Appendix A

Step-by-Step Guide to Creating the Pairwise Data Structure

To create the Pairwise Data Structure, the data sets with individual difference variables and dyadic variables will need to be merged. After this, the resulting data set will need to be restructured so that each record contains the dyadic variables for Person A (the rater) interacting with Person B (the target) as well as the individual difference variables for Persons A and B. Identification variables will also be needed on each record to be able to properly specify the random effects portion of the model. These include the group to which the rater and target belong (i.e., the speed-dating session; GroupID), the subgroup to which the rater and target belong (i.e., male rater with female target vs. female rater with male target; mf and fm, respectively), and the specific combination of rater and target within a particular group (e.g., male rater 1 with female target 2; MaleID and FemaleID). We first discuss the individual difference data set as this will be the most familiar to readers. We then proceed to discuss the dyadic data set. The data sets referenced throughout this appendix can be found at the first author’s website: http://bbs.utdallas.edu/pairlab/materials/.

Individual Difference Data Set

In the individual difference data set (see “example ID data set.sav”), each record of data corresponds to a single research participant and contains that participant’s self-ratings, personality traits, or other individual difference variables. If our study includes a total of 33 speed-dating groups, 16 of which are 6-person groups (i.e., 3 men and 3 women), 13 of which are 8-person groups (i.e., 4 men and 4 women), and 4 of which are 10-person groups (i.e., 5 men and 5 women), there are 240 individuals in the study. The individual difference data set will thus include 240 data records.
There are two essential identification variables that must be included on each record of this data set: A group identification number (\textit{GroupID}; e.g., ranging from 1 to 33), and a rater identification number (\textit{RaterID}; e.g., ranging from 1 to 10 in our example in which the largest group includes 5 men and 5 women). Inclusion of these two variables is important because they are needed for later merging purposes. It is also helpful to have a variable designating rater gender (\textit{gender}; e.g., effects coded such that 1 = male participant and -1 = female participant), and if a researcher is working with groups of varying sizes, we also suggest that he or she include a variable to index group size (e.g., \textit{GroupSize}). Importantly, we recommend computing grand-mean centered and standardized versions of the individual difference variables in this data set before merging it with the dyadic data set (indeed, grand-mean centering and standardizing these variables in the dyadic data set may be problematic if the groups vary in size). Table 7 presents an example data set with self-esteem and agreeableness as individual difference variables (grand-mean centered and standardized versions of agreeableness are also included).
Table 7

*Example Individual-Difference Data Set for Groups that Include 3 Men and 3 Women*

<table>
<thead>
<tr>
<th>GroupID</th>
<th>RaterID</th>
<th>Gender</th>
<th>Self-Esteem</th>
<th>Agreeableness</th>
<th>ZAgree</th>
<th>c_Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>-1</td>
<td>4.30</td>
<td>4.11</td>
<td>0.53</td>
<td>0.28</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-1</td>
<td>3.30</td>
<td>3.56</td>
<td>-0.52</td>
<td>-0.27</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
<td>3.90</td>
<td>3.56</td>
<td>-0.52</td>
<td>-0.27</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>-1</td>
<td>3.10</td>
<td>4.00</td>
<td>0.32</td>
<td>0.17</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>1</td>
<td>3.10</td>
<td>4.11</td>
<td>0.52</td>
<td>0.28</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>1</td>
<td>4.10</td>
<td>3.44</td>
<td>-0.72</td>
<td>-0.39</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>3.22</td>
<td>2.89</td>
<td>-1.77</td>
<td>-0.94</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>-1</td>
<td>4.40</td>
<td>4.44</td>
<td>1.15</td>
<td>0.61</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>1</td>
<td>4.20</td>
<td>4.11</td>
<td>0.53</td>
<td>0.28</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>-1</td>
<td>3.60</td>
<td>3.33</td>
<td>-0.93</td>
<td>-0.50</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>1</td>
<td>3.80</td>
<td>3.56</td>
<td>-0.52</td>
<td>-0.27</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>-1</td>
<td>3.80</td>
<td>4.22</td>
<td>0.74</td>
<td>0.39</td>
</tr>
</tbody>
</table>
Note. GroupID = identification value for rater’s speed-dating session. RaterID = identification value for rater within the speed-dating session. Gender = rater’s gender (effects coded such that 1 = man and -1 = woman). Self-Esteem = rater’s self-esteem. Agreeableness = Rater’s agreeableness. ZAgree = rater’s standardized score on agreeableness. c_Agree = rater’s grand-mean centered score on agreeableness.
Appendix A

Dyadic Data Set

The dyadic data set (see “example dyadic data set.sav”) contains the data generated by each of the “dates”. For instance, in an 8-person group with 4 men and 4 women, each person engages in four interactions and after each interaction rates aspects of the date as well as his or her perceptions of the target. For this example each 8-person group will therefore contribute 32 records to the dyadic data set. For our 33-group study the data set would have 18 records for each of the sixteen 6-person groups, 32 records for each of the thirteen 8-person groups, and 50 records for each of the four 10-person groups, for a total of 904 data records. Note that even if there are missing data (e.g., one rater did not rate one of his or her partners), it is important to still include a data record for that rater-target combination – it will simply report missing values for all variables other than the identification variables.

Each record in the dyadic data set contains all variables generated by one rater’s perceptions of a particular date or target. There are three key identification variables that are needed in this data set: A group identification variable (GroupID), a rater identification variable (RaterID), and a target identification variable (TargetID). Note that the values used for each subgroup to distinguish its members (i.e., the values for RaterID and TargetID) should be unique. For instance, researchers should not use the values 1, 2, and 3 to distinguish the male subgroup members, and also use the values 1, 2, and 3 to distinguish the female subgroup members. As one suggestion to ensure unique values between subgroups, researchers can use odd numbers to distinguish men (e.g., 1, 3, and 5) and even numbers to distinguish women (e.g., 2, 4, and 6). Table 8 presents an example of what such a dyadic data set should look like with attraction towards the target (Attract) and connection towards the target (Connect) as the dyadic variables.
Table 8

*Example Dyadic Data Set for one group that includes 3 men and 3 women*

<table>
<thead>
<tr>
<th>GroupID</th>
<th>RaterID</th>
<th>TargetID</th>
<th>Attract</th>
<th>Connect</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4.25</td>
<td>3.67</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>7</td>
<td>1.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>9</td>
<td>5.00</td>
<td>5.33</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3.75</td>
<td>3.00</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>7</td>
<td>2.50</td>
<td>2.67</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>9</td>
<td>3.25</td>
<td>2.67</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>3.25</td>
<td>1.67</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>4</td>
<td>3.75</td>
<td>3.00</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>6</td>
<td>5.00</td>
<td>4.67</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>5</td>
<td>1.75</td>
<td>2.00</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>7</td>
<td>1.75</td>
<td>3.33</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>9</td>
<td>3.50</td>
<td>3.67</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>2</td>
<td>4.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>4</td>
<td>5.50</td>
<td>4.67</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>6</td>
<td>3.75</td>
<td>3.33</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2</td>
<td>1.50</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>4</td>
<td>1.50</td>
<td>3.00</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>6</td>
<td>2.50</td>
<td>3.33</td>
</tr>
</tbody>
</table>

*Note.* GroupID = identification value for rater’s speed-dating session. RaterID = identification value for rater within the speed-dating session. TargetID = identification value for target within
the speed-dating session. Attract = rater’s attraction towards target. Connect = rater’s connection towards target.

**Merging the Two Data Sets and Creating a Pairwise Data Structure**

Although we have focused on two different data sets, our goal is to end up with a single data set that includes two data records for each dyad (i.e., one record in which the man is the rater and the woman is the target, and another record in which the woman is the rater and the man is the target), with the participants’ individual difference variables repeated on each of the lines on which he or she serves as the rater or the target. This is somewhat problematic, however, because the individual difference and dyadic data sets include different numbers of records (240 and 904, respectively). One solution to this problem for SPSS users is to duplicate the data contained in the individual-difference data set to match the number of records the rater has in the dyadic data set (note that SAS users can simply use a match merge procedure that will replicate the individual data as needed). For instance, individuals in 8-person groups will need to have four replications of their data in the individual-difference data set to match the 32 records in the dyadic data set.

This manipulation of the data is perhaps easiest if separate individual-difference data sets are created for each group size (e.g., “GroupSize6IDdata.sav”, “GroupSize8IDdata.sav”, and “GroupSize10IDdata.sav”). The following SPSS syntax illustrates how to accomplish this with the 6-person group size and can be easily modified to create individual-difference data sets of other sizes (note that the syntax will need to be modified to reflect the desired save location for the output):

```
DATASET COPY GroupSize6.
```
DATASET ACTIVATE GroupSize6.
FILTER OFF.
USE ALL.
SELECT IF (GroupSize = 6).
EXECUTE.
SAVE OUTFILE='C:\GroupSize6IDdata.sav'
/COMPRESSED.

Once these data sets are created, researchers will need to save each of them under a new data file name (e.g., “CopyGroupSize6IDdata.sav”, “CopyGroupSize8IDdata.sav”, and “CopyGroupSize10IDdata.sav”, respectively) so that they can merge the cases from the original data sets. So for our thirteen 8-person groups, we would initially have 104 records in our “CopyGroupSize8IDdata.sav” data file. By merging the cases from the “GroupSize8IDdata.sav” data file three times using “add cases”, we would end up with 416 records (four records for each of 104 individuals because they are each paired with four dates in the dyadic data set). The SPSS syntax used to execute this for the 8-person group is as follows (note that the number of merges required will be dependent on group size):

ADD FILES /FILE=* 
/FILE='C:\GroupSize8IDdata.sav'.
EXECUTE.
ADD FILES /FILE=* 
/FILE='C:\GroupSize8IDdata.sav'.
EXECUTE.
ADD FILES /FILE=*
Likewise, for our sixteen 6-person groups, we would initially have 96 records that would have to be merged twice to create a data set with 288 records. Finally, for our four 10-person groups, we would initially have 40 records that would have to be merged four times to create a data set with 200 records. To avoid confusion given the proliferation of data sets, each of these newly merged data sets can be saved with a name that clearly reflects that it is a copy that has been merged to contain repeated values (e.g., “CopyGroupSize6IDdataReplicated.sav”).

After merging all cases from the individual-difference data sets with the replications together (this should result in a data set with 904 records), we would merge the variables of the resulting individual-difference data set (see “IDMerged.sav”) with the dyadic data set (see “example dyadic data set.sav”) by matching on GroupID and RaterID. Remember to sort both data files with respect to these variables before merging. The SPSS syntax is:

```
SORT CASES BY GroupID(A) RaterID(A).
MATCH FILES /FILE=*  
/FILE='DataSet1'  
/BY GroupID RaterID.
EXECUTE.
```

At this point we have a merged data set (see “DyadicIDMerged.sav”) that includes two records for each date – one when the man was the rater and the other when the woman was the rater. Each record contains the individual difference and dyadic variables for the rater. Because we may be interested in how traits from the target predict relationship initiation outcomes, we need to expand the data set to also include the target’s individual difference variables on the
same record. Kenny and his colleagues (2006) refer to this type of data structure as a pairwise format.

The distinguishing feature of the pairwise format is that each data record contains the variables from both the rater (often called the actor) and the target (often called the partner). More specifically, when the rater is a man, the woman’s individual difference variables need to be included on that man’s record; likewise, when the rater is a woman, the man’s individual difference variables need to be included on that woman’s record. Thus, each person’s scores for individual difference variables need to be included in the data set twice.

To create a pairwise data set, we need to add so-called “partner” variables on any data record for which the person is the target. To do this, we first save our merged data set (see “DyadicIDMerged.sav”) into a file that we typically label the “Actor data set.” In this data set, we then edit all of the names for the individual difference and dyadic variables by adding a prefix of “A” to each (e.g., Aattract, Aselfesteem, and Agender). Note that we do not change the names of the identification variables. We then sort the data on GroupID, RaterID, and TargetID in ascending order, and then save and close the Actor data set (see “ActorDataSet.sav”).

The next step is to create a file we will save as the “Partner data set,” by opening the original merged data set (see “DyadicIDMerged.sav”) and saving it a second time as Partner data (see “PartnerDataSet.sav”). We then add a prefix of “P” to each of the individual difference and dyadic variables (e.g., Pattract, Pselfesteem, and Pgender); again, we do not change the names of the identification variables. The critical next step is that we change the name of the RaterID variable to be the TargetID and we change the name of the TargetID variable to be the RaterID within the partner data set. We then sort the Partner data set on the GroupID, RaterID, and
TargetID variables in ascending order and save the data set. Finally, we merge the variables of the actor and partner data sets by matching on the GroupID, RaterID, and TargetID variables.

Now that we have our data merged, we need to add two dummy-coded variables, \( mf \) and \( fm \), that denote which subgroup generated the data. The dummy-coded variable, \( mf \), will denote a dyad wherein the man is the rater and the woman is the target (\( mf = 1 \) if true, 0 otherwise). The dummy-coded variable, \( fm \), will denote a dyad wherein the woman is the rater and the man is the target (\( fm = 1 \) if true, 0 otherwise). As was mentioned earlier, we will need these variables to properly specify the random effects structure. Additionally, dummy-coded variables for each level of gender (man and woman) should be created to permit subsequent analyses that estimate parameter estimates for the two subgroups separately.

Assuming that the gender of the rater is coded such that men = 1 and women = -1, then the SPSS syntax for creating these variables is:

```
IF (Agender = 1) mf=1.
IF (Agender = -1) mf=0.
IF (Agender = -1) fm=1.
IF (Agender = 1) fm=0.
EXECUTE.
```

```
IF (Agender = 1) man=1.
IF (Agender = -1) man=0.
IF (Agender = -1) woman=1.
IF (Agender = 1) woman=0.
EXECUTE.
```
Next we need to create two additional variables that specify the particular male and female individuals within each subgroup: MaleID and FemaleID. For a dyad in which the rater is a woman and the target is a man, MaleID will refer to the TargetID value for the man and FemaleID will refer to the RaterID value for the woman. To illustrate this, consider the tenth record of data in Table 8. Given how we assigned even numbers to women and odd numbers to men for the RaterID and TargetID variables, we can see that the rater is a woman (RaterID = 6) and the target is a man (TargetID = 5). (Note that the inclusion of the mf and fm variables would also serve this purpose). Accordingly, the MaleID would be 5 and the FemaleID would be 6. The reverse will be true for a dyad in which the rater is a man and the target is a woman: MaleID will then refer to the RaterID value for the man and FemaleID will refer to the TargetID value for the woman. For instance, consider the seventh record of data in Table 7. We can see that the rater is a man (RaterID = 5) and the target is a woman (TargetID = 2); therefore, MaleID would be 5 and FemaleID would be 2.

Assuming that the gender of the rater is coded such that men = 1 and women = -1, then the SPSS syntax for creating these variables is:

```
IF (Agender = 1) MaleID=RaterID.
IF (Agender = -1) MaleID=TargetID.
IF (Agender = 1) FemaleID=TargetID.
IF (Agender = -1) FemaleID=RaterID.
EXECUTE.
```

If researchers are interested in obtaining standardized regression coefficients for any precursor question analyses, they should compute standardized scores for any relevant dyadic variables from the rater or actor at this time (e.g., ZAconnect). In addition, it is a good idea to
compute the effect codes for GroupID at this time. Here is an abbreviated version of the SPSS syntax that we used for our study of 33 groups:

IF (GroupID = 1) GroupID_EC1=1.
IF (GroupID ~= 1) GroupID_EC1=0.
IF (GroupID = 33) GroupID_EC1 =-1.
EXECUTE.

…and so on…IF (GroupID = 32) GroupID_EC32=1.
IF (GroupID ~= 32) GroupID_EC32=0.
IF (GroupID = 33) GroupID_EC32 =-1.
EXECUTE.

Researchers should then save this file as the “Pairwise data set.”