DOCUMENTATION GUIDE #13:
TLT COUPLES DATA

Prepared by Sara Yeatman, Jenny Trinitapoli, and Sarah Garver on September 3, 2018

This guide is designed to help users interested in analyzing TLT's couples data. It contains detailed guidance on the variables in the couple-linking file as well as background on how couples were enrolled. Men interviewed by TLT were enrolled either as a random-sample of men, or as male-partners of women already enrolled in the study. Note that the data from male partners do not represent a random sample of men and thus should not be analyzed as a standalone group nor in combination with the random-sample of men. Here we discuss the enrollment process for male partners as well as how to construct and analyze couple-level data.

Couple identification, recruitment, and retention

Dyadic processes and questions (e.g., how do partners influence each other's HIV risk?) were central to the design of TLT. Thus, in addition to enrolling a small, simple random sample of men aged 15-25 years living within the catchment area in 2009, TLT also enrolled men who were sexual and romantic partners (SRP) of women in the sample. These men were enrolled on an ongoing basis. During women's baseline interview, they were asked about their recent SRPs. For each relationship female respondents described as either "ongoing" or "confusing", the interviewer would give the woman a token. Women received these tokens during the interview -- at the end of the SRP section -- and were instructed to give these tokens to the partner(s) they named, who could then come to the TLT research center to enroll in the study. The tokens were small business cards with the TLT logo and a simple map to the TLT research center (see Figure 1). Interviewers added to the token a unique number (**maleid**) that would later be used to link the male SRP to the female respondent who reported and answered questions about him.

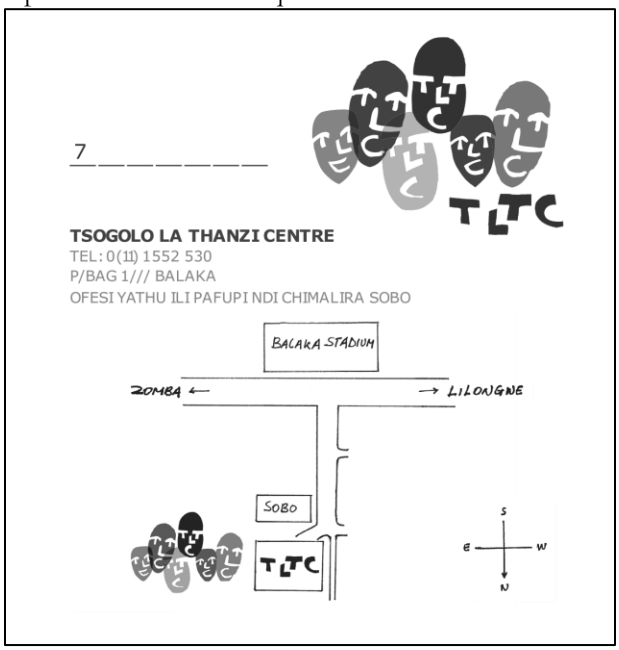

Figure 1. Sample blank token used during TLT-2.

Tsogolo la Thanzi is a research project designed by Jenny Trinitapoli and Sara Yeatman and funded by grants R01-HD058366 and R01-HD077873 from the National Institute of Child Health and Human Development.
For data access see https://www.icpsr.umich.edu/icpsrweb/content/DSDR/index.html.

At each subsequent wave through wave 8 of TLT-1 (2009-2012), women were asked the following questions about the men they mentioned in the SRP section: *did you tell me about this partner last time you came in (s6p), did you give this partner a token (s6t),* and *did the partner come in for an interview (s6r)?* Women were then given new tokens for each partner who was new or who had been reported but had not yet come to TLT for an interview.

At each wave, women could report on up to three SRPs, though the vast majority reported on only one man per wave and often told us over multiple or all waves about their relationship with one particular man. Women could enroll multiple partners, either at the same wave or at different waves. During TLT-1 & 2, 931 women recruited one partner, 116 women recruited two, and 19 women recruited three unique partners.

Once enrolled, male partners stayed in the study regardless of the status of their relationship. When a married couple divorced or a romantic relationship otherwise dissolved after the man enrolled, he remained in the study through wave 8.

Refresher sample: When the women's refresher sample was added to TLT-1 in 2012, women reported on their SRPs but no tokens were given. At TLT-2, however, refresher-sample women used tokens to enroll their partners according to the same procedures as the core-sample women (described below).

TLT-2: In 2015, all women were asked to give tokens to their current SRPs regardless of whether these men had been part of the first phase of the study. ***Male partners who had previously been in TLT-1 were only interviewed in 2015 if they were in ongoing relationships with a female respondent.***

Enrollment

To enroll, a male partner would simply come to the TLT research center with his token at any time the center was open. (TLT was typically staffed in cadence with business hours at the adjacent market: Mondays – Saturdays, 7am – 6pm, closed on Sundays.) The man presented his token to the front-desk manager who would then ask a few questions to check its validity, explain the parameters of the study, take his photograph to confirm his identity for future visits, assign him a respondent id based on the token he brought in, and connect him with a male interviewer to complete a baseline interview. Prior to the interview itself, the assigned interviewer would explain the study more fully and obtain informed consent. Regardless of when male partners were recruited during TLT-1, they always completed a baseline interview at their first visit. At subsequent interviews, they were given the corresponding questionnaire for that wave. For example, a man enrolling for the first time during wave 3 would be given the wave 1 (baseline) survey for that visit. He would then complete wave 4 survey at his next interview in accordance with the rest of the study cohort. These baseline data are deemed the "alternate baseline" datasets and must be merged to the main data waves. In 2015 (TLT-2), all men received the same questionnaire, regardless of prior participation.

Respondent IDs: Male partners first enrolled in TLT-1 are uniquely identified in the data by 7-digit respondent ids that begin with the number "6". Partner id's took the following format in TLT-1:

[6] + [wave # at which token was given]+ [unique 4-digit number shared with the referring woman] + [1, 2 or 3 depending on the number of tokens given to the woman at that wave]

The last digit served to differentiate the male partners of women who reported more than one new partner at a given wave; however, we could not control which partner the woman gave the token to. ***Analysts should not assume the last digit necessarily corresponds with the partner number in***

the SRP section. E.g., respid=6432512 would indicate a man in the partners' sample who was given a token from wave 4 by a woman who reported at least 2 partners. (And she could have reported 3.)

Figure 2. Example Construction of Male-Partner Respondent ID in TLT-1

Male partners who enrolled for the first time in 2015 (TLT-2) were given token numbers that followed the same logic but begin with 7s and have a '0' in place of the wave number (e.g. 7032511 would be a newly enrolled male partner at TLT-2 of the same woman as above who was given at least one token at TLT-2). Ongoing partners who had participated in the study during TLT-1 were linked and logged using their *original* **respid** (beginning with either a "6" or a "5" if they were from the core sample of men, see below). In the couple linking dataset (and all TLT data), analysts can infer certain things about respondents based on their **respid**.

- Core-sample women have 6-digit respondent ids
- Refresher-sample women have 4-digit respondent ids that begin with 9
- All men have 7-digit respondent ids.
    - If a **respid** begins with 5, the respondent is in the core sample of men, enrolled via a simple random sample of men aged 15-25.
    - If a **respid** begins with 6, the respondent is a man who enrolled as part of the couples' sample and was first interviewed during TLT1.
    - Men with a **respid** beginning with 7 are part of the couples' sample and were first interviewed during TLT-2 (2015).

Core-sample men in the couple-linking file: We correctly anticipated that men in the random sample would also occasionally appear as the SRPs of core-sample women. Women were instructed to deliver tokens to all the partners they named – even if they knew he was already a part of TLT. During TLT-1, fifteen men in the random sample came in with tokens they had received from serious or casual partners. During TLT-2, an additional six men from the random sample came in with tokens. These men are linked using their tokens to the referring woman in the couple database using their original respid (which starts with a "5"). Core-sample men in relationships with core-sample women were not re-interviewed if they had already completed an interview during that wave. Their survey data at each wave can be found in the core men's files (i.e., their data are not duplicated in the male partner files.)

Partners of two women: Given the high sampling fraction of women in the catchment area (1 in 4 women 15-25 years old were sampled), a single man could be romantically linked to multiple women in the study. This occurred only once. The multiple tokens enabled us to link this man to both of the women who reported and referred him, but he was interviewed only once at each wave and is in the dataset under a single respondent id.

The Couple-Linking File

To help analysts use the couple-level data, we created a couple-linking file that matches partners who are both respondents and their relationships across the course of the study. In this section, we lay out the structure of the linking file, the variables included, and then provide case studies that illustrate the nuance and complexity of the data (and relationships) through time.

- A total of 1220 unique couples are linked during the full TLT-1 & TLT-2 period.
- Regardless of when the man joined the study or how long the couple stayed together, each unique couple contributes 10 observations (i.e., couple-waves) to this file (eight waves during TLT-1, one for the RS wave, one for TLT-2).
- The full 10-wave dataset is designed to:
    - give analysts a bird's eye view of the female respondent's first mention of the male partner;
    - indicate the status of the relationship over time as reported by each member of the couple (sometimes discordant) and indicate whether data from both/either parties is available;
    - and differentiate between unit non-response (someone not interviewed) and relationship dissolution when links between partners are not made in a given wave.

The couple-linking file includes the following variables:

The couple-wave dataset is in long form (as opposed to wide). Each observation indexes the relationship status of a particular couple during each wave of the TLT study.

coupleid	Unique id for each couple
femaleid	ID of female member of the couple (matches **respid** in main datasets)
maleid	ID of male member of the couple (matches **respid** in main datasets)
wave	Wave
femalelink	Indicates the column of the SRP section in which the woman answered questions about this man. Wave-specific. 1 indicates p1; 2 indicates p2, 3 indicates p3.
malelink	Indicates the column of the SRP section in which the man answered questions about this woman. Wave-specific. 1 indicates p1; 2 indicates p2, 3 indicates p3.
femalereport	Relationship status as reported by the female respondent. Based on s9p in the SRP section of the questionnaire (linked to the appropriate column using **femalelink**). This variable also distinguishes if the (female) respondent did not report on that relationship during a completed interview (**femalereport**=4), or if the (female) respondent herself was not interviewed at that wave (**femalereport**=5). [1=ongoing; 2=report ended; 3=report confusing; 4=not reported; 5= not interviewed]
malereport	Relationship status as reported by the male respondent. Based on s9p in the SRP section of the questionnaire (linked to the appropriate column using **malelink**). This variable also distinguishes if the (male) respondent did not report on that relationship during a completed interview (**malereport**=4), or if the (male) respondent himself was not interviewed at that wave (**malereport**=5). [1=ongoing; 2=report ended; 3=report confusing; 4=not reported; 5= not interviewed]

MUST READ FOR ANALYST: Notes for use of Couple-Data

- To use these data responsibly, analysts need a thorough understanding of the entire timeline of the study (2009-2015) and its various facets. For example, analysts should know that refresher-sample women were first interviewed in 2012 (wave 9 in the data). By design, core-sample women and all men were not interviewed at wave 9 and appear as "not interviewed" in the wave 9 observation of the couple-linking file. See example under **Case #2** later in this document for help.
- Proceed cautiously with these data. Think of the couple-linking file as a tool that allows analysts to relate survey data from members of an actual or eventual couple, before, during, and after that relationship.
    - As structured, this dataset includes waves in which one or both members of a couple were never interviewed and others in which one or both members report that the relationship has ended (see **Case #3** below for an example).
    - Women occasionally reported on a male partner waves before he joined the study.
- Analysts will have to make choices about which waves to drop in light of their specific, analytic goals. Depending on the research objectives, an analyst might choose to exclude all waves that precede one partner's mention of the other, exclude all waves that don't include both partners' reports of the relationship, or exclude waves that follow any suggestion that a particular union has dissolved.
- The linking file is designed to be merged with the main TLT datasets for non-partner-specific variables (e.g., fertility preferences, religion, self-rated health) using **femaleid** and **maleid** (couple-linking file) and **respid** (main TLT datasets).
    - Note: **femaleid** is the female partner's **respid** and will need to be renamed to merge with main TLT datasets (and similarly with **maleid**).
- Most variables in the SRP section (beginning with "s") are responses to partner-specific questions and should not be used as characteristics of the respondent him/herself. Respondents could report on up to 3 partners at a time, including partners they recently stopped seeing (see survey instruments for details). The reference partner in this couple is just one of three possible partners the other reported on. The **femalelink** and **malelink** variables in the couple-linking file indicate the partnership number (1=p1, 2=p2, 3=p3) of the partner identified in SRP section of the main questionnaire *for that wave*. **Femalelink** indicates the column in which the linked male partner appears (p1, p2, or p3) in the SRP section of *her* questionnaire. Similarly, **malelink** indicates whether the female partner was reported as p1, p2, or p3 in his SRP section.
    - E.g., if a couple consists of Mary and Mphatso and Mary reported on two men at wave 4 -- Jonathan first and then Mphatso – **femalelink** would equal 2 to indicate that Mphatso was the 2nd partner reported on in the SRP section at wave 4. This indicates that variable s7p2, for example, refers to Mphatso, while s7p1 refers to Jonathan who may or may not be in the study. At his baseline interview during wave 5 (in this case, it took him a while to bring the token in), Mphatso mentions Mary as his only partner (**malelink**=1 and **malereport**= "ongoing"). At his wave 6 interview, Mphatso didn't even mention Mary because he considered their relationship over. In that case, **malelink** would be missing and **malereport** would indicate that Mphatso did "not report" on Mary at that wave (differentiating a non-report from a non-interview). See **Cases #3 and #4** for other illustrative examples.
- Not all SRP questions were asked at each wave, and there are wording changes to take seriously between the baseline and subsequent waves. In particular, sometimes the reference period changes from 1 year to past 4 months. The couple-linking file should always be used alongside the data key and actual survey instruments.

- Some SRP questions were asked only at the first time a partner was reported. Take care when using these questions; the first report may have happened before the partner entered the study (usually this is before the male partner entered but in one case the male partner was in the study first as the partner of a different female respondent). This information can be ascertained in the couples linking file using the **femalelink**/**malelink** variables and **wave**.
    - e.g. The following Stata code will create a variable indicating the wave in which a partner was first reported.

```
sort coupleid wave
gen tempf=wave if femalelink~=.
gen tempm=wave if malelink~=.
// for female partner reporting on male:
       bysort coupleid: egen female1stSRP=min(tempf)
// for male partner reporting on female:
       bysort coupleid: egen male1stSRP=min(tempm)
drop temp*
```

We recommend analysts interested in using the relationship-specific SRP-variables within couples, use the Couple_Data_Generating Stata do file we created. This file will combine all TLT datasets and reshape the data so that each observation contains a wave of data and both members of a couple's reports on their specific relationship.

Illustrative cases

Case 1: This is a straightforward case. A woman in the sample (110116) was first interviewed at wave 1, during which time she reported on her partner and he was given token 6101161. He enrolled in TLT at wave 1, which is clear because he reported on her at that wave (**malelink**). They are each interviewed at every wave with the exception of wave 9 when, by design, only the refresher sample was interviewed. At every wave in which they are interviewed, 110116 and 6101161 each report on the other in the first partner column of the SRP section (**femalelink**, **malelink**). They both consistently report that the relationship is ongoing (**femalereport**, **malereport**). This is still the case at wave 10 (TLT-2).

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
1210.	1	1001	110116	6101161	1	1	ongoing	ongoing
1713.	2	1001	110116	6101161	1	1	ongoing	ongoing
3167.	3	1001	110116	6101161	1	1	ongoing	ongoing
3713.	4	1001	110116	6101161	1	1	ongoing	ongoing
5591.	5	1001	110116	6101161	1	1	ongoing	ongoing
6185.	6	1001	110116	6101161	1	1	ongoing	ongoing
8406.	7	1001	110116	6101161	1	1	ongoing	ongoing
8809.	8	1001	110116	6101161	1	1	ongoing	ongoing
10257.	9	1001	110116	6101161	.	.	not interviewed	not interviewed
11031.	10	1001	110116	6101161	1	1	ongoing	ongoing

Case 2: This is a case from a woman in the refresher sample. This woman reported the man at wave 9 when she was first interviewed. Women were not given tokens at that wave, so he could not have joined the study at wave 9. They are still together at wave 10 (TLT-2), at which time she gives him a token and he enrolls in the study. They both report each other in the first column of SRP (**femalelink**=1 and **malelink**=1) and report the relationship as ongoing **(femalereport** and **malereport**).

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
12191.	1	2601	9044	7090441	.	.	not interviewed	not interviewed
12192.	2	2601	9044	7090441	.	.	not interviewed	not interviewed
12193.	3	2601	9044	7090441	.	.	not interviewed	not interviewed
12194.	4	2601	9044	7090441	.	.	not interviewed	not interviewed
12195.	5	2601	9044	7090441	.	.	not interviewed	not interviewed
12196.	6	2601	9044	7090441	.	.	not interviewed	not interviewed
12197.	7	2601	9044	7090441	.	.	not interviewed	not interviewed
12198.	8	2601	9044	7090441	.	.	not interviewed	not interviewed
12199.	9	2601	9044	7090441	1	.	ongoing	not interviewed
12200.	10	2601	9044	7090441	1	1	ongoing	ongoing

Case 3: This is a more complex case. The male partner was first interviewed in wave 6 (see **malelink** or **malereport**). At that wave, the woman reported this partner in the first column of her SRP section (**femalelink**) and the man reported her in first column of his SRP section (**malelink**). At wave 7, she was interviewed but no longer reported on him (**femalelink**=. and **femalereport**= "not reported"). He, on the other hand, reported on her in the second column of the SRP section (**malelink**=2) and reported that their relationship had ended (**malereport**= "report ending"). At wave 8, both partners were interviewed but neither reported on each other. As designed, neither partner was interviewed in wave 9. The woman was reinterviewed in wave 10 (TLT-2) but did not report on this man, who therefore could not be interviewed at wave 10 as he was no longer a current partner (criterion for partner recruitment at wave 10/TLT-2).

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
788.	1	1916	114671	6646711	.	.	not reported	not interviewed
2426.	2	1916	114671	6646711	.	.	not reported	not interviewed
3276.	3	1916	114671	6646711	.	.	not reported	not interviewed
4286.	4	1916	114671	6646711	.	.	not reported	not interviewed
5819.	5	1916	114671	6646711	.	.	not reported	not interviewed
6233.	6	1916	114671	6646711	1	1	ongoing	ongoing
7461.	7	1916	114671	6646711	.	2	not reported	report ending
9224.	8	1916	114671	6646711	.	.	not reported	not reported
10539.	9	1916	114671	6646711	.	.	not interviewed	not interviewed
12157.	10	1916	114671	6646711	.	.	not reported	not interviewed

Case 4: Here is a case in which a woman in the sample was in a relationship with a man already enrolled in TLT as part of the random men's sample (**maleid** begins with "5"). She reports on him consistently (except for wave 9, when neither were interviewed), although she reports at wave 4 that their relationship is confusing. He reports on her in waves 1 and 2, then indicates in wave 3 that their relationship ended, and does not mention her again until wave 6 when he reports their relationship and again characterizes it as ongoing. This is a good reminder that men and women in the same relationship might perceive the state of the relationship differently (or report on it so). Discrepancies like this can also reflect temporal processes of union formation and dissolution. All interviews in a given wave occurred within the same ~3-4-month period but could have been conducted up to four months apart. Exact interview dates can be derived from the main datasets.

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
586.	1	2599	122090	5142140	1	1	ongoing	ongoing
1450.	2	2599	122090	5142140	1	1	ongoing	ongoing
2516.	3	2599	122090	5142140	1	1	ongoing	report ending
4010.	4	2599	122090	5142140	1	.	confusing	not reported
5248.	5	2599	122090	5142140	1	.	ongoing	not reported
6609.	6	2599	122090	5142140	1	1	ongoing	ongoing
7415.	7	2599	122090	5142140	1	1	ongoing	ongoing
9724.	8	2599	122090	5142140	1	1	ongoing	ongoing
10520.	9	2599	122090	5142140	.	.	not interviewed	not interviewed
11719.	10	2599	122090	5142140	1	1	ongoing	ongoing

Case 5: In this example, the couple is not matched until wave 10 (TLT-2) when the male partner comes in for the first time (**maleid** begins with "7"). However, we were able to link the man to an earlier report the woman made of him at wave 8.

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
11291.	1	2202	115707	7057071	.	.	not reported	not interviewed
11292.	2	2202	115707	7057071	.	.	not reported	not interviewed
11293.	3	2202	115707	7057071	.	.	not reported	not interviewed
11294.	4	2202	115707	7057071	.	.	not reported	not interviewed
11295.	5	2202	115707	7057071	.	.	not reported	not interviewed
11296.	6	2202	115707	7057071	.	.	not reported	not interviewed
11297.	7	2202	115707	7057071	.	.	not reported	not interviewed
11298.	8	2202	115707	7057071	1	.	ongoing	not interviewed
11299.	9	2202	115707	7057071	.	.	not interviewed	not interviewed
11300.	10	2202	115707	7057071	1	1	ongoing	ongoing

Case 6: Similar to the case above, if more unusual, the female respondent first reported on this partner at wave 1, at which time she reported on him in the second column of the SRP section (**femalelink**=2). She didn't mention him again until wave 5. This is also when he first joined the study and reported on her (both reported on each other in the first column of SRP: **femalelink**=1 and **malelink**=1). They were still in a relationship at wave 10 (TLT-2) when he was again recruited as a partner and logged using his original respondent id (**respid** in main datasets; **maleid** in the couple-linking file).

	wave	coupleid	femaleid	maleid	femalelink	malelink	femalereport	malereport
7951.	1	1826	118676	6586761	2	.	ongoing	not interviewed
7952.	2	1826	118676	6586761	.	.	not reported	not interviewed
7953.	3	1826	118676	6586761	.	.	not reported	not interviewed
7954.	4	1826	118676	6586761	.	.	not reported	not interviewed
7955.	5	1826	118676	6586761	1	1	ongoing	ongoing
7956.	6	1826	118676	6586761	1	1	ongoing	ongoing
7957.	7	1826	118676	6586761	1	1	ongoing	ongoing
7958.	8	1826	118676	6586761	1	1	ongoing	ongoing
7959.	9	1826	118676	6586761	.	.	not interviewed	not interviewed
7960.	10	1826	118676	6586761	1	1	ongoing	ongoing

Case 7: In this case, the woman reports on the man at wave 2, but he does not enroll until wave 4 (with a token from wave 3 **maleid**=6**3**98501). Because TLT surveys asked more detailed questions on partners the first time he or she was reported in SRP (e.g. start date of relationship, partner's education, etc.), analysts in this case may want to use her wave 2 report on this partner when she first tells us about the relationship. He would have answered similar questions about her at wave 4.

	wave	coupleid	femaleid	maleid	female~k	malelink	femalereport	malereport
7411.	1	1771	119850	6398501	.	.	not reported	not interviewed
7412.	2	1771	119850	6398501	1	.	ongoing	not interviewed
7413.	3	1771	119850	6398501	1	.	ongoing	not interviewed
7414.	4	1771	119850	6398501	1	1	ongoing	ongoing
7415.	5	1771	119850	6398501	1	1	ongoing	ongoing
7416.	6	1771	119850	6398501	1	1	ongoing	ongoing
7417.	7	1771	119850	6398501	1	1	ongoing	ongoing
7418.	8	1771	119850	6398501	1	1	ongoing	ongoing
7419.	9	1771	119850	6398501	.	.	not interviewed	not interviewed
7420.	10	1771	119850	6398501	1	1	ongoing	ongoing

The following text is a STATA do-file that can be used for merging the couple data into the main survey waves. File paths and filenames should be checked based on the most recently available data. If working with only a subset of the main waves, the user should: (1) empty the file-path in part I, but leave the macro name as-is (e.g. `local file2w ""`), and (2) revise the highlighted line of text to drop any unnecessary waves.

Do-file version is from 2019-09-11, S. Garver

```
// do-file starts here

*******************************************************************************
* Setup Directory
*******************************************************************************
cd "" // This should be the file path for all of your data files
global D ""  // Use this to designate where any newly created files will be saved

ssc install mmerge
clear matrix
clear mata
set maxvar 32000            // since dataset is large with all waves, and then gets reshaped
wide for couple data, (doubled in variables)

*******************************************************************************
*                           PART I: COMPILE RELEVANT DATASETS
*******************************************************************************
// Identify datasets to append
      /*
      1. file0 should be the couple dataset
      2. The datasets are listed in three parts based on the three subset of public datasets
      3. You can leave blank the local macros if you don't need them. The "capture"
      command should generate any errors that exist but will continue to execute the loop
      */
/* Couple-Linking Data */
local file0 "$D/Full_TLT_Couples_Data_180814.dta"

use "`file0'", clear
keep wave coupleid female*
rename femaleid respid
rename female* *
tempfile femalecouplelink
save `femalecouplelink'

use "`file0'", clear
keep wave coupleid male*
rename maleid respid
rename male* *
tempfile malecouplelink
save `malecouplelink'

*******************************************************************************
/* List Women's Data */
local file1w "W1_public_women_20190722.dta"
local file2w "W2_public_women_20190722.dta"
local file3w "W3_public_women_20190722.dta"
local file4w "W4_public_women_20190722.dta"
local file5w "W5_public_women_20190722.dta"
local file6w "W6_public_women_20190722.dta"
local file7w "W7_public_women_20190722.dta"
local file8w "W8_public_women_20190722.dta"
local file9w ""

forvalues i = 1/9 {
      if `i'==1 {
      use `file`i'w', clear
            }
      else {
      capture noi append using `file`i'w'
            }
```

```
	}

	mmerge respid wave using `femalecouplelink'
	tempfile womensdata
	save `womensdata'

*******************************************************************************

/* List Male Partner's Data */
local file1m "W1_public_malepartners_20190722.dta"
local file2m "W2_public_malepartners_20190722.dta"
local file3m "W3_public_malepartners_20190722.dta"
local file4m "W4_public_malepartners_20190722.dta"
local file5m "W5_public_malepartners_20190722.dta"
local file6m "W6_public_malepartners_20190722.dta"
local file7m "W7_public_malepartners_20190722.dta"
local file8m "W8_public_malepartners_20190722.dta"
local file9m ""

/* List Random Men's Data */
local file1mr "W1_public_randommen_20190722.dta"
local file2mr "W2_public_men_20190722.dta"
local file3mr "W3_public_men_20190722.dta"
local file4mr "W4_public_men_20190722.dta"
local file5mr "W5_public_men_20190722.dta"
local file6mr "W6_public_men_20190722.dta"
local file7mr "W7_public_men_20190722.dta"
local file8mr "W8_public_men_20190722.dta"
local file9mr ""

/* List Alternate Data Files */
local file1ab "BL_public_20190722.dta"

forvalues i = 1/9 {
	if `i'==1 {
	use `file`i'm', clear
	capture noi append using `file`i'mr'
		}
	else {
	capture noi append using `file`i'm'
	capture noi append using `file`i'mr'
		}
	}

	append using `file1ab'

	mmerge respid wave using `malecouplelink'
	tempfile mensdata
	save `mensdata'

// Append the Women's Data  for FULL, COMBINED dataset
	append using `womensdata'

*******************************************************************************

// Make sure gender, coupleid, and wave are non-missing for reshape
	drop if coupleid==.
*	drop if wave>2 // Drop whichever waves are not needed for analysis purposes

	*add gender for unreported. Gender used as _j in reshape. Can't be missing.
	replace gender=1 if respid>5000000 & gender==.
	replace gender=2 if respid<500000 & gender==.

save "$D/coupledata_long.dta", replace

*******************************************************************************
*                           PART II: GENERATE Couple Dataset
*******************************************************************************
use "$D/coupledata_long.dta", clear

// Reshape based on coupleid & gender
```

```
rename * *_

rename (wave_ coupleid_ gender_ link_ report_ respid_) (wave coupleid gender link report
respid)
order coupleid gender wave link report respid, first

drop _* // for reshape
reshape wide link respid report *_, i(wave coupleid) j(gender)

      rename link1 malelink
      rename link2 femalelink
      rename respid1 maleid
      rename respid2 femaleid
      rename report1 malereport
      rename report2 femalereport

      order coupleid wave *link *report *id , first

save "$D/coupledata_wide.dta", replace
use "$D/coupledata_wide.dta", clear


// Generate new *MENS* variables for couple linking
unab vars: *p1_1
local stubs: subinstr local vars "p1_1" "", all
noi di "`stubs'"

foreach s of local stubs {
    noi display "`s'"
      capture confirm numeric var `s'p1_1

                            if _rc==0 {
                                    gen `s'_1=.
                                                replace `s'_1=`s'p1_1 if 1==malelink
                                                replace `s'_1=`s'p2_1 if 2==malelink
                                                replace `s'_1=`s'p3_1 if 3==malelink
                                            }

                            if _rc!=0 {
                                    gen `s'_1 = ""
                                                replace `s'_1=`s'p1_1 if 1==malelink
                                                replace `s'_1=`s'p2_1 if 2==malelink
                                                replace `s'_1=`s'p3_1 if 3==malelink
                                            }
      drop `s'p1_1 `s'p2_1 `s'p3_1
                            }

// Generate new *WOMENS* variables for couple linking
unab vars: *p1_2
local stubs: subinstr local vars "p1_2" "", all
noi di "`stubs'"

// This loop generates a single partner, SRP variable and matches that variable to the correct
column-data of the SRP
foreach s of local stubs {
    noi display "`s'"
      capture confirm numeric var `s'p1_2
                            if _rc==0 {
                                    gen `s'_2=.
                                                replace `s'_2=`s'p1_2 if 1==femalelink
                                                replace `s'_2=`s'p2_2 if 2==femalelink
                                                replace `s'_2=`s'p3_2 if 3==femalelink
                                            }

                            if _rc!=0 {
                                    gen `s'_2 = ""
                                                replace `s'_2=`s'p1_2 if 1==femalelink
                                                replace `s'_2=`s'p2_2 if 2==femalelink
                                                replace `s'_2=`s'p3_2 if 3==femalelink
                                            }
       drop `s'p1_2 `s'p2_2 `s'p3_2
```

```
                            }

save "$D/coupledata_public_wide_20190722.dta", replace

// CLOSE OUT
global macro drop // this drops all global macros so as not to mess up further data-runs

// end of do-file
```

Further questions or challenges met while working with the couple-data should be directed to the User Forum on TLT's study website: http://tsogololathanzi.org