything else that you can describe in a couple of words or a short sentence.

SynCaps is a method of doing card sorting that combines the easy handling and familiarity of paper with a robust calculation package. It also supports online sorting, so that you can compare or combine data from multiple sources.

Here is an outline of the steps in the SynCaps card sort process:

1. **Decide on type of sort:** There are two main types of card sorting
 - ▶ Closed: you decide on the groups and their names.
 - ▶ Open: participants determine the number and names of groups (although you may provide guidance and limitations).

SynCaps supports both types as well as a hybrid sort where participants are given some group names but are allowed to make up others.

2. **Print cards or set up online sort:** Print participant, group and item cards (or adhesive labels). We'll explain how to do this and how many cards you need. Alternatively, online sorting sites WebSort and OptimalSort both export data in SynCaps format. Just set up the item and group names and invite participants.
3. **Sort cards:** Get participants to sort the paper or online cards. For paper cards with quality of fit, ask participants to choose how well each item fits the group in which they've place it once sorting is complete. (If you are using the alternatively labelled 'frequency of use' cards, you may wish to consider asking participants to mark this first as it provides better engagement with the topics on the cards.)
4. **Capture data:** It's easier and quicker to capture data using a bar code reader, but if you don't have one then you can type it in. For online sorts, just export the data in SynCaps format (details below).
5. **Prepare and analyze data:** SynCaps imports the raw data and does the calculations for you. It prepares a cluster-analysis dendrogram for inclusion in reports and exports additional data for you to analyze. SynCaps V3 also produces specially-designed charts to help see the patterns of items and groups.
6. **Analyze additional data:** You can perform further analysis using software such as Excel. The data exported by SynCaps includes a variety of measures by item, by participant and proximity matrices (don't worry if this means nothing to you, it is explained later). Also, SynCaps V3 creates script files for SPSS should you want to run your own analyses.
7. **Perform participant filtering (SynCaps V3 only):** Omit participant outliers and/or identify separate participant communities who may have differing views of the problem domain. Participant filters produce alternative analyses that can be viewed side-by-side.

This document describes these steps in order. Refer to the separate Release Notes for SynCaps V3 if you want to concentrate on the new features offered.

2 Decide on the type of sort

If you are at an early stage in your project and do not have any names for your categories, then **open card sorting** is likely to be best for you. You ask your participants to sort the cards as they wish and to create their own names for the groups.

If you already have a set of category names, for example the titles of menus, then **closed card sorting** will help you to find out if those names are likely to work for your participants. You can also give participants group names that you are fairly certain about and let them make up the others. We call this a **hybrid sort**.

In all cases, you need a list of the items that you want your participants to sort. (Participants can also make new items up as they are going along if you are using SynCaps V3.)

3 Print Cards

If you are going to use a bar code reader, you need to print the bar codes and text onto cards or adhesive labels. The Word mail-merge documents we provide² for cards expect you to use card stationery that splits into four cards per page. Commonly available stationery includes Avery 3381 (US-letter) and Sigel LP 711 (A4). However, we provide very cost-effective quantities of cards for customers in Europe (see <http://www.syntagm.co.uk/design/cardsorting.htm>).

We also provide templates for two common sizes of Avery self-adhesive labels: L7163 (for A4 printers) and 8162 (for US Letter printers). These both have 14 labels per sheet. Note though that the labels do not include the 'quality of fit' choices available on the cards. We suggest that you stick the labels onto index cards, or onto the photographs or objects being sorted.

Calculate how many sheets you need, for example:

FOR CARDS (4 per sheet)

- Take the number of items you have and divide by 4, rounding up to the nearest whole number (also allow some spare for participants to add their own cards).
- Divide the number of groups you have (or the maximum you think you may have in an open card sort) and divide by 4, rounding up to the nearest whole number.
- Add the number of item and group sheets you get from the above together and multiply by the number of participants.

² The templates are automatically installed with SynCaps V3 and can be found in the My Documents\SynCapsV3\Templates folder under the account used for installation. For other versions of SynCaps you need to [download the templates](#) separately.

- In some cases you may want to use sub-group cards as well (see section 9.2). You can afford to be generous with the number of sub-group cards you print as they are all the same.
- Also include one participant card each (again, divide by 4 and round up).

FOR LABELS (14 per sheet)

- Take the number of items you have and divide by 14, rounding up to the nearest whole number (also allow some spare for participants to add their own cards).
- Divide the number of groups you have (or the maximum you think you may have in an open card sort) and divide by 4, rounding up to the nearest whole number.
- Add the number of item and group sheets you get from the above together and multiply by the number of participants.
- In some cases you may want to use sub-group cards as well (see section 9.2). You can afford to be generous with the number of sub-group cards you print as they are all the same.
- Also include one participant card each (again, divide by 4 and round up).

If you have a lot of items, you may find that you'll need a surprisingly large number of cards and it sometimes takes a few days to get hold of the right card stationery. Meanwhile, you can prepare the data for printing.

If you already have experience of Mail Merge in Word, you should find the process quite easy.

If you are not familiar with Word or Mail Merge, the instructions may seem daunting at first, but it will feel very straightforward after the first couple of times you try it.




<p>Item name</p> <p>A line or two to provide missing context (optional)</p> <p>Alternatives: Other words that users might prefer (optional)</p>	<p>Fit</p> <p><input type="checkbox"/> Perfect</p>  <p><input type="checkbox"/> Good</p>  <p><input type="checkbox"/> Fair</p>  <p style="text-align: right; margin-top: 20px;">1</p>
--	---

Figure 1, Sample Item Card (with Quality of Fit)

Item and Group Cards

- 3.1 Get hold of your list of items that you want to sort. You need to turn it into SynCaps Item Data. The easiest way to do this is by editing the sample file ItemData.doc that came as part of your download. Your Item Data will be used as the 'Data Source' later, so remember where you put it.

ItemData.doc is a table with numbered entries. You need to put an item in each line of the table, and you can add optional description and alternative description if you wish. You can use any character sets you like (including, for example, Greek and Cyrillic) but you should save files using any non-Latin characters as Unicode.

- 3.2 The table has 100 lines pre-formatted. You can add extra lines if you need them or delete unused lines. (It is important to delete unused lines since Word will otherwise try to print additional cards for each one.) **Be careful to ensure that all of your items and groups are consecutively numbered without duplications or omissions.**

No	Item	Description	Alternatives
1	Put your item here	Add an optional description here	Add other words that participants may prefer (optional)

3.3 If you are doing a closed sort, prepare your group list by editing GroupData.doc. If you are doing an open sort, just use blank cards and ask participants to write their name for the group on the card.

3.4 Locate the appropriate master file for the paper size you are using. It is important that you start with the correct files – otherwise Word may resize the pages to fit your paper, rendering the barcodes unscannable.

- If you are using A4: you need ItemMasterA4.doc for item cards and if doing a closed sort, GroupMasterA4.doc

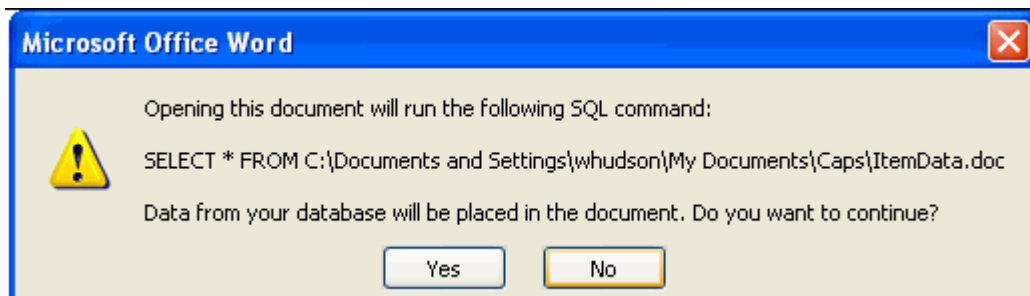
- If you are using US Letter, you need ItemMasterLt.doc for item cards and if doing a closed sort, GroupMasterLt.doc.

- In both cases there are versions of the item cards without the quality of fit choices. These have the words 'NoFit' appended to the filenames. While we do find that quality of fit improves results, particularly for smaller sample sizes (fewer than 30 participants), you are not obliged to use it.

- There are also alternative versions of the Quality-of-Fit scale that measures frequency-of-use instead. The files for these end with the word 'Use'.

Do not edit the master files unless you want to change the layout or content of all cards. They are formatted for use with Mail Merge. Save the data files using any name you like (and can remember). We will refer to them as ItemMaster and GroupMaster from now on.

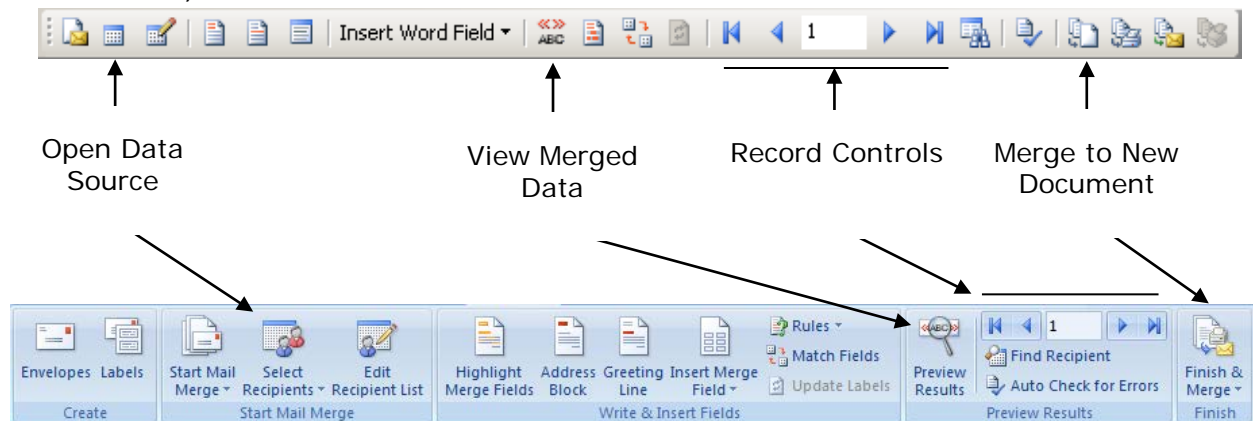
3.5 Open your chosen ItemMaster in Microsoft Word. You will get a message box like this:



Simply click 'No'. You will instruct Word about where to get the data in a couple of steps, below.

3.6 In Word 2003 or earlier, if the Mail Merge toolbar is not visible, then enable it by either: using the View -> Toolbars -> Mail Merge menu item
or: by right-clicking in the menu and toolbar area and checking Mail Merge.

3.7 You should then see something similar to this (Word 2003 – top; Word 2007 – bottom):



(Versions of Word prior to 2003 are very similar to the top toolbar shown here.)

3.8 Open your item list by clicking on the **Open Data Source** button shown above and navigating the appropriate folder.

3.9 If you have opened your item list successfully, you should see some of your items in the ItemMaster. If you get field names in <<double angle brackets>> instead of your items, then click the **View Merged Data** button to see the first page of your items (called **Preview Results** in Word 2007 and later).

3.10 It is advisable to look through all the cards to check that they all worked OK. We generally use the 'Record Controls' to do this.

If you find any problems in the cards, then you need to adjust ItemData.doc:

either: close everything down and go back to edit ItemData.doc in your preferred editor

or: click on the button to the right of **Open Data Source** (confusingly described as **Mail Merge Recipients** in the popup text).

3.11 Once the cards look okay, create a file for printing by clicking the **Merge to New Document** button in Word 2003 and earlier versions. Oddly, in Word 2007 this is called **Edit Individual Documents** and can be found under the **Finish & Merge** button in the **Mailings** ribbon. .

Quickly check that the file appears exactly as you want it. You can make alterations directly to this document if you like, but do not forget to adjust your data file if necessary.

3.12 Select any options provided in your printer's property dialog for card or other thick material. (This is very important for some printers as otherwise the barcodes may smudge.) In particular be sure to turn off double-sided printing if it is on by default.

3.13 Print out a few test sheets on the material you plan to use.

- 3.14 Check the resulting cards with the barcode reader by opening a new document within Word and scanning some of the codes. Each code should appear on a new line, start with a capital letter ('I' for item or 'G' for Group) and be followed by the number that is printed in the bottom-right corner of each card.
- 3.15 Assuming that all is well, print copies of the merged document for each of your participants plus one or two spare sets just in case.
- 3.16 Repeat this process for the group cards, if necessary (only for a closed sort).

Note that because the barcode fonts are embedded in the original template, you cannot save the merged results as a Word document (the barcodes disappear). You can however print them to an image device such as the Adobe PDF printer or Microsoft XPS Document Writer. Both of these will create files complete with barcodes that may be saved for later or sent via e-mail.

Sub-Group Cards

- 3.17 Sub-groups are an advanced feature that you may wish to consider (see section 9.2). Sub-group cards have no unique information, so you can print a large quantity and use them as required (or re-use them if participants do not write on them). To print these, simply open the appropriate sub-group Word document (A4 or letter) and print as many copies of the single sheet as needed.

Nested Group Cards

- 3.18 SynCaps V3 allows groups to be nested (up to 9 deep). These require the corresponding NestedGroup templates. The Open version of these templates require no data file – just print as many cards as you calculate you need for participants to provide their own group names. The Closed template requires a data file in just the same way as unnested groups, described above. Note that the group levels are not pre-coded. Participants or researchers mark (or write-in) the appropriate group level. The data capture of nested groups is described in section 5.8.

Participant Cards

- 3.19 Note that you do not necessarily need to print participant cards. You can just ask participants to write their names on the top of their results and then type the letter 'P' followed immediately by the name when you're doing the data collection. (PWilliam for example.)
- 3.20 The participant cards do not have a data file associated with them. Just open the file and print as many pages as required (there are four cards per sheet). The best way to do this is to copy just the sheets you need to a new, empty document, since Word will ignore a request to print just selected pages from the participant document since each page is in a separate section. (This always happens with mail-merge documents.)
- 3.21 Keep the item and group cards in their sheets of four as printed until you are ready to ask participants to sort them (they are easier to transport and handle this way).

- 3.22 We also recommend that you have a supply of elastic/rubber bands or something similar that you can use to keep the cards in their sorted groups. (If you are using elastic/rubber bands make sure they are fresh. India rubber does not keep terribly well and perished elastic bands simply break, often without warning.)
- 3.23 Now it's time to recruit your participants and complete the other arrangements for your card sort.

4 Sort Cards

The process described here is for perforated cards. If you are using paper or adhesive labels, you will need to adapt them to your circumstances. If you are conducting online sorts, skip to section 6, Online Sorting.

- 4.1 If you have printed adhesive labels, ask participants to apply these to index cards or similar.
- 4.2 Put the participant cards (if used) to one side for later use.
- 4.3 Give participants their group and item cards. For perforated cards, ask participants to fold the sheets horizontally and vertically and then to gently separate them. Note that it is usually possible to do this to several sheets at a time. Participants should keep group cards (printed for a closed sort or blank for an open sort) to one side at this stage.
- 4.4 Give participants their card sorting instructions. You can ask them to write queries or comments on the cards or to circle alternative terms if you wish, but you will have to analyze this information manually.
- 4.5 Do encourage participants to create new cards or groups if necessary. You can print item and group cards with no names for this purpose, or just use blank cards (you will have to do some manual data entry if you choose the latter approach).
- 4.6 If you are using SynCaps V3 for analysis you can also encourage participants to make up new items.
- 4.7 If you are using quality of fit, once the cards have been sorted into their groups, ask participants to look at each item card and mark its quality of fit in the group they selected for it: Poor, Fair or Perfect. (The process is similar for frequency of use, although you might request that participants mark frequency of use before they sort the cards, rather than after.)
- 4.8 Now ask participants to create a deck of cards with the participant card (or their name written on the top group card) at the front. Collect decks from all participants and check that they have ticked the quality of fit / frequency of use boxes (if used). Also, remove or query any cards that appear not to have been sorted or are in groups with names like 'Don't know', 'Miscellaneous', 'No idea' and so on. Once you are certain that the cards are correct and complete, carefully contain them with at least two separate elastic bands or similar.

5 Capture Data

- 5.1 You must create a simple text file for processing by SynCaps. All of the data that you want to analyze together should be in the same file. Notepad under Windows is ideally suited, but other text editors can be used as long as the output is plain text. Note that V3 adds an easy-to-use syntax editor for this purpose.
- 5.2 The SynCaps input file is conceptually in two parts. Header and Data.

Header

- 5.3 The header simply defines any prepared groups and items. This starts with a list of group names for a closed or hybrid card sort followed by the item names (for both open and closed card sorts). Each name appears on its own line with the letter 'G' ('g') or 'I' ('i') as the first character (upper or lowercase), immediately followed by the name. For example:

```
GAnimal
GVegetable
GMineral
IPeanut butter
IPhil Silvers
ITomato
IPotato
```

You can type in your header or use your GroupList.doc and ItemList.doc files as a source. In SynCaps V3, click the Groups or Items command buttons in the Data group to copy and paste these in from any source. Note that SynCaps V1 does not support 'ad hoc' items (those made up by participants during the sort). You will need to add them to the header during data capture).

Note that groups do not have to come before items although they do need to be in the order that corresponds to the printed cards (or the list order used in online sorts). Also, for users of V3, group cards in the header should not have a level indicator (L1, L2, L3 and so on) after the initial 'G' – group names can be used at any level selected by a participant.

The end of the header is marked by the presence of the first participant.

Data

- 5.4 In a closed card sort, all of the data following the header comes from scanning the cards with a barcode reader (or typing it in). Just put the cursor in Notepad (or equivalent) on the first line following the last item name and start scanning the cards. The scanned data follows on from the header without a break (although blank lines have no effect on the analysis):

```
GAnimal
GVegetable
GMineral
IPeanut butter
IPhil Silvers
ITomato
```

```

IPotato
P1 (scanned data starts here)
G2
I4F3
I3F2
I1F1
G1
I2F3

```

- 5.5 The new data is for participant 1 and starts with group 2 (Vegetable), into which is placed items 4, 3 and 1 (Potato, Tomato and Peanut Butter). Notice that Potato is a perfect fit (F3), Tomato is a fair fit (F2) and Peanut butter is a poor fit (F1). The fit data is determined by scanning the appropriate bar code on each item card.
- 5.6 In SynCaps V3 the process is identical using the inbuilt syntax editor, except that text will automatically be coloured according to its type. Items and group numbers are expanded using the header contents (see example below). Also, in V3 only, the first participant can use an 'R' instead of a 'P' (for Reference and Participant, respectively). This means you don't have to remember to check the 'First participant is expert/reference sort' option during import. Be aware that 'R' and nested groups (GL1, GL2 and so on) are not backward compatible with earlier versions of SynCaps, so if you try to process a file containing these record types, you will receive data errors from SynCaps V1.

```

27 IRed Cabbage
28 IRed Peppers
29 ISquash / Marrows
30 ISwede / Rutabaga
31 ISweet Potatoes / Yams
32 ITomatoes
33 ITurnips
34 IWatercress
35 IWhite Cabbage
36 RReference
37 GL1Pears
38 I24 {Pears}
39 GL1Exotic Fruit
40 I15 {Kiwi Fruit}
41 I19 {Lychees}
42 GL1Mushrooms & Squash
43 I20 {Mushrooms}
44 I29 {Squash / Marrows}
45 I26 {Pumpkin}
46 GL1Potatoes & Onions
47 I31 {Sweet Potatoes / Yams}
48 I21 {Onions}
49 I25 {Potatoes}

```

Figure 2, Syntax Editor in SynCaps V3

- 5.7 In Version 3 anonymous subgroups can be used to further refine relationships between items. Note that they only affect the weights applied to items during cluster analysis. So a pair of items that appear together in the same subgroup has twice the weight as a pair that appears in separate subgroups. As only the weights

are affected, there will be no apparent difference in the various analyses other than how items are clustered. For open sorts, subgroups are most effective for small numbers of participants (say 30 or fewer). In SynCaps V3, a similar effect is achieved through nested groups, although unnamed subgroups are also supported. Nest groups have the advantage of allowing up to nine levels and can be named (whereas anonymous subgroups cannot, as the term implies).

The start of a subgroup is indicated with just the letter 'S' (or 's') with no further information on the line. So, supposing that we wished to put Tomato and Potato into the same same-group (separate from Peanut Butter), the file would appear as follows:

```
GAnimal
GVegetable
GMineral
IPeanut butter
IPhil Silvers
ITomato
IPotato
P1
G2
S (first subgroup starts here)
I4F3
I3F2
S (second subgroup starts here)
I1F1
G1
I2F3
```

The first subgroup line (the 'S' immediately following G2) is optional – SynCaps will put all of the items following 'G2' into a separate subgroup when it finds the first explicit subgroup line (the second one in the example above).

- 5.8 Nested groups (SynCaps V3) add a level indicator (L1..L9) following the initial 'G' but are otherwise similar to the standard group cards. They can be numbered (as in a closed sort) or group names can be typed in during data capture. An example is shown in the Syntax Editor example on page 14. Nested groups can appear without intervening item cards, for example:

```
GL1Fruit
GL2Apples & Pears
GL3Apples
ICoxes
IRed Delicious
IEarly Windsor
GL3Pears
IAnjou
IGreen Williams
```

- 5.9 Scan all of the cards.

Note that if you would like to do an expert or reference comparison (where alignment is calculated between each participant and a reference sort), the reference participant should be the first in the file (see the Advanced Concepts

chapter starting on page 38). All other participant data should follow it. The reference data will not be included in the overall results.

- 5.10 If a participant has not marked a quality of fit value, you need to decide which barcode to scan. If you are not using the quality of fit data, it does not matter, but if you are, it might be best to consistently choose 'fair' (the middle barcode).

If any barcode refuses to scan, simply type in the missing data: 'P' followed by the participant name or number; 'G' followed by the group number or 'I' followed by the item number (participant, group and items numbers are usually in the bottom right-hand corner of the card). For item cards, add the quality of fit (if used): 'F' followed by 1, 2 or 3 for poor, fair or perfect (as shown in the last line above).

- 5.11 In an open sort, you will have group cards without barcodes and just a hand-written group name. The simplest thing to do is to just type 'G' (or 'g') on a line on its own followed immediately by the name. For example 'GFruit and veg' (without the quotes).
- 5.12 If you are using SynCaps V3, you can also create item cards this way (that is, 'I' or 'i' followed immediately by an item name). If quality of fit is used, just type it as the last two characters on the line – 'F1' for a fair fit, 'F2' for a good fit and 'F3' for a perfect fit. For example,

IRestaurant Opening TimesF3

is an ad hoc item with the name 'Restaurant Opening Times' and a perfect fit. (SynCaps V3 also allows the fit to appear immediately after the first character, so 'IF3Restaurant Opening Times').

- 5.13 If you are using nested groups in V3, type the level indicator immediately after the 'G', so 'GL3Pears' as in the example above.
- 5.14 While capturing data, save your file from time to time. A file type of '.txt' is recommended.
- 5.15 Before you finish check:
- ▶ No more than one entry on a single line
 - ▶ An adequate number of group and item names in the header.
It does not matter if some of these are never referred to, but you will get errors if the scanned card sort data uses group or item numbers that do not have names (except for open card sorts for which there should be no group names prior to the first scanned 'P' card).
 - ▶ Both of these checks are fairly easy in the V3 syntax editor. Just look for data records with '???' where the group or item name would normally appear or for suspicious indentation (normally groups would be indented under participants and items under groups – however, unrecognized lines are not indented at all).

6 Online Sorting

OptimalSort provides an option to export sorted data in SynCaps format. See the screen shot below.

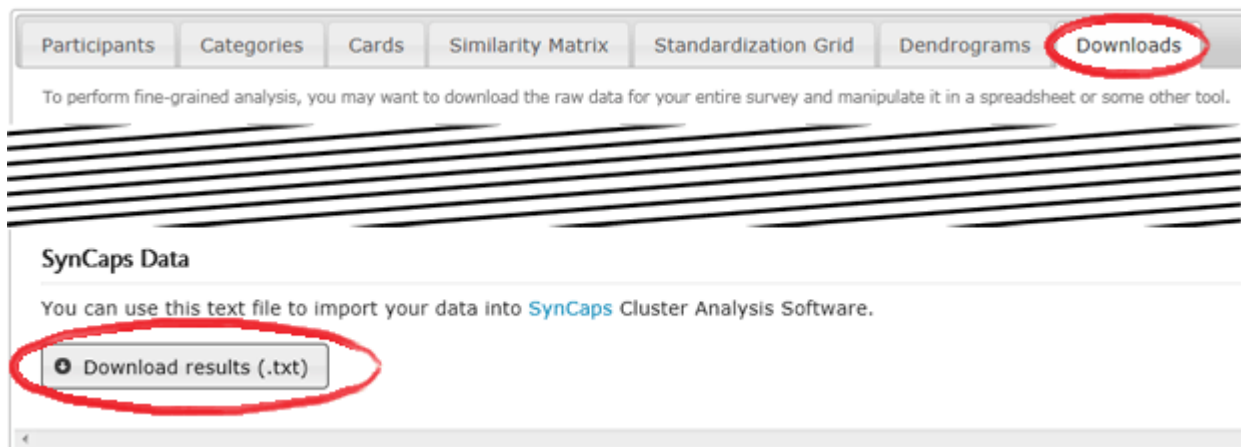


Figure 3, SynCaps Export Feature in OptimalSort

You will end up with a simple .txt file that can either be opened directly by SynCaps (just by dragging it and dropping it onto the application window or icon) or by text editors such as Windows Notepad. SynCaps V3 includes its own syntax editor which has many useful features - just choose 'Edit' when opening a file (or click 'Edit' in the Data group of the V3 Home ribbon).

7 Prepare and Analyze Data

7.1 Import and Export for SynCaps V1

Open SynCaps V1 (the process is slightly different in V3). SynCaps will automatically start the import and export process, displaying the dialog shown. Alternatively, you can drag and drop your input file onto SynCaps or any shortcut to it you may have created.

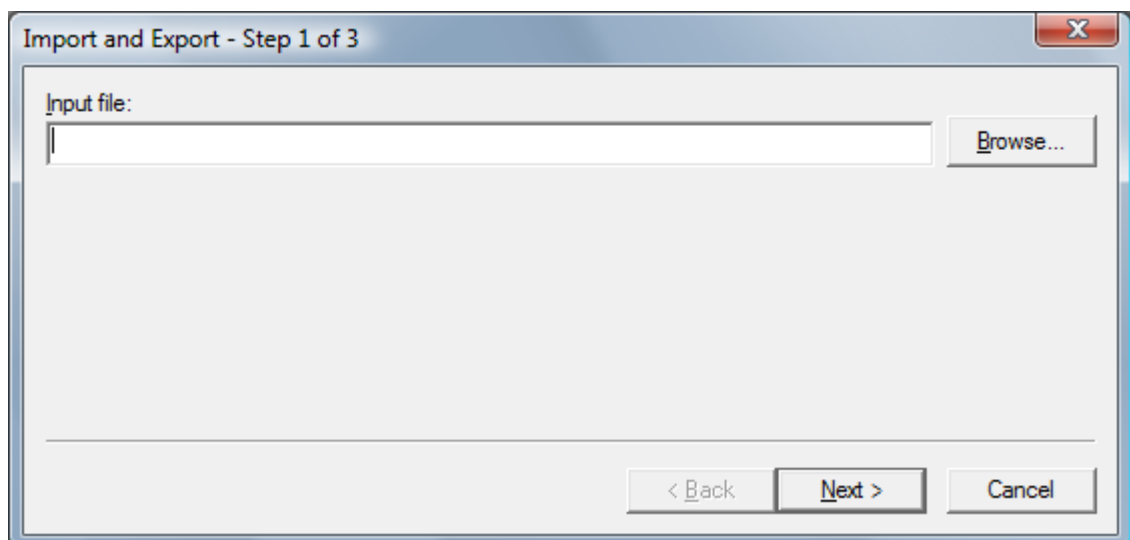


Figure 4, Import and Export Step 1

You can enter file details either by typing its full name in the input file field shown above, or by clicking the Browse button. (The recommended file type for SynCaps is .txt as this is the default for simple text editors like Notepad. However, any file type can be used in practice.)

If you would like to open an existing .caps chart file instead of starting the import and export process, just cancel this initial dialog (by clicking on the Cancel button or pressing Esc on the keyboard).

Click the Next button when you have selected a suitable import file. SynCaps will briefly check it and report any serious problems. If all is well, the step 2 dialog will be displayed (note that this is slightly different depending on which version of SynCaps you're using):

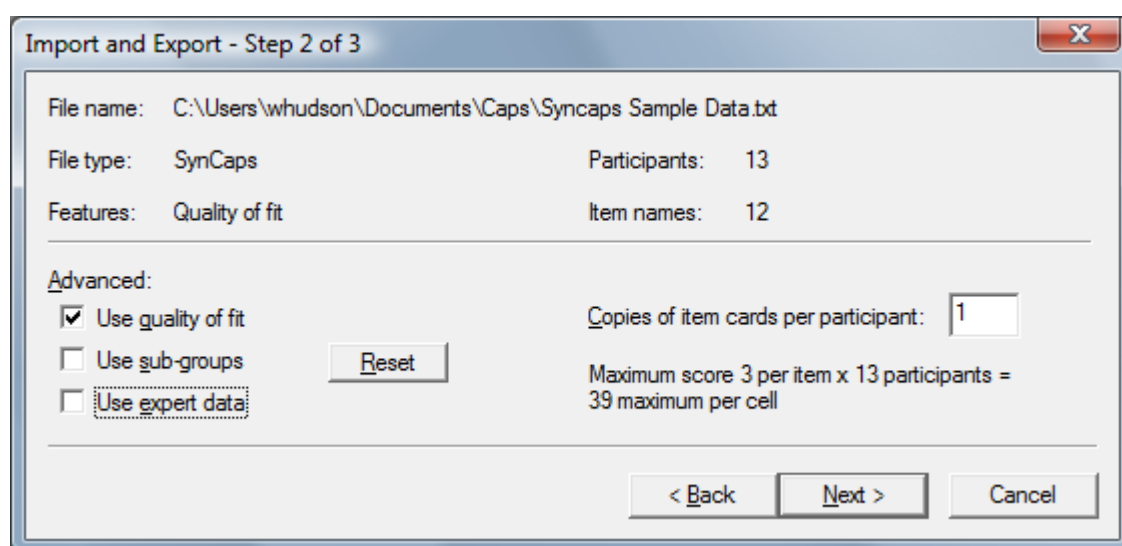


Figure 5, Import and Export Step 2 for SynCaps V1

This dialog shows the basic characteristics of the file: the type of file, the numbers of participants and the number of item names. The options that appear in the bottom of the dialog show what SynCaps features are present in the import file. They can normally be left unchanged. If you change them accidentally, click the Reset button to restore them to their default settings for your import file. (You can find out more about these check boxes in the Advanced Features section on page 38.) In SynCaps V1 unless you have supplied your participants with more than one copy of item cards each, leave the 'Copies...' field at 1.

In SynCaps V3 the 'Copies ...' field is replaced by a simple checkbox. This can be used to generate a warning if a participant has sorted the same item more than once. (SynCaps V3 allow multiple copies of cards to be sorted by default. See section 9.5 for a discussion.)

In the bottom right-hand corner of the dialog is the calculation of the maximum score that can be present in each cell of the weighted proximity matrix. This is based on the quality of fit and subgroup options and will be of interest if you plan further analysis of the weighted proximity matrix.

Check that the file information is correct and click 'Next'. (If it is not correct, click 'Back' and choose another file.) If you are using SynCaps V3 it will then

be possible to merge or ignore groups and items using two similar dialogs (only the first is shown):

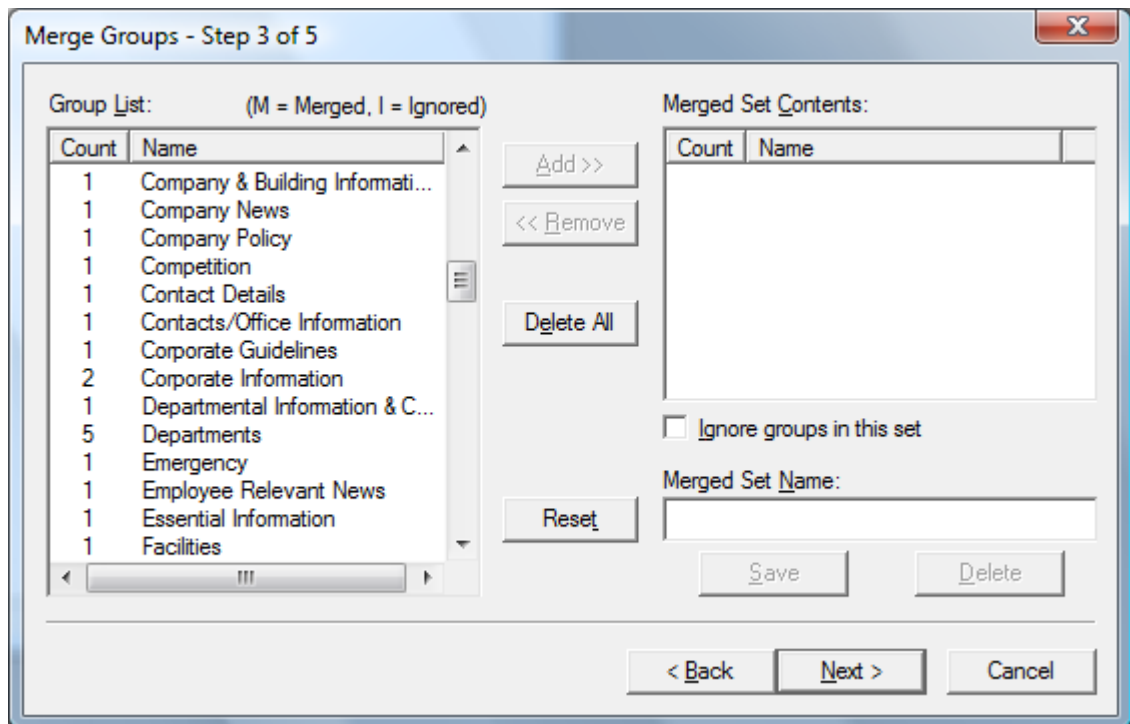


Figure 6, Merge Groups for SynCaps V3 (Merge Items is similar)

In the two merge dialogs the original groups or items appear in the left-hand panel. If you notice that two or more items are conceptually very similar you can merge³ them into a single set. Either select each name and press the Add button or hold down the Ctrl key and select multiple items, then click the Add button only once to move them all to the merged set contents. Then give the set a name (as you would like it to appear in the analyses) and click Save. If you would like the groups or items to be ignored, click the Ignore checkbox. (You might do this for groups labelled 'Don't Know', 'Miscellaneous', 'No Idea' and similar or items which have been used only rarely.)

The final dialog is the same in all versions of SynCaps, but is numbered as step three in V1 and step five in V3:

³ The merge information is stored in a text file with the same name and location as the import data file – the file type is '.mrg'. If you rename or copy the import data file you may wish to do the same with the .mrg file. You can also edit the .mrg file in a text editor but be sure to save an unedited copy just in case.

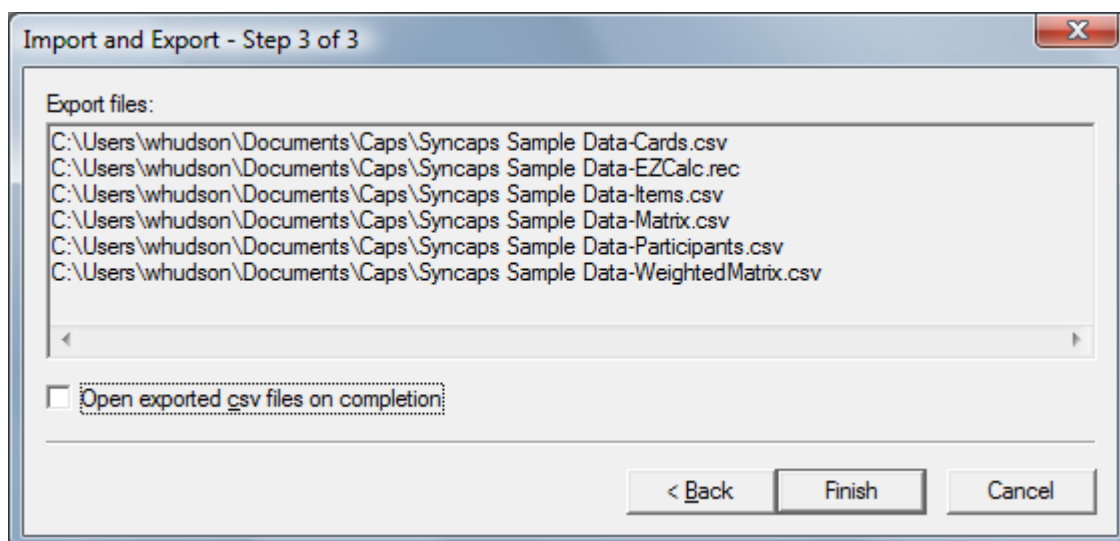


Figure 7, Export Files (Step 3 for V1; Step 5 for V3)

This shows the additional data files that will be exported. A checkbox provides the option to have the csv (comma-separated-values⁴) files opened in their associated application. This is usually Microsoft Excel. V3 also includes SPSS script files for cluster analysis, but these are not opened automatically. (They may require editing to make variable names conform to SPSS requirements. The variable names are taken from the item data and may contain special characters not acceptable to SPSS.)

7.2 Import and Export for SynCaps V3

SynCaps V3 uses a New / Open dialog to give users the choice of creating new text files or opening existing text or analysis files (.txt or .caps3, respectively). As with earlier versions of SynCaps, you can drag and drop a file onto the application window or shortcut. If it is a saved analysis (.caps3) it will open immediately without any further user interaction. (Note that V1 .caps files are not compatible with SynCaps V3. Simply start a new analysis and reimport the original data file.)

However, for .txt files you will have the choice of editing it or opening it for analysis:

⁴ The file is referred to as comma-separated, but in reality it is tab-separated since Excel processes tab-separated data more consistently.

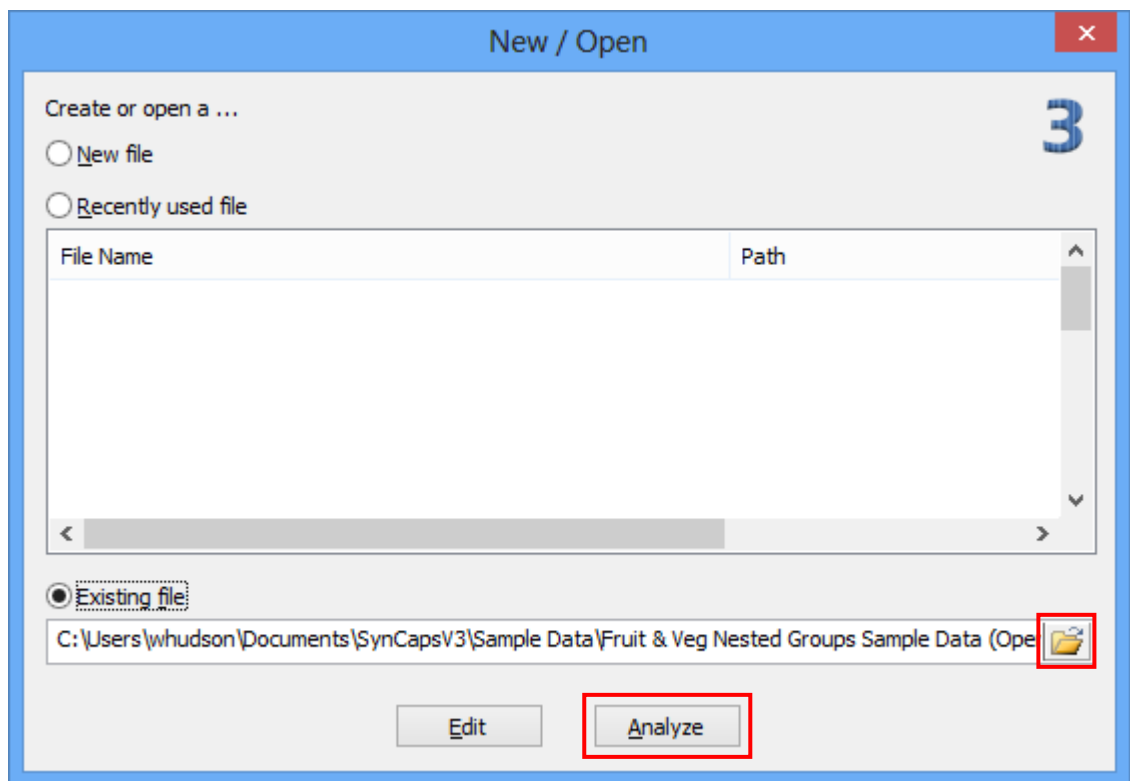


Figure 8, New / Open Dialog in SynCaps V3

You can also browse for an existing file by clicking the folder icon to the right of the Existing file edit box (marked in red).

To start the Import/Export process, select Analyze. (On the other hand, if you want to view or edit the raw data, select Edit.) The Import/Export process is very similar to V1 except that there are some additional options at step 2:

Import and Export - Step 2 of 5

File name: C:\Users\whudson\Documents\Sy...\Fruit_Veg Nested Groups Sample Data (Open - V3 Only).txt

Present: Expert/reference participant, Nested groups

Max Weight: 2 per cell Participants: 7 Item names: 35

Options:

- ☐ Quality of fit
- ☒ First participant is expert/reference sort
- ☐ Warn of repeated items within participant
- ☐ Ignore missing items (sampled sort)
- ☐ Anonymize participants in charts

Groups:

- ☐ No subgroups
- ☐ Unnamed subgroups
- ☒ Nested groups

[Reset Options & Groups](#)

Maximum Label Length:

[< Back](#) [Next >](#) [Cancel](#)

Figure 9, Import and Export Options in SynCaps V3

The first new feature for V3 is indicated by the checkmark in the '**First participant...**' box. SynCapsV3 allows the first participant to be entered with an 'R' instead of a 'P' in the data file to indicate that it is a reference sort (they can be in lowercase too). This means you don't have to remember to check this option although you can uncheck it if you wish. The other new V3 options shown in this dialog are:

Anonymize participants in charts: If participant names or email addresses are present in the data they will normally be displayed in the Participant Dendrogram chart. Select this option if you would prefer them to be numbered instead (if you are sharing the charts with a customer, for example). The actual participant details can always be found in the exported participants' spreadsheet file.

Groups: The correct option will normally be selected according to the contents of the data file. However, if you wish to directly compare the results of two sorts – one with and one without subgroups or nested groups you can change the selection here. Note that turning off nested groups when they are actually present in the data will cause warning messages about empty groups to be displayed, but these can be safely dismissed.

Maximum Label Length: This option allows you to truncate very long item or group names. However, if it is important that truncation does not take place, ensure that this figure is large enough – it is measured in characters.

Notice that the maximum weight per cell (called maximum score in SynCaps V1) has moved towards the top left-hand corner of the dialog. This piece of information is needed to understand the values present in the weighted proximity spreadsheet.

7.3 Item Dendrogram

(All versions of SynCaps produce an item dendrogram as shown below. V3 include further charts that are described in sections 7.4 through 7.7)

Once the import and export dialogs are closed, SynCaps V1 displays the cluster analysis dendrogram. It is drawn as a Windows metafile, so it can be copied and pasted or dragged and dropped into Microsoft Word, Excel or equivalent application (Copy can be found on the SynCaps Edit menu or you can use the usual Windows keyboard shortcuts, Ctrl+C for copy and Ctrl+V for paste). Changing the size of the image will cause it to be scaled appropriately, resulting in high-quality output under most circumstances.

V3 displays an Excel-style workbook of three and five chart tabs, respectively (again, as Windows metafiles for high-quality copy-and-paste). If you cannot see the chart tabs at the bottom of window, resize or maximize it.

The appearance of the charts can be altered by changing the font, font size and the two colours used. Also, in the dendrograms, the text labels can be moved horizontally by dragging the vertical grey bar to the immediate right of the dendrogram itself. For the item dendrogram the number of groups displayed initially is the average across all participants (3 in the example shown below). For the participant dendrogram (SynCaps V3 only) the default number of groups is 1. (Note that in V3 the corresponding term used is 'clusters'). For open card sorts, you can experiment with changing the number of groups to provide an optimal configuration. It is also possible to change the number of groups for a closed card sort, but unless participants have been encouraged to create new groups, altering this value does not generally provide useful results.

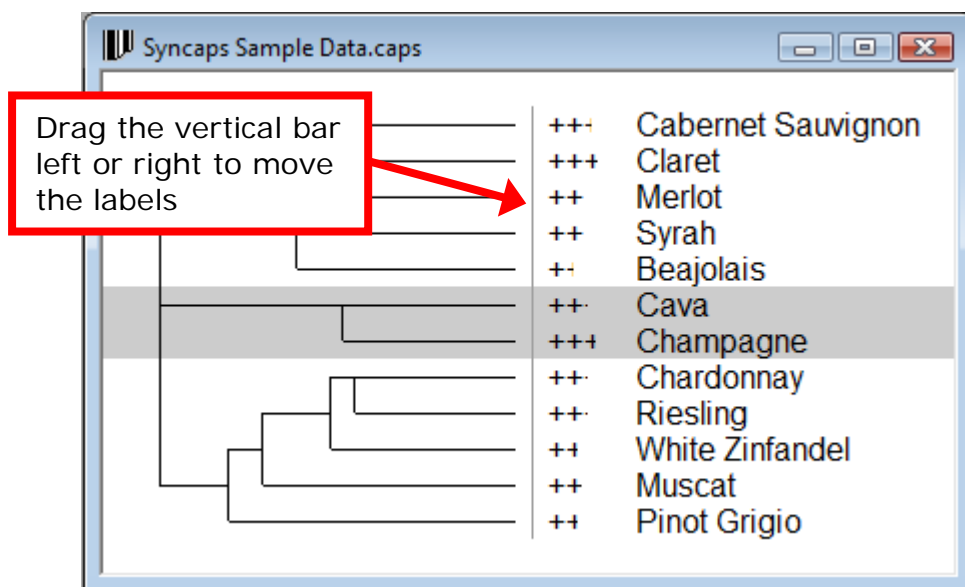


Figure 10, SynCaps Item Dendrogram (All Versions)

The vertical lines in the dendrogram itself correspond to the weighted scores calculated when each item or cluster is merged. When a pair of single items are merged (as is the case for Cabernet Sauvignon and Claret), the score will be the

one that appears in the weighted proximity matrix. In all other cases, the score is the average between across all of the items in the existing cluster and the new item being merged. (This can be calculated from the weighted proximity matrix.)

The maximum value represented by the right-most point in the dendrogram is that shown in the bottom right corner of import and export dialog step 2 (above).

The small bar chart to the left of the labels is provided when the 'Use fit data' option is selected during step 2 of import and export. It is the average quality of fit for each item.

7.4 Items x Items (V3 Only)

The Items x Items chart shows the frequency with which every possible pair of items appeared together in the same groups. The chart uses colour saturation to display this so that a white cell shows a pairing that did not occur while a dark cell shows a pairing that was made by most participants. Clicking with the mouse on a cell will show or hide its details (two cells with details displayed can be seen below). Clicking outside of the chart area will clear all detail popups. (You also need to click outside of the chart area in order to drag and drop a chart into another program – a text editor or spreadsheet, for example).

The bottom row of the Pairs Map shows relative alignment. Here, darker cells represent items that were consistently grouped while lighter cells indicate less agreement between participants.

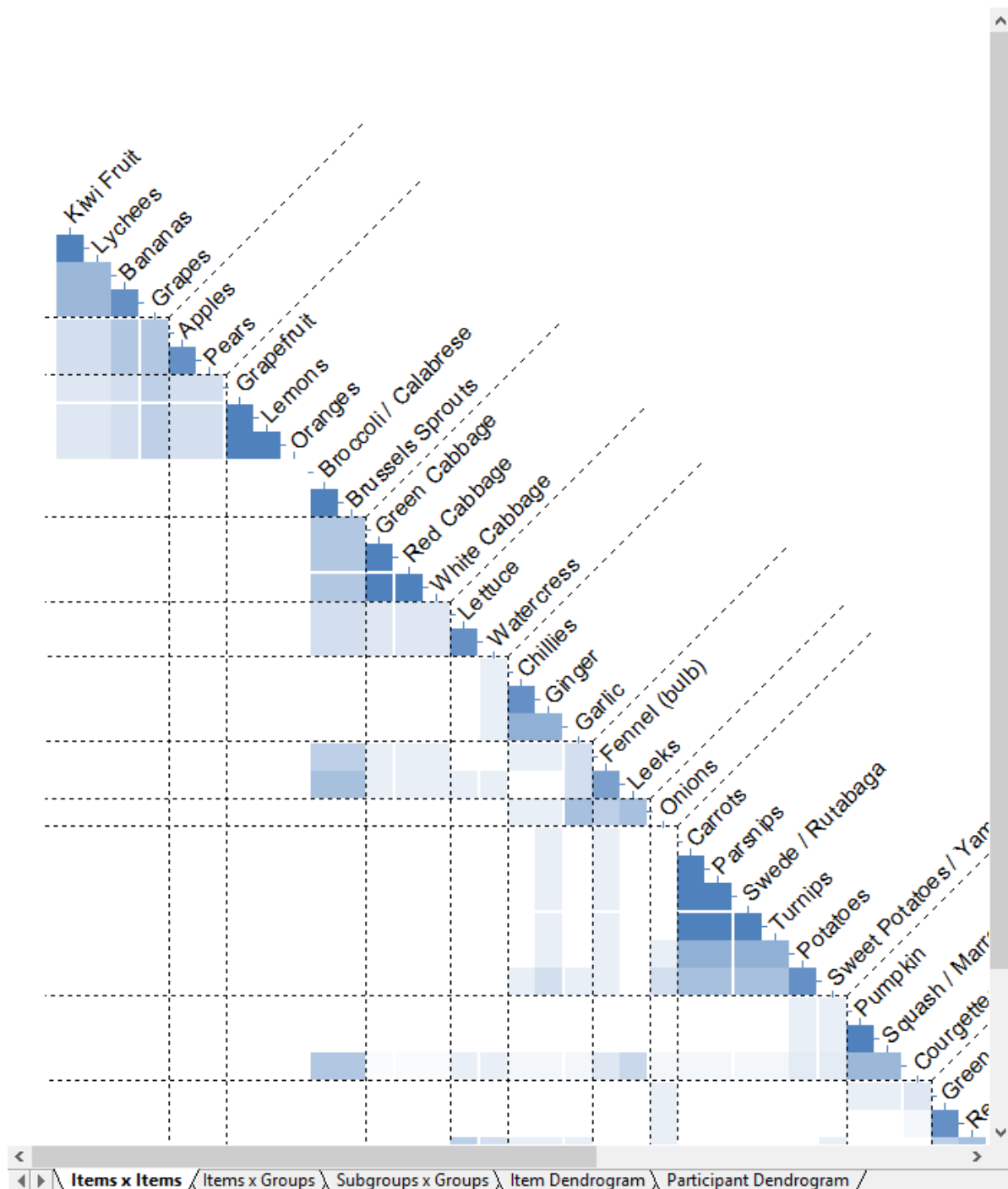


Figure 11, Items x Items Chart from SynCaps V3

Note that if you chose to use expert data in step 2 (above), the central portion of each cell will indicate the expert pairings and alignment. This is illustrated in the Items x Groups chart below.

The order of items is the same in all charts and has been determined by the cluster analysis (the dendrogram shows this more clearly).

7.5 Items x Groups (V3 Only)

This chart shows items down the left-hand side and groups across the top. The method of display is the same as for the Pairs Map / Items x Items chart, as

described above. However, here 'expert' data is also shown in the central part of each cell. In the example shown, the expert data represents the current intranet design. So, 'Communication Tools' is found in the 'Communicating' group while 'Pay and Reward' through 'Leaver Procedures' are in the 'HR' group. (Since it is based on just one participant, the expert data is typically shown as either 0% or 100%, although intermediate values can appear if the quality of fit / frequency of use varies.)

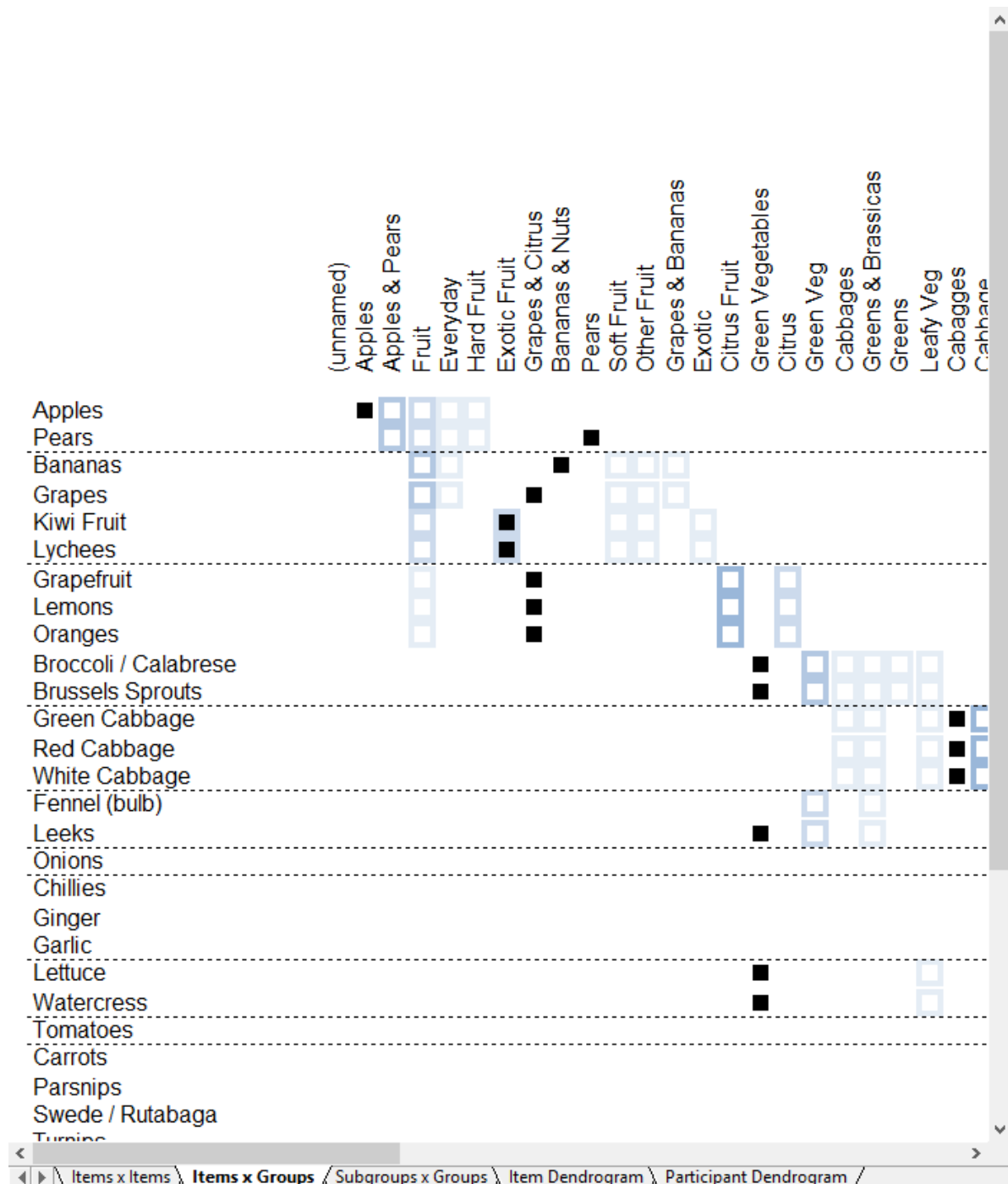


Figure 12, Items x Groups with Expert Data (SynCaps V3)

7.6 Subgroups x Groups (V3 Only)

If you are using the nested groups feature of SynCaps V3, the Subgroups x Groups chart will show how participants organized groups at each level. Note that any given group can appear multiple times. A simple example is shown below.

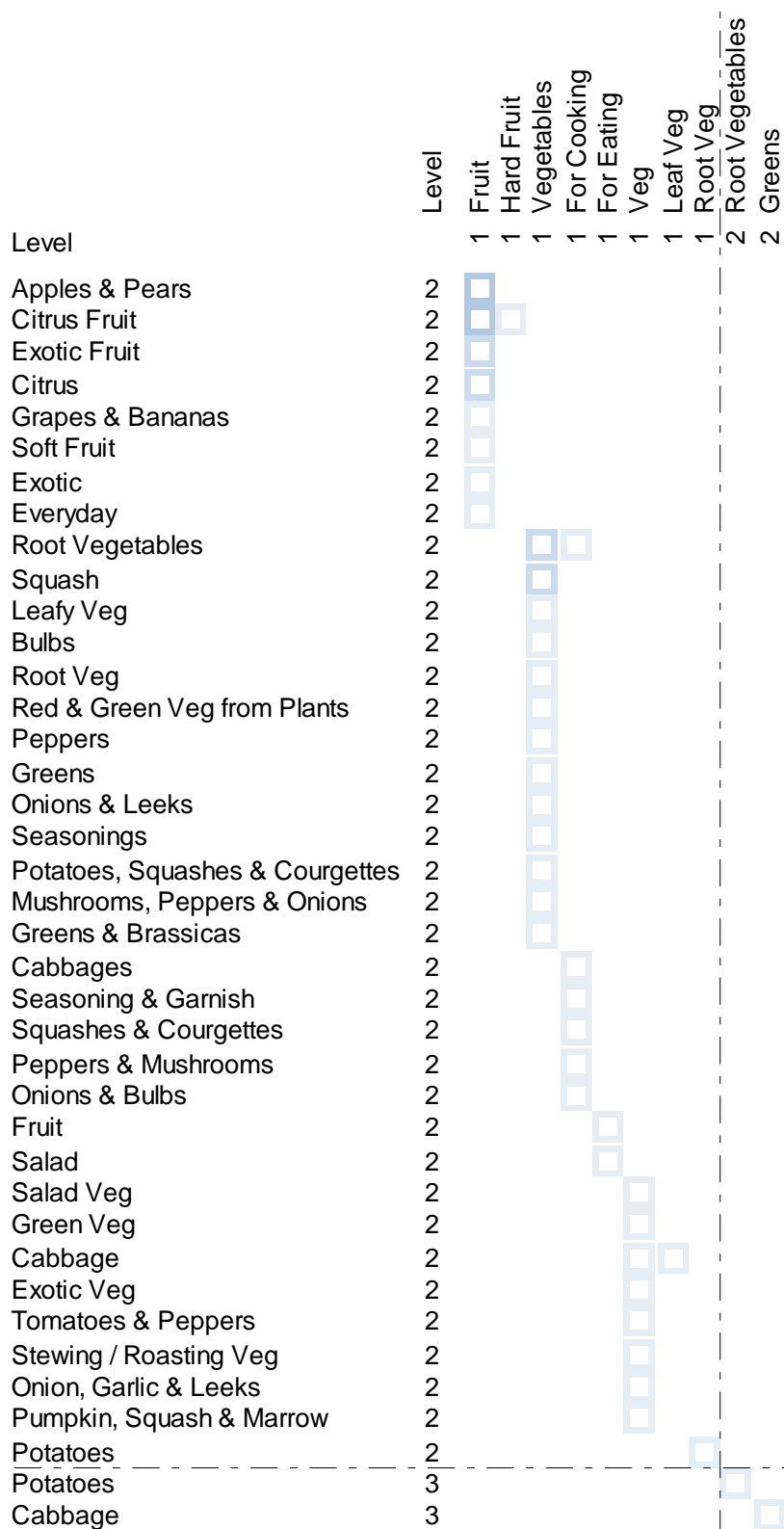


Figure 13, Subgroups x Groups Chart (SynCaps V3 Only)

In this chart the 'parent' groups are listed across the top while 'child' subgroups are listed down the side. Dashed lines separate the different levels. In this example eight subgroups are shown for 'Fruit' at levels 1 and 2, while the parent group 'Hard Fruit' has only one child called 'Citrus Fruit'. Levels 2 and 3 contain only four groups in total: 'Potatoes' were nested under 'Root Veg' while 'Cabbages' appeared within 'Greens'.

7.7 Participant Dendrogram (V3 Only)

The participant dendrogram in SynCaps V3 is conceptually identical to the items dendrogram except that participant proximity is based on how similar participants were in sorting cards. We measure this as the percentage of cards that would have to be moved to make two sort results identical. So if 100 cards were sorted and 10 cards would have to be moved between groups to make two participants' results equivalent, they would have a proximity score of 90%.

8 Analyze Additional Data

Once the import and export process has completed you can analyze the exported data using your preferred analysis software.

- **Excel** – open the .csv files in Excel (or similar) by selecting File | Open and selecting files of type 'Text Files'. You can then select the columns of interest (hold down the Ctrl key if they are not consecutive) and create graphs as required. You can also use the Data | Sort menu to arrange the results as needed (bear in mind that the charts usually put the independent variable values in reverse order). Further details of useful Excel analyses are given below.

8.1 Excel Analyses

There are several extremely helpful charts that can be produced in Microsoft Excel or a similar package. These are described for each output data file:

8.1.1 Cards File (-Cards.csv)

Fields:

- ▶ **Participant Name** – shown in column 1 as it appears in the input file. Unless names are typed in during data entry, it is usually of the form 'Pn' where n is an arbitrary number.
- ▶ **Group Name** – shown in column 2. For closed card sorts with number group cards, providing the group names at the top of the input file will allow the actual names to be shown, otherwise just the group number will appear. Unnumbered group cards appear as a series of five dashes ('-----') if the group name is not typed in during data entry.
- ▶ **Item Name** – shown in column 3. As items are always numbered, the corresponding names will appear. Subgroup cards ('S') appear as a series of three dashes ('---').

8.1.2 Item File (-Items.csv)

Fields:

- ▶ **Item No** – this field shows the position of the item in the input data file header. As the item file is output in dendrogram order, the original item number order can be obtained by using Excel (or similar) to sort by item number.
- ▶ **Item Name** – as it appeared in the input file.
- ▶ **Count** – the number of times this item appeared in the sort results. If you instructed participants to omit cards they are uncertain of, any omissions will be reflected in this value. If you are expecting all items to be present for all participants, any discrepancies (caused by missing cards, for example) can be found here.
- ▶ **Pairs** – How many other items this item was paired with. This number will be larger when the item appears in large groups and smaller for small groups. It is the sum of each row (or column) in the unweighted proximity matrix. Note that the total number of pairs shown here is twice the number that appears in the participant file since each pair is counted twice – once for each item in the pair.
- ▶ **Weighted** – whereas the **Count** field shows the number of times an item appeared, the **Weighted** field is the sum of the quality of fit and subgroup weights for each item (quality of fit ranges in value from 1 to 3; this is multiplied by a subgroup value of 1 or 2 depending on whether an item appeared in the main group or subgroup).
- ▶ **AveFit** – the average quality of fit for each item based on the number of participants. Note that this will be the same as **Weighted/Count** when **Count** equals the number of participants (that is, when no items are missing).
- ▶ **AveDev (not shown but used in calculation of %AveAlign)** – average deviation. This measure is calculated as the sum of the squares of the difference of each cell between each participant's proximity matrix and a proximity matrix averaged across all participants:

$$\text{AveDev} = \sqrt{\frac{\sum (P_{ij} - A_{ij})^2}{N_{sets} N_{pairs}}}$$

where P_{ij} represents each unique cell of the participant's proximity matrix and A_{ij} , the corresponding cell from the averaged matrix. (Note that 'unique' means that only half of the matrix is considered since the entry for item pair {3,4} is identical to that for {4,3}.) N_{sets} is the product of the number of participants and the number of card sets used by each (the latter is 1 in most cases). N_{pairs} is the sum of $N_{items} (N_{items} - 1) / 2$ for each group created by a participant.

The average and expert deviation figures show how much a participant's results vary from the average or expert results, respectively. Larger values show greater variance.

- ▶ **%AveAlign** – average deviation expressed as a percentage of the maximum theoretical deviation (the maximum score that can appear in each cell of the proximity matrix).
- ▶ **ExpertDev (not shown but used in calculation of %ExpertAlign)** – ‘expert’ deviation. This is calculated as described above, but using a proximity matrix for an ‘expert’ participant instead of an averaged matrix. This may be the sort results for a real expert or could correspond to the organization of an existing topology (web site hierarchy for example). This field only appears in the participants file when the ‘Use expert data’ option is selected during import and export. (See the Advanced Concepts chapter starting on page 38.)

The formula is:

$$\text{ExpertDev} = \sqrt{\frac{\sum (P_{ij} - E_{ij})^2}{N_{sets} N_{pairs}}}$$

where E_{ij} represents the proximity matrix entries for the ‘expert’ sort results. (The ‘expert’ is the first set of results in the input data file.)

- ▶ **%ExpertAlign** – expert deviation expressed as a percentage of the maximum theoretical deviation (the maximum score that can appear in each cell of the proximity matrix).

An ‘average fit’ bar chart shows in greater detail the average fit values entered on each item card. This is a large version of the small bar chart that appears in the body of the SynCaps dendrogram (to the immediate left of the item names).

To create it:

- 1 Open the csv file in Excel.
- 2 Hold down the Ctrl key and click on the column heading above ‘Item Name’ in the spreadsheet. This will select the item name column.
- 3 While still holding the Ctrl key down, also click on the column heading above ‘AveFit’.
- 4 Click on the Insert → Chart menu item, select a chart type (‘Bar’ is quite suitable) and then click ‘Finish’.

This will place the chart on the same page as the spreadsheet. If you would like axis titles and other options, the steps vary according to what version of Excel you are using:

- For Excel 2003 and earlier, use the chart wizard to adjust features and right click on chart components to modify them further.
- For Excel 2007 use the Chart Tools ribbon tabs Design, Layout and Format

In most cases if there are components of the chart that you do not want (such as the title or key) just click on them and press the 'Del' key on the keyboard. A quality of fit bar chart typically looks like this:

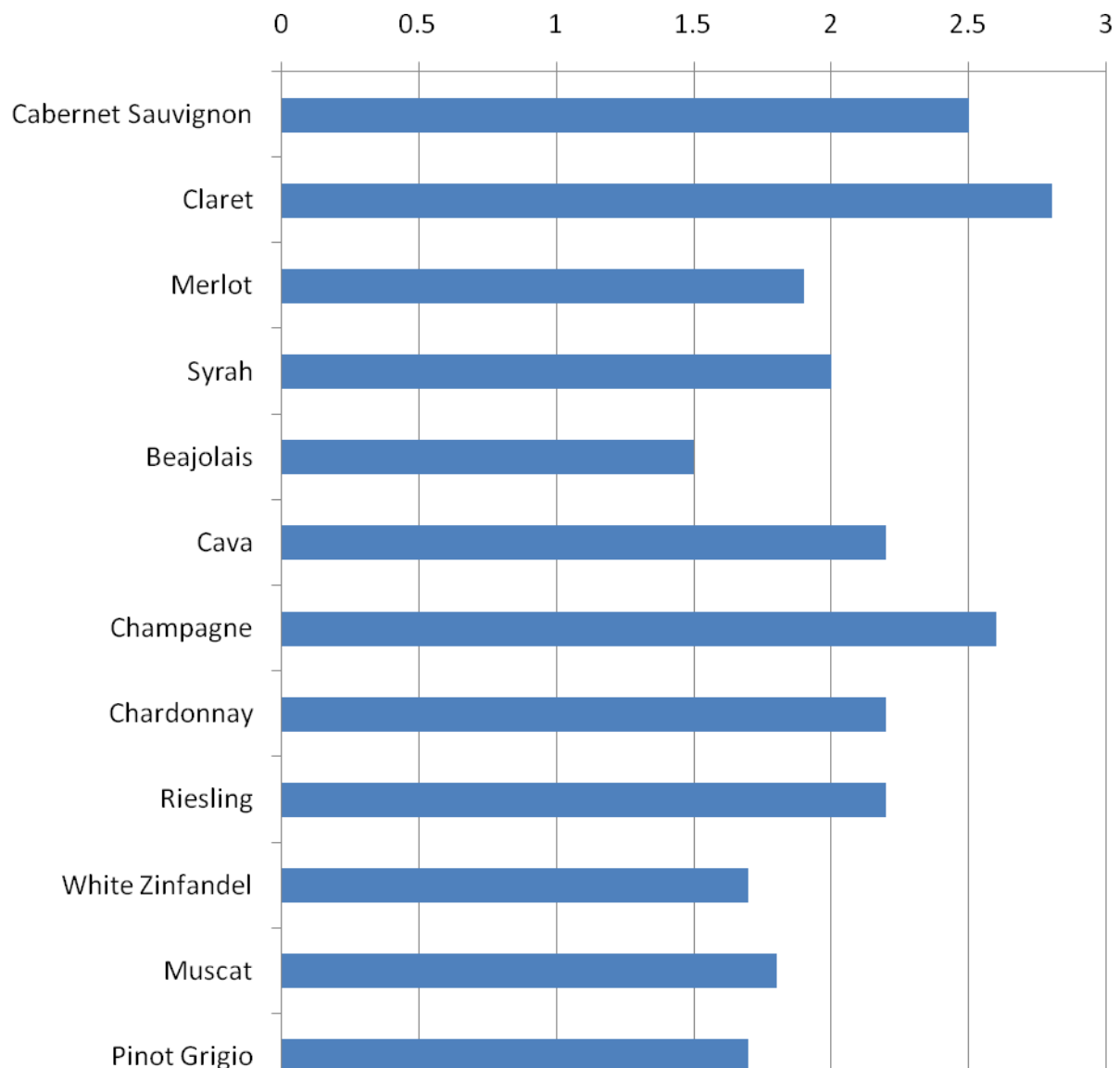


Figure 14, Average Quality of Fit by Item

An 'alignment' bar chart shows agreement between participants versus agreement with a reference ('expert') dataset. Agreement between participants is shown by the %AveAlign measure while agreement with the expert data set is shown by %ExpertAlign. In the example shown below, items have been sorted into descending order by %AveAlign (see below).

To create it:

- 5 Open the csv file in Excel.
- 6 Hold down the Ctrl key and click on the column heading above 'Item Name' in the spreadsheet. This will select the item name column.

- 7 While still holding the Ctrl key down, also click on the column heading above '%AveAlign ' and '%ExpertAlign'.
- 8 Click on the Insert → Chart menu item, select a chart type ('Bar' is quite suitable) and then click 'Finish'.

This will place the chart on the same page as the spreadsheet. If you would like axis titles and other options, go through each stage of the wizard or use the Chart → Options menu. If there are components of the chart that you do not want (such as the title or grey background) just click on them and press the 'Del' key on the keyboard. Note that Excel is fond of putting items in charts like this in reverse order. Use the axis options to correct this.

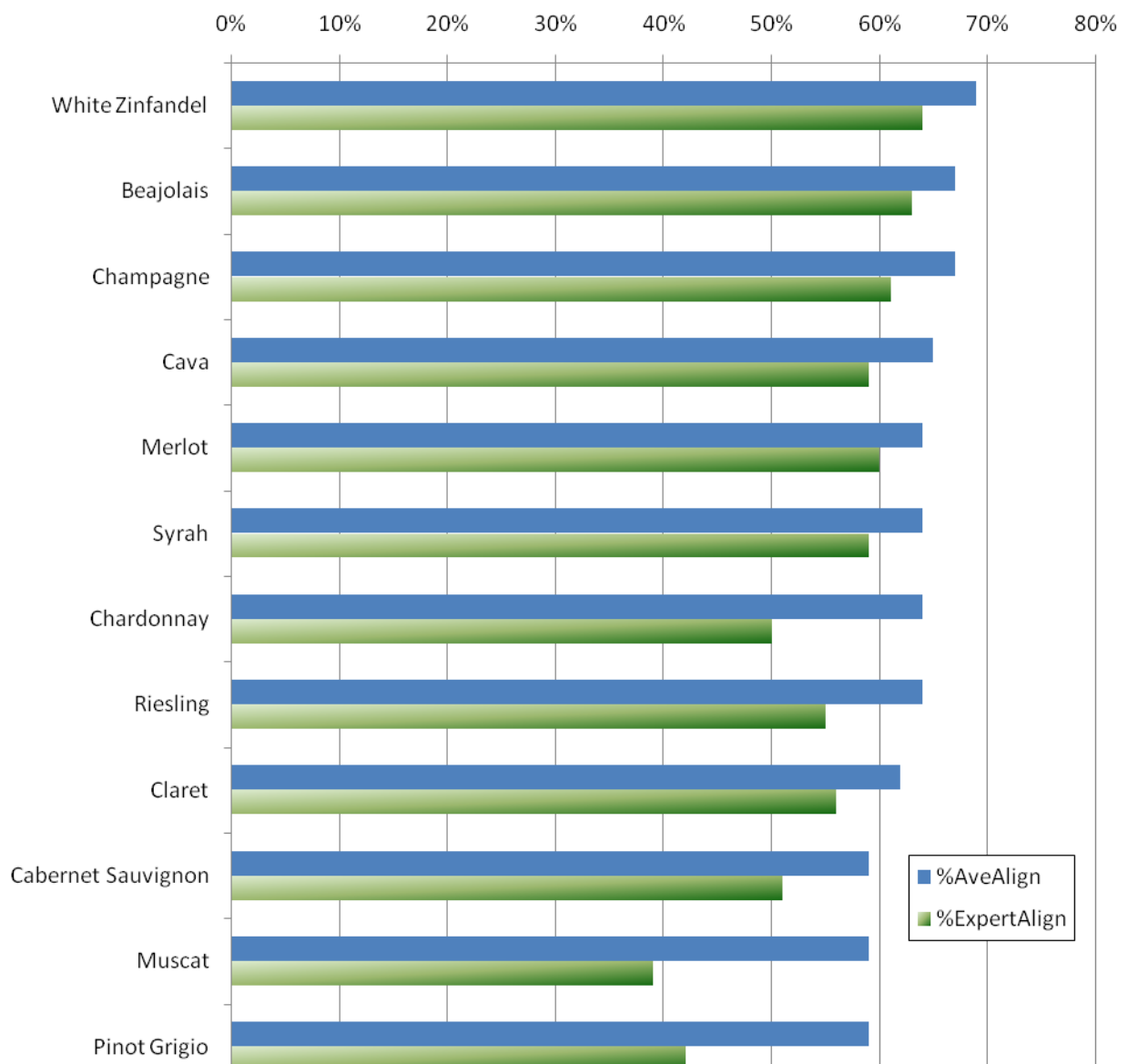


Figure 15, %AveAlign and %ExpertAlign by Item

note that in this example participants show a fair amount of agreement for Muscat, but agreement with the expert data set is substantially lower (this is also true for Pinot Grigio). Consequently, these items would warrant further investigation.

8.1.3 Participants File (-Participants.csv)

Fields:

- ▶ **ID** – the participant ID (number or name) from the input file.
- ▶ **Groups** – the number of groups used by this participant.
- ▶ **Items** – the number of items sorted by this participant.
- ▶ **Pairs** – How many pairs of items were created for this participant. This number will be smaller if a participant creates small groups and larger for large groups. It is the count of the number of unique entries in this participant's proximity matrix (each item appears twice in the matrix, but it is counted only once for this field). For example, if a participant creates only one group with two items in it, that will produce one pair. Three items will result in three pairs and four items creates six pairs. (If the items were labelled A, B, C and D, the four pairs would be AB, AC, AD, BC, BD and CD.)
- ▶ **AveSize** – Items/Groups
- ▶ **AveFit** – quality of fit averaged across all items for this participant. It can be useful in deciding whether a participant was in any way extreme in their levels of confidence.
- ▶ **AveDev** – average deviation. This measure is calculated as the sum of the squares of the difference of each cell in the participant's proximity matrix and a proximity matrix averaged across all participants:

$$\text{AveDev} = \sqrt{\frac{\sum (P_{ij} - A_{ij})^2}{N_{pairs}}}$$

where P_{ij} represents each unique cell of the participant's proximity matrix and A_{ij} , the corresponding cell from the averaged matrix. (Note that 'unique' means that only half of the matrix is considered since the entry for item pair {3,4} is identical to that for {4,3}.) N_{pairs} is the sum of $N_{items} (N_{items} - 1) / 2$ for each group created by a participant.

The average and expert deviation figures show how much a participant's results vary from the average or expert results, respectively. Larger values show greater variance.

- ▶ **%AveAlign** – average deviation expressed as a percentage of the maximum theoretical deviation (the maximum score that can appear in each cell of the proximity matrix).
- ▶ **ExpertDev** – 'expert' deviation. This is calculated as described above, but using a proximity matrix for an 'expert' participant instead of an averaged matrix. This may be the sort results for a real expert or could correspond to the organization of an existing topology (web site hierarchy for example). This field only appears in the participants file when the 'Use expert data' option is selected during

import and export. (See the Advanced Concepts chapter starting on page 38.)

The formula is:

$$\text{ExpertDev} = \sqrt{\frac{\sum (P_{ij} - E_{ij})^2}{N_{pairs}}}$$

where E_{ij} represents the proximity matrix entries for the 'expert' sort results. (The 'expert' is the first set of results in the input data file.)

- ▶ **%ExpertAlign** – expert deviation expressed as a percentage of the maximum theoretical deviation (the maximum score that can appear in each cell of the proximity matrix).

A scattergram can be very helpful in discovering participants whose sort results are unusual, indicating that they did not understand the problem, saw it from a very different perspective from others or perhaps were not adequately motivated.

To create a scattergram:

- 1 Open the csv file in Excel.
- 2 Choose two columns of interest by holding down the Ctrl key and clicking their column headings. (Consider for example, 'AveFit' and 'AveAlign' or 'AveAlign' and 'ExpertAlign'.)
- 3 Select Insert → Chart from the Excel menus and choose 'XY (Scatter)'.
- 4 Be sure to label the X and Y axis so that you do not have to remember which is which. The X axis will be the left-most column selected in the spreadsheet.
- 5 Click 'Finish'

When the scattergram is displayed (example below), you can find out which participant a marker corresponds to by putting your mouse over it.

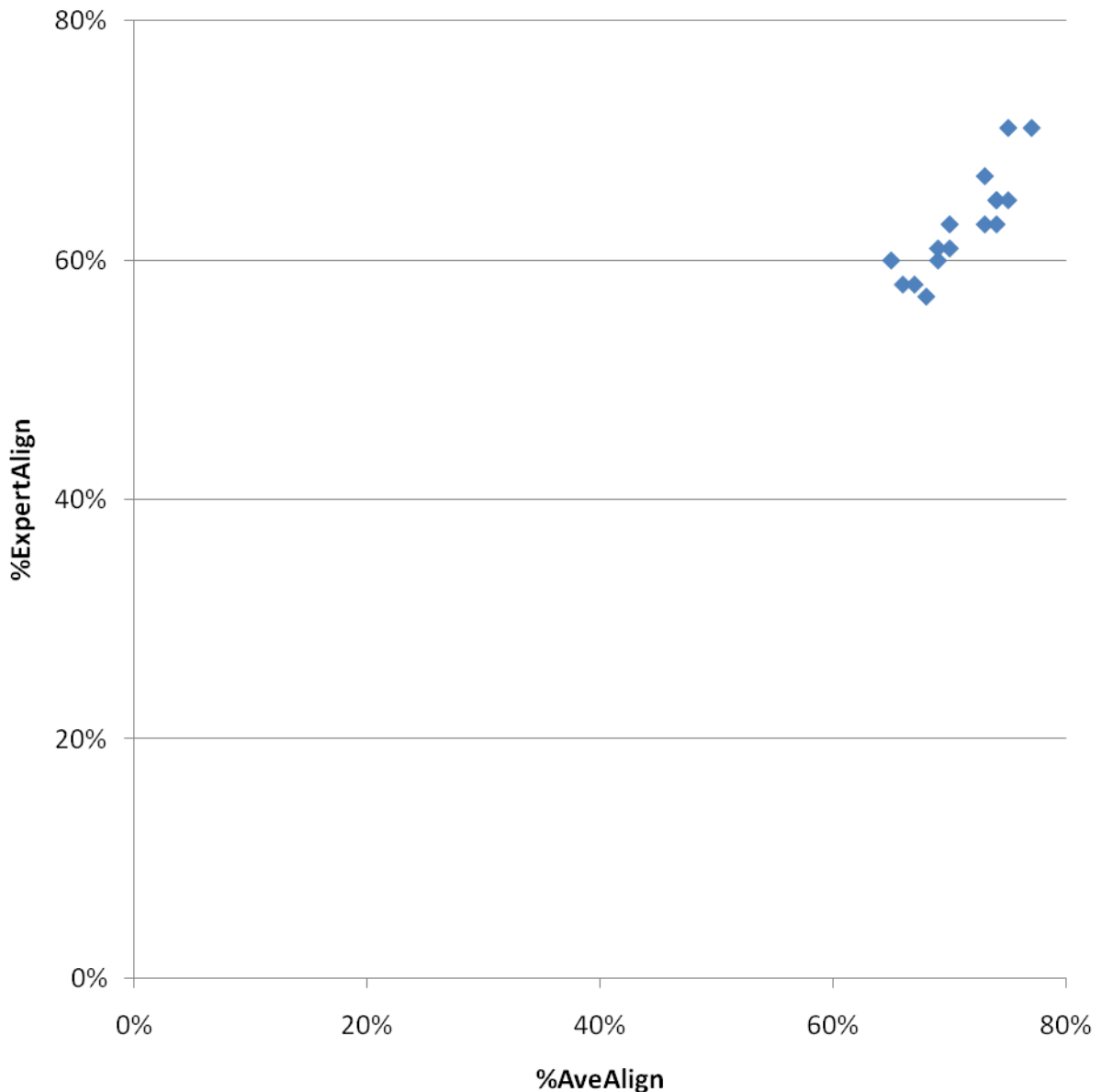


Figure 16, Scattergram of Participants

The example above shows no outliers, but a pretty healthy range of values between 60 and 80%. Note that %AveAlign and %ExpertAlign are fairly closely correlated as it is not possible to have a high expert figure without a high average figure. However, it is those participants low average figures that are worthy of further attention.

8.1.4 Proximity Matrix (-WeightedMatrix.csv)

The proximity matrix represents the raw data for the cluster analysis. In the unweighted version (-Matrix.csv) each cell contains the count of the number of times a given pair of items was grouped together by participants. The weighted version is similar in concept, but the combined quality of fit and subgroup weightings are taken into account as follows:

$$WeightedScore_{ij} = \sum_P \left(\frac{F_i + F_j}{2} \times S_{ij} \right)$$

where F_i and F_j are the quality of fit measures (1 to 3) for the two items in question and S_{ij} is the subgroup value. What this shows is that if two items appear in the same group for a given participant P , their quality of fit values are averaged together and multiplied by their subgroup value. The subgroup value S_{ij} is defined as follows for any given item pair:

- 1 - Items appear in the same group but in different subgroups.
- 2 - Items appear in the same group and subgroup

(For further explanation of how subgroups work, see the Advanced Concepts chapter starting on page 38.) If either feature is unused, its term in the above expression is set to 1.

It is relatively simple to create a surface map from the weighted proximity matrix (an example is shown below). This can be much more revealing than a dendrogram for items that have not been grouped consistently by participants. Note though, that in SynCaps V3, the Items x Items chart shows this in a much more compact form.

These instructions are for Excel 2003 (and earlier). Surface maps in Excel 2007 and later are somewhat less useful due to a change in formatting by Microsoft. See section 8.1.5 for an alternative method.

- 1 Open the csv file in Excel.
- 2 Select the whole matrix starting with the (untitled) second column.
- 3 Go to the Insert → Chart menu item. When the dialog appears choose 'Surface'. In the chart sub-type field, choose the 2D colour surface chart in the bottom left corner.
- 4 Click 'Finish'.
- 5 On the chart itself, first double-click one of the Y-axis labels. This will show a Format Axis dialog box. Go to the Scale tab and change both numeric fields to '1'. Also check the 'Series in reverse order' box at the bottom and click 'OK'.
- 6 Repeat this for the X axis but **do not** check 'Series in reverse order'.

The surface chart that follows shows three distinct groups, but the bottom group has two odd outcrops. The larger is because of Pinot Grigio which participants seem to have confused with Pinot Noir (Pinot Grigio is white but Pinot Noir is red). Similarly, a few participants have grouped Muscat with the sparkling wines. In this case they may be thinking of 'Moscato' which is an inexpensive sparkling wine.

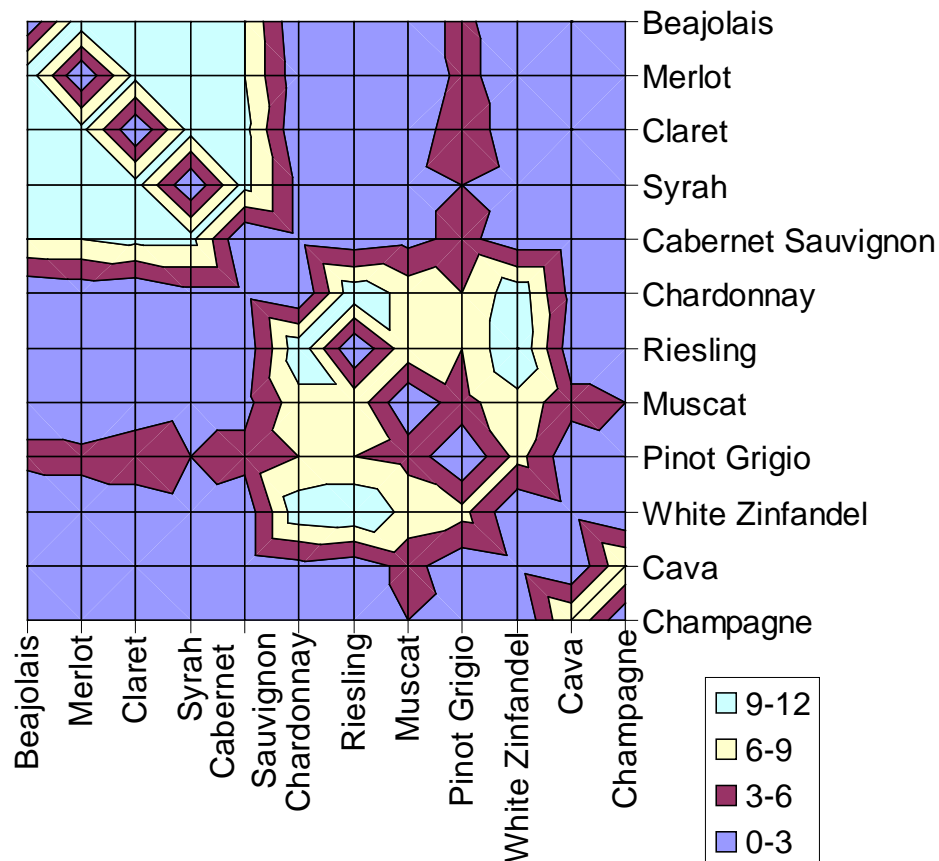


Figure 17, Surface map showing three distinct groups and some confusion around 'Muscat' and 'Pinot Grigio' (Excel 2003)

8.1.5 Conditional Formatting as an Alternative to a Surface Map

In more recent version of Excel (2007 & 2010), the surface map is visually over-complex. Simple colour-coding of cells is much more effective and is described below. It is also procedurally simpler than the surface map.

- 1 Open the csv file in Excel.
- 2 Select the whole matrix starting with the (untitled) second column.
- 3 Go to the Home → Styles → Conditional Formatting ribbon item. In the resulting drop-down menu, choose 'Color Scales'. The choice of colour scales is yours, but monochrome scales (for example, the Green to White choice offered) are recommended. (The actual colours can be adjusted.) The results will be very similar to the SynCaps Items x Items chart in V3).

9 Advanced Concepts

Basic card sorting simply takes note of which items participants put in the same groups. The resulting proximity matrix contains counts of the number of times each possible pair of items appeared together.

However, this 'all or nothing' approach (for any given participant two items are either in the same group or they are not) does not allow participants to directly express any degree of uncertainty or dissatisfaction in the groupings. Even if participants are encouraged to omit cards which they feel do not belong in any group, some will view this as a sign of failure and will group all cards regardless.

There are several techniques that can be used to better reflect participant's intentions. These are discussed below along with notes on how to use each with SynCaps.

The use of these concepts is controlled from the second import and export dialog, described in section 6 and also shown below. The quality of fit and subgroup check boxes are initially set to reflect what advanced features have been found in the data, so they can be left alone in most cases.

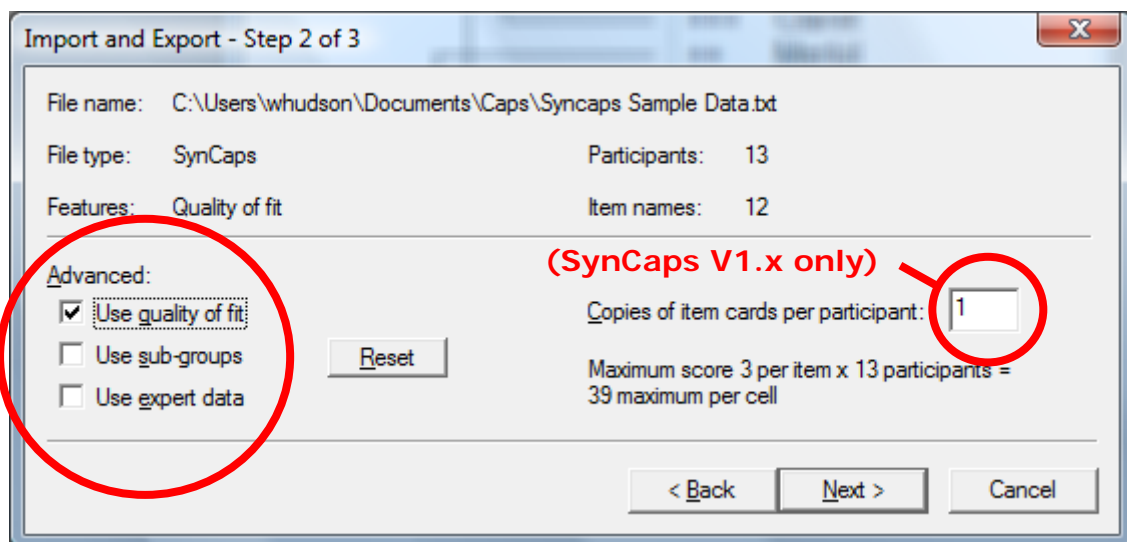


Figure 18, Import and Export Step 2 (SynCaps V1 - V3 is Similar)

Note that while it is possible to use these techniques in combination we strongly recommend that you choose only one for any given session unless you are confident that your participants will not feel stressed or confused. If in doubt, give them a simple sort to perform as a trial.

9.1 Quality of fit

This technique is unique to Caps. It provides many useful measures, allowing analysis by item and participant. The basic principle is that participants are allowed to say how well each item fits its group on a scale of 1 to 3 (1 = fair, 2 = good, 3 = perfect).

The Caps item templates assume that you will be using quality of fit.

SynCaps will automatically use the quality of fit data if it is present in an input file.

If the 'Use quality of fit' option is checked, this shows that quality of fit data is present in the import file. You may uncheck this box if you would like the fit data to be ignored. If the box is not checked, you may check it. This will result in all items being treated as if marked with a 'perfect' quality of fit if the data is not present in the import file ('3' is the numeric value). Consider doing this if you have files with and without quality of fit data and want to compare them on an equal basis. (For example, online sorting sites never contain quality of fit data so checking 'Use quality of fit' would allow dendrograms to be prepared for comparison with SynCaps data files which normally do include it.)

9.2 Subgroups

If participants feel that a group should be further sub-divided, they can insert a subgroup card. Items that appear in the same subgroup are treated as having a stronger relationship than those that are in the same group but different subgroups. The stronger relationships are given a value of 2 in the calculations while 1 is used in the weaker cases. Consider the following example:

Group
Item 2
Item 4
Item 5
Subgroup
Item 3
Item 6
Item 7
Subgroup
Item 1
Item 8

Item pairs {2,4}, {2,5} and {4,5} are treated as being in the same subgroup so each of their entries in the proximity matrix will be assigned the value 2. The same is also true for item pairs {3,6}, {3,7}, {6,7} and {1,8}. However, the item pair {2,3} is assigned a value of 1 since the items are in the same containing group, but different subgroups {as are all the remaining pairings}.

The use of subgroups is controlled from the second import and export dialog (shown above). If the 'Use subgroups' option is checked initially, this shows that subgroup data is present in the import file. You may uncheck this box if you would like the subgroup data to be ignored. If the box is not checked, you may check it. This will result in all items being given a value of 2 when they occur in the same group.

If subgroups are used in combination with quality of fit, the two scores are multiplied.

The reasons for checking the 'Use subgroup' box even if your import file does not contain subgroup data are the same as for quality of fit (see above).

9.3 Nested Groups

SynCaps V3 allows groups to be nested up to 9 levels deep. They are described in section 5.8 and a Subgroups x Groups chart is shown in section 7.6. The concept is similar to unnamed subgroups (described above) but the groups can be named at all levels. A Subgroups x Groups chart shows how the nesting was done by participants.

Nested groups are described in more detail in the Release Notes for SynCaps V3.

9.4 Expert /reference data

This feature (among several others) is unique to SynCaps. When the 'Use expert/reference data' option is selected, the first participant results are compared with all other participants to produce the expert deviation field ('ExpertDev') in the participants data file. The 'expert' is not included as a participant in the analyses.

9.5 Copies of item cards

In **SynCaps V1** this technique allows you to give each participant more than one set of item cards to sort. So, for example, you could print 2 sets of cards per participant and ask them to put both copies in one group if that is the only one they feel is appropriate or one in another group if it belongs in more than one place in their view. Beware that the traditional cluster analysis dendrogram does not give the full picture when sorting with multiple copies. The surface map described in section 8.1.4, coloured proximity matrix described in section 8.1.5, a SynCaps V3 Items x Items chart would be a better choice.

The number of copies of the item cards provided to each participant should be entered in the field shown on the second import and export dialog.

SynCaps V3 allows a similar, but more flexible approach. Instead of printing multiple sets of cards for each participant, you can allow any cards to be repeated within a sort. These can either be pre-printed or simply hand written (as long as you enter the item name exactly as it appears on the original cards). SynCaps V3 will then simply split the grouping of an item as appropriate. So if a participant puts the same item in two different groups, it will be counted at half value in each; if it is found in three groups, each will receive a third and so on. This is the default behaviour and requires no pre-planning or special processing options.

If you have any questions or comments, you can email me directly, but it would benefit the larger community of card sorters if you joined the Caps discussion list and posted queries there. To join, visit

<http://www.syntagm.co.uk/design/disclists.htm#caps>

9.6 Participant Similarity Filters

SynCaps V3 introduces the ability to exclude outlying participants from an analysis. This is described in detail in the Release Notes for SynCaps V3 document but the concept is very straightforward. Each participant has a maximum similarity score calculated across all other participants. So, a participant that sorts cards in exactly the same way as at least one other participant, will have a score of 100% while a participant who sorts their cards very differently from all other participants will have a very low score. By selecting an appropriate minimum score, outlying participants can be omitted. This usually produces sort results with greater clarity and fewer exotic group headings. Bear in mind, though, that outliers should not be dismissed out of hand. Make sure that you understand the motives of participants who sort differently from others. (Often the motives are innocuous – too little time, too little engagement with or understanding of the activity – but this is not always the case.)

9.7 Participant Cluster Filters

In card sorting we use cluster analysis to decide which items are most frequently grouped together by participants. But we can also use the way in which cards have been sorted to decide whether there is more than one significant community of participants. Along with participant similarity, SynCaps V3 introduces participant cluster filtering. This uses cluster analysis on the participant proximity matrix outlined in section 9.6. Automatic cluster calculation will determine whether there are any significant participant clusters and allow you to view the results on a per-cluster basis. They can be used in combination with participant similarity filters although you need to bear in mind that the latter will reduce the number of participants for analysis and in so doing may reduce the viability of a cluster analysis for small numbers of participants (say fewer than 30).

Details of the process are provided in the Release Notes for SynCaps V3.

William Hudson

william.hudson@syntagm.co.uk