

ICPSR 26841

**Midlife in the United States  
(MIDUS 2): Daily Stress  
Project, 2004-2009**

Carol D. Ryff  
*University of Wisconsin-Madison*

David M. Almeida  
*Pennsylvania State University*

Field Notes

Inter-university Consortium for  
Political and Social Research  
P.O. Box 1248  
Ann Arbor, Michigan 48106  
[www.icpsr.umich.edu](http://www.icpsr.umich.edu)

# Terms of Use

The terms of use for this study can be found at:  
<http://www.icpsr.umich.edu/cocoon/ICPSR/TERMS/26841.xml>

## Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

### NOTICE

#### WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.



## MIDUS Project 2:

### *National Study of Daily Experiences*

#### Wave 2: Field Notes

For file: M2\_P2\_DAILY DATA\_7-14-09.sav

**July 14, 2009**

---

Address correspondence to David M. Almeida, Department of Human Development and Family Studies, The Pennsylvania State University, 114-H Henderson Building, University Park, Pennsylvania 16802 (email: [dalmeida@psu.edu](mailto:dalmeida@psu.edu)). This project was supported by the Network on Successful Mid-Life Development of the John D. and Catherine T. MacArthur Foundation and Grants R01 AG19239 and P01 AG0210166-02, from the National Institutes of Health.

## ***National Study of Daily Experiences***

### **Wave 2: Field Notes**

This document highlights aspects of the daily telephone diary data that analysts should be aware of prior to working with the file. We start by first explaining the variable naming scheme and the administrative variables which are the first few in the dataset. Issues specific to the entire dataset such as missing values are next described. Finally, we explain the difference between daily diary data and the aggregated data, as well as how the variables should be interpreted when using aggregated data.

#### Variable Naming Scheme

The document “Naming and Coding Conventions” (M2\_Variable Naming and Coding Conventions\_10-26-05.doc) specifies that the third character of the variable name be a letter that identifies the type, or name, of the instrument used to collect the data. In keeping with this specification, the following naming conventions have been used in Project 2:

D	= Daily Telephone Interview
A	= Time Use, Giving, and Volunteering
B	= Physical Health Symptoms and Cigarette and Alcohol Usage
C	= Non-Specific Psychological Distress and Positive Affect
D	= Work Productivity and Cutbacks
F	= Daily Inventory of Stressful Events (DISE) and Positive Events
H, I, J, & K	= Final Day Questions
S	= Discrimination Questions

#### Administrative Variables

The following administrative variables appear toward the beginning of the file:

Examples:

- B2DDAY = Interview Day (Days 1 to 8)
- B2DDATE = Interview Date
- B2DMISS = Missed Phone Interview

#### Constructed Variables

Constructed variables appear toward the end of the data set. Information on the constructed variables is provided in “M2\_P2\_Scales Documentation\_7-14-09.pdf”.

Examples:

- B2DN\_STR = Number of Stressors
- B2DA\_STR = Any Stressor
- B2DN\_SYM = Number of Symptoms
- B2DA\_SYM = Any Symptom

In a latter section, we also show how one can construct aggregated variables from “day-level” variables.

### Missing Values

In accordance with M2\_Variable Naming and Coding Conventions\_10-26-05.doc, missing values appear as follows:

- 7, 97, 997, etc. = Do not know
- 8, 98, 998, etc. = Refused/Missing
- 9, 99, 999, etc. = Inapplicable

### Daily Level Variables & Aggregated Data

#### **The Person-Day Dataset**

The NSDE daily data file (M2\_P2\_DAILY DATA\_7-14-09.sav) is structured as a “person-day” dataset such that each row of data corresponds to a single day for a given individual. Data regarding days are ‘nested’ within individuals over time. Micro-level information (data collected on a given day) comes from a macro-unit, in this case a single individual. In contrast to typical multivariate datasets, where each row of values corresponds to one single individual and that individual’s values for some set of variables, each row of a person-day dataset corresponds to an individual’s values for some set of variables, on that day. Figures 1 and 2 depict the traditional multivariate and person-day datasets, respectively. Both figures 1 and 2 represent data from 5 participants, actually, the same 5 participants (note the *same* IDs). The figures below depict data from two components of a study. First, Figure 1 shows data about participant’s age, gender, and depression scores taken at the baseline assessment (i.e., MIDUS P1). These data can be seen as characteristics that vary across individuals in the study (traditional individual differences research).

*Figure 1.* Traditional multivariate dataset.

ID	Age	Gender	Depression
101	45	M	4
102	86	F	7
103	37	F	11
104	72	M	8
105	66	M	18

Figure 2 depicts data collected from the same participants assessed on three consecutive days. Participants were asked whether they experienced a stressor on that day (variable ‘Any Stress’; Yes = 1, No = 2), as well their negative affect (Neg. Affect) over the past 24 hours (sum of NA items). The day of assessment is indicated by the variable ‘DAY’.

*Figure 2. Person-day dataset (3-days of assessment).*

ID	Day	Any Stress	Neg. Affect
101	1	1	7
101	2	2	5
101	3	1	7
102	1	1	11
102	2	1	15
102	3	2	10
103	1	1	9
103	2	1	9
103	3	1	8
104	1	2	5
104	2	2	6
104	3	1	9
105	1	2	5
105	2	2	5
105	3	2	6

Figure 2 shows participant 101 experienced a stressful event on the first and third days that they were assessed, whereas participant 105 did not report experiencing a stressor on any of the days. Figure 2 also shows that across all 5 participants, their negative affect varies from one day to the next. Figure 2 displays quite clearly the concept of the ‘nested’ data structure. Notice that there are multiple observations for each individual, and these observations are organized by ID and DAY of assessment.

### **Linking Traditional Multivariate and Person-Day Datasets**

It may be of interest to link the data from the baseline assessment to the daily assessments to answer certain research questions (e.g., Are age and depressive symptomatology related to the likelihood of experiencing a stressor on any given day?). To answer these questions, we would need to merge the two datasets into one. Let’s assume that the data from the traditional multivariate dataset (baseline variables, including age, gender, and depression scores) are in a dataset called ‘baseline’, and the data from the daily component of the study are contained in a dataset called ‘daily’. SAS and SPSS codes to merge these two datasets into one are shown below:

**[NOTE: Prior to creating the aggregated dataset, the following recode occurred (YES: 1=1; NO: 2=0)].**

## SAS Codes

**Data** all; \*Create/name a dataset;

**Merge** baseline daily; \*Tells SAS that the new dataset 'all' is going to be a combination of the 'baseline' and 'daily' datasets;

**By** ID; \*Tells SAS that 'ID' is the relevant variable by which the data should be linked;

**Run**;

## SPSS Codes

MATCH FILES /FILE=\*

/TABLE='C:\NSDE\aggr2.sav'

/BY ID

EXECUTE.

When these commands are executed, the resulting dataset should look similar to Figure 3.

*Figure 3.* Combined multivariate (baseline) and person-day (daily) datasets.

ID	Day	Any Stress	Neg. Affect	Age	Gender	CESD
101	1	1	7	45	M	4
101	2	0	5	45	M	4
101	3	1	7	45	M	4
102	1	1	11	86	F	7
102	2	1	15	86	F	7
102	3	0	10	86	F	7
103	1	1	9	37	F	11
103	2	1	9	37	F	11
103	3	1	8	37	F	11
104	1	0	5	72	M	8
104	2	0	6	72	M	8
104	3	1	9	72	M	8
105	1	0	5	66	M	18
105	2	0	5	66	M	18
105	3	0	6	66	M	18

Figure 3 shows that age, gender, and CESD variables from the 'baseline' dataset have been appended to the 'daily dataset. Furthermore, notice that each participant's age, gender, and CESD score have been appended to EACH row. Because age, gender, and CESD only were assessed once each variable has a single value for each participant, and that value is included at each observation ('DAY') for each participant ('ID'). Although age, gender, and CESD only were assessed once, the fact that they appear at each observation after merging the datasets is normal. Because these measures were not time-varying (i.e., collected more than once), in the context of this example, we would not expect them to have differing values across DAYS for any given participant. If a value

did vary from day to day (say CESD for participant 103 was 11, 14, 11, for days 1 to 3, respectively), something went wrong in the data merging process.

### Creating an Aggregate Dataset

If one is interested in examining between-person correlations using the ‘Any Stress’ variable, then an aggregate value needs to be computed. Because ‘Any Stress’ and ‘Negaff’ are time-varying, that is, not a single value for any one participant, we need to create a person-level aggregate of this variable. This is done using the following SAS and SPSS codes:

#### SAS Codes

```
Proc means data=daily;*Execute MEANS procedure in SAS using the 'daily' dataset';
By id;*Tells SAS to group observations by ID;
Var anystress negaff;*Variable you want the Means for;
Output out=agg (drop = _TYPE_ _FREQ_) mean(anystress) mean(negaff) = stress_mn
na_mn;*Tells SAS to output the mean for each individual, variables named ‘stress_mn’
and ‘na_mn’ to a dataset call 'agg';
Run;
```

#### SPSS Codes

```
AGGREGATE
/OUTFILE='C:\MIDUS\NSDE I\aggr.sav'
/BREAK=B2DM2ID
/B2DA1_mean = MEAN(B2DA1) /B2DA2_mean = MEAN(B2DA2) /B2DA3_mean =
MEAN(B2DA3).
```

The resulting dataset ‘agg’ should look like Figure 4.

*Figure 4.* Aggregate of ‘any stress’ from ‘daily’ dataset.

ID	Stress_mn	NA_mn
101	0.67	6.33
102	0.67	12.00
103	1.00	8.67
104	0.34	6.67
105	0.00	5.33



Because daily stress was coded dichotomously, the resulting aggregate value ‘Stress\_mn’ reflects the proportion of days participants reported experiencing a stressor. Whereas the ‘Any Stress’ variable in the daily dataset corresponded to a particular value for a particular individual for a particular day (whether they experienced a stressor), the aggregate variable ‘Stress\_mn’ reflects how frequently each individual reported experiencing a stressor. Similarly, the variable ‘NA\_mn’ reflects each participant’s average negative affect across the study days. We have created a between-person, or macro-unit, estimate of daily stress and negative affect. These new variables, ‘Stress\_mn’ and ‘NA\_mn’ then could be appended to the ‘baseline’, ‘daily’, or ‘all’ datasets by using code similar to what is provided above in the section on linking datasets. Combining the ‘agg’ dataset with the ‘baseline’ dataset would provide an aggregate dataset (one that can be used for inter-individual differences analyses), as shown in Figure 5.

*Figure 5. Aggregate Dataset Including Baseline and Aggregate Daily Variables*

ID	Age	Gender	Depression	Stress_mn	NA_mn
101	45	M	4	0.67	6.33
102	86	F	7	0.67	12.00
103	37	F	11	1.00	8.67
104	72	M	8	0.34	6.67
105	66	M	18	0.00	5.33