Values Affirmation Database README

02 November, 2015

Contents

Remaining tasks and Residual concerns — 27

Conclusions — 28

Introduction

This document describes the database compiled as part of the NSF and ROADS sponsored research for the investigation of the values affirmation essays using natural language processing methodology.

The database consists of four files:

1. **demog** - This file features the demographic characteristics of the participants in the studies. As indicated by the number string in the filename, it was last updated on 9/11/2015. Similar dates are appended to the other files.
2. **essays** - This file includes the essays and their corresponding attributes.
3. **grades** - This file includes grades and other academic outcome measures for participants in the studies
4. **prompts** - This file includes characteristics of the prompts for each of the possible interventions.

Each file's name indicates the content of the data, plus the date on which the file was generated. At the time of this writing, the files have the names `demog9.11.15.csv`, `essays9.30.15.csv`, `grades10.5.15.csv`, and `prompts10.8.15.csv`.

File format and structure

In order to make these files usable across a variety of software types and system environments, they have been stored as basic csv files. However, using commas to separate the field values in the essays and prompts file will obviously lead to problems, because the content of the fields occasionally contain commas. For that reason, these files are stored as pipe-separated values (a pipe is a |). It's possible to see the difference in structure by opening the files using a basic text-editor (textEdit on mac, notepad, sublimetext, textwrangler, etc)

The format of each file follows the general principles of relational databases. Each file is stored in what is often referred to as a 'long' format, and each file features a specific type of observational unit (i.e. essays, participants, outcomes, or prompts). Within each file, these observational units are organized by the variables, which form columns, and the observations, which form rows.

The variables for each file are listed below.

Essays

```
## [1] "ID"              "Study"           "Intervention_number"
## [4] "Essay"           "Condition"       "Intervention_Date"
## [7] "Cohort"          "annotated"       "corrected"
```

Participants

```
## [1] "ID"        "Study"     "Cohort"    "Ethnicity" "Gender"
```

Outcomes

```
## [1] "ID"              "Study"           "Grade"
## [4] "Grade_type"      "est_Grade_date"  "intervention_year"
```

Prompts

```
## [1] "Intervention_number" "Condition"            "Intervention_Date"
## [4] "Study"                "Cohort"               "Essay.Prompt"
## [7] "Values"
```

Note that this file structure makes it very clear what information is contained where. Very little explication of the content of each file is needed here.

All files with the exception of the prompts share a common *key* of the participant ID number. This allows the analyst to join together data from each file using basic join operations. Where studies had duplicate ID numbers, we appended a `.x` to the end of the ID, where x is an integer, incremented each time we encountered a file that had ID numbers that duplicated ID numbers already incorporated into the database.

The prompts file, since the unit of measure is the intervention and not the individual, does not have a specific participant ID associated with each row. Instead, the information from this file can be joined to the others using one of the multiple keys it has in common with the other files. For instance, if the analyst were interested in joining the prompts to the Essays table, one could join by the combination of `Intervention_number`, `Intervention_date`, and `Condition`, which would correctly join the information in the two, yielding a new table that contained the prompt information for each essay oin the Essays file.

Furthermore, these data are stored in what is sometimes referred to as a 'long' format. This means that there are a minimal number of columns, and each table features measurements on a single element (i.e. essays, participants, outcomes, or prompts). The rows are the individual observations

File contents

I now use these files to illustrate some basic properties about the database. We will investigate these files one-by-one. I will also go in depth for the Windsor data in some areas.

Essays

- **Total Number of essays**

```
## [1] 9528
```

- **Number of essays per study**

```
##
##   Awareness&Affirmation           Belonging                Blindness
##                     227                 165                       35
##            CO BMI Study         CO/CA Latino              Connecticut
##                     205                1161                     6256
##                  Debate     Group Affirmation NeuroBioThreat Study 1
##                     143                 104                      129
##                  Physics  Political Affirmation                Surgeons
##                     798                  93                      212
```

- **Number of essays per Cohort - Windsor data only**

```
## 
##  Cohort 1  Cohort 2  Cohort 3  Cohort 4  Cohort 5 Cohort 6   Cohort 7
##       932      1121      1014      1339      1064       352       317
##  Cohort 8
##      117
```

- **Number of essays per interventions per Cohort - Windsor data only**

```
## 
##      Cohort 1 Cohort 2 Cohort 3 Cohort 4 Cohort 5 Cohort 6  Cohort 7
##  1        133      173      129      159      162      133       168
##  2        132      154      129      154      156       90       149
##  3        131      153      124      157      162      129         0
##  4        132      159      127      150      158        0         0
##  5        123      165      128      150      139        0         0
##  6        139      158      125      149      138        0         0
##  7        138      156      125      148      146        0         0
##  8          1        3      127      142        3        0         0
##  9          1        0        0      130        0        0         0
##  10         1        0        0        0        0        0         0
##  11         1        0        0        0        0        0         0
## 
##      Cohort 8
##  1       117
##  2         0
##  3         0
##  4         0
##  5         0
##  6         0
##  7         0
##  8         0
##  9         0
##  10        0
##  11        0
```

- **Number of essays per intervention date per Cohort - Windsor data only**

```
## 
##             Cohort 1 Cohort 2 Cohort 3 Cohort 4 Cohort 5 Cohort 6
##  2003-10-22      133        0        0        0        0        0
##  2003-11-20      132        0        0        0        0        0
##  2003-12-16      131        0        0        0        0        0
##  2004-01-15      132        0        0        0        0        0
##  2004-05-20      123        0        0        0        0        0
##  2004-09-20      139      173        0        0        0        0
##  2004-10-26      138      154        0        0        0        0
##  2005-01-19        0        0        1        0        0        0
##  2005-05-04        0      153        0        0        0        0
##  2005-09-22        0        0      127        0        0        0
##  2005-09-23        0        0        2      159        0        0
##  2005-09-30        1      159        0        0        0        0
##  2005-10-21        0        0      127        0        0        0
##  2005-10-26        1      162        0        0        0        0
```

```
##    2005-10-28         0         0         2       154         0         0
##    2006-01-06         0         0         3       157         0         0
##    2006-01-19         0         0        45         0         0         0
##    2006-01-20         0         0         1         0         0         0
##    2006-01-25         0         0        38         0         0         0
##    2006-01-26         0         0         1         0         0         0
##    2006-01-27         0         0        35         0         0         0
##    2006-02-15         1       155         0         0         0         0
##    2006-03-06         1       153       127       150         0         0
##    2006-09-20         0         0         0         1         0         0
##    2006-09-21         0         0         0        25         1         0
##    2006-09-22         0         0         0        24        51         0
##    2006-09-23         0         0         0         0         1         0
##    2006-09-25         0         3       127        51         1         0
##    2006-09-26         0         0         1        32       108         0
##    2006-09-27         0         0         0         3         0         0
##    2006-09-28         0         0         0         2         0         0
##    2006-10-16         0         0         0         0         1         0
##    2006-10-25         0         3       121         0         0         0
##    2006-10-27         0         0         4       148       156         0
##    2006-10-30         0         0         0         0         2         0
##    2006-11-03         0         0         0         0         1         0
##    2007-01-05         0         0         0         0       162         0
##    2007-01-09         0         3       121         0         0         0
##    2007-01-10         0         0         1         0         0         0
##    2007-01-20         0         0         3       147         0         0
##    2007-03-05         0         3       127       142       161         0
##    2007-08-31         0         0         0         1       138         0
##    2007-09-28         0         0         0         1        34         0
##    2007-10-01         0         0         0         0         1         0
##    2007-10-02         0         0         0         0        49         0
##    2007-10-03         0         0         0         0        49         0
##    2008-01-30         0         0         0         0         0       133
##    2008-03-06         0         0         0       130         0        89
##    2008-08-28         0         0         0         1       145       129
##    2008-08-29         0         0         0         0         1         0
##    2009-11-02         0         0         0         0         0         0
##    2009-11-09         0         0         0         0         0         1
##
##             Cohort 7 Cohort 8
##    2003-10-22        0        0
##    2003-11-20        0        0
##    2003-12-16        0        0
##    2004-01-15        0        0
##    2004-05-20        0        0
##    2004-09-20        0        0
##    2004-10-26        0        0
##    2005-01-19        0        0
##    2005-05-04        0        0
##    2005-09-22        0        0
##    2005-09-23        0        0
##    2005-09-30        0        0
##    2005-10-21        0        0
##    2005-10-26        0        0
```

```
##    2005-10-28         0         0
##    2006-01-06         0         0
##    2006-01-19         0         0
##    2006-01-20         0         0
##    2006-01-25         0         0
##    2006-01-26         0         0
##    2006-01-27         0         0
##    2006-02-15         0         0
##    2006-03-06         0         0
##    2006-09-20         0         0
##    2006-09-21         0         0
##    2006-09-22         0         0
##    2006-09-23         0         0
##    2006-09-25         0         0
##    2006-09-26         0         0
##    2006-09-27         0         0
##    2006-09-28         0         0
##    2006-10-16         0         0
##    2006-10-25         0         0
##    2006-10-27         0         0
##    2006-10-30         0         0
##    2006-11-03         0         0
##    2007-01-05         0         0
##    2007-01-09         0         0
##    2007-01-10         0         0
##    2007-01-20         0         0
##    2007-03-05         0         0
##    2007-08-31         0         0
##    2007-09-28         0         0
##    2007-10-01         0         0
##    2007-10-02         0         0
##    2007-10-03         0         0
##    2008-01-30         0         0
##    2008-03-06         0         0
##    2008-08-28       168         0
##    2008-08-29         0         0
##    2009-11-02         0       117
##    2009-11-09       149         0
```

Annotations

We have manually annotated essays from the connecticut cohorts for grammatical and spelling errors. After several false starts, we settled on a scheme in which research assistants attempted to convert all writing into sentences that were grammatically well-formed. They accomplished this by indicating through three types of edits - additions, deletions, and substitutions. For each essay, they marked the spot of the edit through the use of symbols for each of the edit type. Additions were marked with $ symbols, deletions with ^ symbols, and substitutions with @ symbols, In addition, for additions and substitutions, they provided the new or replacement text.

For instance, here is an entry that required two substitutions:

Essay

```
## Having membership in a social group may be important to someone
## else because they probably don't like to be left out.
```

7

In the `annotated` column of the same table, we can find where the changes were marked, as well as what the 'fix' is.

```
## Having membership in a social group may be important to someone
## else because they @he or she@ probably don't @doesn't@ like to be
## left out.
```

Finally, the `corrected` column displays the essay with the suggested edits implemented.

```
## Having membership in a social group may be important to someone
## else because he or she probably doesn't like to be left out.
```

An example of an essay that includes deletions and substitutions:

Essay

```
## My grades are important to me because you need to have good grades
## so you can get into a good collge.
```

annotated

```
## My grades are important to me because you need to have good grades
## so ^so^ you ^you^ can @to@ get into a good collge. @college@
```

corrected

```
## My grades are important to me because you need to have good grades
## to get into a good college
```

And an example essay with all three edit types:

Essay

```
## Being smart and getting Good grades is important to me because if
## you don't do good you won't be able to succeed in life.  For
## instance if you don't do good in school you'll never get that job
## that you wanted so much and won't do good in life.  You won't be
## able to get enough money, or you would live in a poor
## neighborhood.
```

annotated

```
## Being smart and getting Good @good@ grades is @are@ important to
## me because if you don't do good @well@ $,$ you won't be able to
## succeed in life.  For instance $,$ if you don't do good @well@ in
## school $,$ you'll never get that job that you wanted so much $;$
## and ^and^ $You$ won't do good @well@ in life.  You won't be able
## to get enough money, or you would @will@ live in a poor
## neighborhood.
```

corrected

```
## Being smart and getting good grades are important to me because if
## you don't do well, you won't be able to succeed in life.  For
## instance, if you don't do well in school, you'll never get that
## job that you wanted so much; You won't do well in life.  You won't
## be able to get enough money, or you will live in a poor
## neighborhood.
```

Typically, the addition, deletion, or substitution referred to only the previous word. However, on occasion it was necessary to perform an edit on a group of words. In these instances, the group of words which the edit refers to are grouped by curly brackets (i.e. **{}**). Here is an example:

Essay

```
## The value might be important to someone else because it may give a
## feeling of superiority and joy to them, and it may help the
## feeling of intelligence because it may help with collage.
```

annotated

```
## The value might be important to someone else because it may give
## {a feeling} @feelings@ of superiority and joy to them, @.@ and
## @And@ it may help the feeling of intelligence because it may help
## with collage. @college.@
```

corrected

```
## The value might be important to someone else because it may give
## feelings of superiority and joy to them. And it may help the
## feeling of intelligence because it may help with college.
```

It should be apparent from even these few examples that the edits are not always perfect. However, it is also equally apparent that, on the whole, they dramatically improve the grammar and spelling for each essay. At this time, we have performed annotations for the windsor data only. We have nearly finished with this task.

- **Proportion of essays annotated for correction - Windsor data only**

```
##     FALSE
## -0.8921036
```

Demographics

- **Number of participants**

```
## [1] 3799
```

- **Number of participants per study**
```

```
## 
##  Awareness&Affirmation             Belonging               Blindness
##                91                       169                      35
##          CO BMI Study           CO/CA Latino             Connecticut
##               205                       348                    1306
##            Debate        Group Affirmation NeuroBioThreat Study 1
##               148                       105                     142
##           Physics   Political Affirmation                Surgeons
##               399                       111                     470
##      Women-Math Study
##               270
```

- **Breakdown of participant race by study**

```
## 
##                    Awareness&Affirmation Belonging  Blindness
##   African American                     5         2          6
##   Asian                                3        13          1
##   Asian American                       0         0          0
##   Black/Non-American                   0         0          0
##   Hispanic/Latino                      6         8          3
##   Middle Eastern                       0         0          1
##   Native American                      0         0          1
##   Other/Mixed                         10         9          3
##   Reported Unknown                     0         0          0
##   White                               65       137         18
## 
##                    CO BMI Study CO/CA Latino Connecticut Debate
##   African American            0            0         581      0
##   Asian                       1            0          19      0
##   Asian American              0            0          41      0
##   Black/Non-American          0            0           0      0
##   Hispanic/Latino            88          137         128      0
##   Middle Eastern              0            0           0      0
##   Native American             0            0           0      0
##   Other/Mixed                 0            0           8      0
##   Reported Unknown            2            0           0      0
##   White                     105          200         514      0
## 
##                    Group Affirmation NeuroBioThreat Study 1 Physics
##   African American                60                     96       0
##   Asian                            0                      0       0
##   Asian American                   2                      0       0
##   Black/Non-American               0                      0       0
##   Hispanic/Latino                  2                      6       0
##   Middle Eastern                   0                      0       0
##   Native American                  0                      0       0
##   Other/Mixed                      8                      2       0
##   Reported Unknown                 0                      0       0
##   White                           32                     33       0
## 
##                    Political Affirmation Surgeons Women-Math Study
##   African American                    66       16                3
##   Asian                                7      120               27
```

```
##   Asian American                         0        0            0
##   Black/Non-American                     3        0            0
##   Hispanic/Latino                        2       29            9
##   Middle Eastern                         0       18            0
##   Native American                        0        0            0
##   Other/Mixed                            2       70           11
##   Reported Unknown                       0        0            0
##   White                                  0      214          220
```

- **Breakdown of participant gender by study**

```
##
##                              f   m
##   Awareness&Affirmation     36   52
##   Belonging                 62   55
##   Blindness                 21   14
##   CO BMI Study             127   78
##   CO/CA Latino             178  159
##   Connecticut              675  617
##   Debate                     0    0
##   Group Affirmation         55   50
##   NeuroBioThreat Study 1    69   67
##   Physics                  116  283
##   Political Affirmation     59   39
##   Surgeons                 187  205
##   Women-Math Study         130  140
```

Outcomes

- **Total number of outcome observations**

```
## [1] 28803
```

- **Number of outcome type per study**

```
##
##                          Belonging Blindness CO BMI Study CO/CA Connecticut
##   ENGLISH GPA                    0         0            0     0        3350
##   FCME SCORE                     0         0            0     0           0
##   Lab Test 1                     0         0          205     0           0
##   Lab Test 2                     0         0          205     0           0
##   Lab Test 3                     0         0          205     0           0
##   MATH GPA                       0         0            0     0        5316
##   OVERALL GPA                    0        68            0     0        3371
##   OVERALL QUARTER GPA            0         0            0   898           0
##   OVERALL YEAR GPA               0         0            0   540           0
##   PHYSICS GRADE                  0         0            0     0           0
##   POST GPA                       0         0            0     0           0
##   SCIENCE GPA                    0         0            0     0        5170
##   STEM GPA                       0         0            0     0        5418
##   UCSB MATH ASSESSEMENT        117         0            0     0           0
##
```

```
##                         Group Affirmation NeuroBioThreat Physics
##    ENGLISH GPA                          0              0       0
##    FCME SCORE                           0              0     698
##    Lab Test 1                         104              0       0
##    Lab Test 2                         104              0       0
##    Lab Test 3                         103              0       0
##    MATH GPA                             0              0       0
##    OVERALL GPA                        724            578       0
##    OVERALL QUARTER GPA                  0              0       0
##    OVERALL YEAR GPA                     0              0       0
##    PHYSICS GRADE                        0              0     399
##    POST GPA                             0              0       0
##    SCIENCE GPA                          0              0       0
##    STEM GPA                             0            361       0
##    UCSB MATH ASSESSEMENT                0              0       0
##
##                         Women-Math Study
##    ENGLISH GPA                        0
##    FCME SCORE                         0
##    Lab Test 1                         0
##    Lab Test 2                         0
##    Lab Test 3                         0
##    MATH GPA                           0
##    OVERALL GPA                      692
##    OVERALL QUARTER GPA                0
##    OVERALL YEAR GPA                   0
##    PHYSICS GRADE                      0
##    POST GPA                         177
##    SCIENCE GPA                        0
##    STEM GPA                           0
##    UCSB MATH ASSESSEMENT              0
```

- **Outcomes by date - Windsor data only**

```
##
##                ENGLISH GPA MATH GPA OVERALL GPA SCIENCE GPA STEM GPA
##    2004-02-01            0      150           0         150      138
##    2004-06-30            0      150           0         150      150
##    2005-02-01            0      335           0         335      335
##    2005-06-30            0      335           0         335      335
##    2005-11-15          161      165         163         163      176
##    2006-02-01          160      500         164         500      514
##    2006-04-01          159      160         162         163      176
##    2006-06-30          158      502         161         500      514
##    2006-11-15          321      333         330         173      347
##    2007-02-01          330      485         330         485      500
##    2007-04-01          327      326         327         326      343
##    2007-06-30          326      482         326         482      498
##    2007-11-15          176      176         176         176      174
##    2008-02-01          176      176         176         176      174
##    2008-04-01          176      176         176         176      175
##    2008-06-30          176      161         176         176      175
##    2008-11-15          176      176         176         176      175
##    2009-02-01          176      176         176         176      173
```

```
##    2009-04-01           176        176           176          176        173
##    2009-06-30           176        176           176          176        173
```

Prompts

This file is still being created. The windsor data is finished, so we will limit our description to these data

- **Number of unique values combinations**

```
## [1] 26
```

- **Number of different prompts**

```
## [1] 39
```

- **Prompt type per Cohort**

```
##       Intervention_number Cohort 1 Cohort 2 Cohort 3 Cohort 4 Cohort 5
## 1                       1     C-11      C-1      C-1      C-1      C-2
## 4                       2      C-4     C-15     C-14      C-2      C-2
## 7                       3      C-8      C-1     C-17     C-17     C-17
## 10                      4      C-7      C-1      C-9      C-9      C-9
## 13                      5     C-10     C-14      C-1     C-13      C-1
## 16                      6      C-5     C-17     C-14     C-14      C-1
## 20                      7      C-6      C-9     C-17     C-17      C-1
## 175                     8     <NA>     <NA>      C-9      C-9     <NA>
## 330                     9     <NA>     <NA>     <NA>     C-13     <NA>
##       Cohort 6 Cohort 7 Cohort 8
## 1        C-12     C-18     C-16
## 4         C-3     C-16     <NA>
## 7        C-13     <NA>     <NA>
## 10       <NA>     <NA>     <NA>
## 13       <NA>     <NA>     <NA>
## 16       <NA>     <NA>     <NA>
## 20       <NA>     <NA>     <NA>
## 175      <NA>     <NA>     <NA>
## 330      <NA>     <NA>     <NA>


##       Intervention_number Cohort 1 Cohort 2 Cohort 3 Cohort 4 Cohort 5
## 2                       1     T-14     T-16      T-2     T-16     T-17
## 5                       2      T-6     T-15     T-16     T-17     T-17
## 8                       3     T-11     T-16     T-18     T-18     T-18
## 11                      4     T-10     T-16      T-9      T-9      T-9
## 14                      5     T-13     T-16     T-16      T-2      T-2
## 17                      6      T-7     T-18     T-16      T-2     T-16
## 21                      7      T-8      T-9     T-18     T-18     T-16
## 176                     8     <NA>     <NA>      T-9      T-9     <NA>
## 331                     9     <NA>     <NA>     <NA>      T-2     <NA>
##       Cohort 6 Cohort 7 Cohort 8
## 2         T-1     T-20      T-5
## 5        T-12      T-4     <NA>
```

```
## 8          T-2      <NA>      <NA>
## 11        <NA>      <NA>      <NA>
## 14        <NA>      <NA>      <NA>
## 17        <NA>      <NA>      <NA>
## 21        <NA>      <NA>      <NA>
## 176       <NA>      <NA>      <NA>
## 331       <NA>      <NA>      <NA>
```

The reader may notice that the treatments for Cohorts 7 and 8 do not entirely match those listed below. For Cohort 7, Intervention 2 and Cohort 8, Intervention 1, subjects were randomly assigned to two different types of prompts - some were about politicians and others were about friends and family.

C-1

Look at the values you picked as least important to you. Think about times when these values would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why these values would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 2

- Intervention Number(s): 1 , 3 , 4

Cohort 3

- Intervention Number(s): 1 , 5

Cohort 4

- Intervention Number(s): 1

Cohort 5

- Intervention Number(s): 5 , 6 , 7

C-2

Remember the two or three things from page 1 that you just circled as not very important to you? What are they? (Please list them) In a few sentences, describe why the things you circled would be important to someone else (like another student in your school, or a person you've heard about.) Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, write down the 2 or 3 things you circled on page 1 that are not very important to you.

Cohort 4

- Intervention Number(s): 2

Cohort 5

- Intervention Number(s): 1 , 2

C-3

Describe how you get home at the end of the day (by bus or by car). How long does it take you to get home? Do you have a snack when you get home? What time do you go to sleep?

Cohort 6

- Intervention Number(s): 2

C-4

In Social Studies we study things that are important to people. Put a circle around the ONE thing that is LEAST important to you in your life. Pick only ONE of them to circle. The way your day begins can influence the rest of your day. Did you eat breakfast before you got to school today? If so, what did you eat? Describe how you got to school (such as: on the bus, by car, or walked). How long did it take you to get to school?

Cohort 1

- Intervention Number(s): 2

C-5

In this assignment you will be answering several questions about your ideas, your beliefs, and your life. It is important to remember while you are answering these questions that there are no right or wrong answers. Please read carefully over this list of personal values and think about each of the values. Then circle the one value that is LEAST important to you. We understand that all of these values may be important to you. Even if you feel that two or more of the value are least important, please pick only ONE of them to circle. Look at the value you picked as least important to you. Think about times when this value would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why this value would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 6

C-6

In this assignment you will be answering several questions about your ideas, your beliefs, and your life. It is important to remember while you are answering these questions that there are no right or wrong answers. Please read carefully over this list of personal values and think about each of the values. Then circle the two or three values that are LEAST important to you. Even if you feel that several of the values are not very important, please pick only TWO or THREE of them to circle. Look at the values you picked as least important to you. Think about times when these values would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why these values would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 7

C-7

Last month, you wrote about how you get home from school. This time, write about how long it takes you to get from class to class by describing the shortest trip and the longest trip you have between classes.

Cohort 1

- Intervention Number(s): 4

C-8

Last month, you wrote about how you get to school in the mornings. This time, tell how you get home at the end of the day (by bus or by car). Do you have a snack when you get home? What time do you usually eat dinner? When do you do your homework?

Cohort 1

- Intervention Number(s): 3

C-9

Last time, you wrote about how you get to school in the mornings. This time, describe how you get home at the end of the day (by bus or by car). How long does it take you to get home? Do you have a snack when you get home? What time do you go to sleep?

Cohort 2

- Intervention Number(s): 7

Cohort 3

- Intervention Number(s): 4 , 8

Cohort 4

- Intervention Number(s): 4 , 8

Cohort 5

- Intervention Number(s): 4

C-10

Look at the group you picked as least important to you. Think about times when this group would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why this group would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 5

C-11

Look at the value you have picked as least important to you. Think about times when this value would be very important to someone else. In a few sentences, describe why this value would be important to someone else. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. In this assignment you will be answering several questions about your ideas, your beliefs, and your life. It is important to remember while you are answering these questions that there are no right or wrong answers. Please read carefully over this list of personal values and think about each of the values. Then mark the one value that is most important to you with the letter M and the one value that is least important to you with the letter L. We understand that all of these values may be important to you. But if you feel that two or more of the values are most (or least) important to you, please just pick one.

Cohort 1

- Intervention Number(s): 1

C-12

Look at the values you picked as least important to you. Pick one of them. Think about times when this value would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why this value would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, think about the value that you just picked. List the top two reasons why someone else (like another student at your school or a person you've heard about) would pick this as their most important value:

Cohort 6

- Intervention Number(s): 1

C-13

Look at the values you picked as least important to you. Think about times when these values would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why these values would be important to someone else (like another person at your school or a person you've heard about). Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, look at your least important values. List the top two reasons why someone else (like another student at your school or a person you've heard about) would pick these as their most important value:

Cohort 4

- Intervention Number(s): 5 , 9

Cohort 6

- Intervention Number(s): 3

17

C-14

Look at the values you picked as least important to you. Think about times when these values would be important to someone else (like another student at your school or a person you've heard about). In a few sentences, describe why these values would be important to someone else. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 2

- Intervention Number(s): 5

Cohort 3

- Intervention Number(s): 2 , 6

Cohort 4

- Intervention Number(s): 6

C-15

Look at the values you picked as least important to you. Think about times when this these values would be important to someone else. In a few sentences, describe why these values would be important to someone else. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 2

- Intervention Number(s): 2

C-16

Please circle ONE politician that you do NOT care about. Circle your answer below. Look at the person that you circled on PAGE 2. Think about how this person might be important to someone else. Write a few sentences that describe why this politician might be important to someone else, like someone at your school or a person you've heard about. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 7

- Intervention Number(s): 2

Cohort 8

- Intervention Number(s): 1

C-17

Sometimes, the way that the morning begins can affect a person's day. What time did you get up today? How long did it take to get ready? If you had breakfast, what did you eat? How did you get to school (bus, car, walk)? In the space below please answer each of these questions about what you did before you got to school this morning. Feel free to write as much or as little as you want.

Cohort 2

- Intervention Number(s): 6

Cohort 3

- Intervention Number(s): 3 , 7

Cohort 4

- Intervention Number(s): 3 , 7

Cohort 5

- Intervention Number(s): 3

C-18

Version A: Look at the things you picked as NOT IMPORTANT on PAGE 1. Write a few sentences describing why these things would be important to someone else. Focus on your thoughts and feelings, and don't worry about spelling or how well written it is. Please write about someone else, like someone at your school or a person you've heard about. Again, think about the things you picked as NOT important on Page 1. List the top two reasons why these things would be important to someone else, like another student at your school or a person you've heard about:

Cohort 7

- Intervention Number(s): 1

C-19

Version B: Look at the things you picked as NOT IMPORTANT on PAGE 1, and CHOOSE ONE. Write a few sentences describing why this one thing would be important to someone else. Focus on your thoughts and feelings, and don't worry about spelling or how well written it is. Please write about someone else, like someone at your school or a person you've heard about. Again, think about the thing you just wrote about. List the top two reasons why this thing would be important to someone else, like another student at your school or a person you've heard about:

Cohort 7

- Intervention Number(s): 1

T-1

Look at the values you picked as most important to you. Pick one of them. Think about times when this value was or would be very important to you. In a few sentences, describe why this value is important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, think about the value that you just picked. List the top two reasons why this value is important to you:

Cohort 6

- Intervention Number(s): 1

T-2

Look at the values you picked as most important to you. Think about times when these values were or would be very important to you. In a few sentences, describe why these values are important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, look at the values you picked as most important. List the top two reasons why these values are important to you:

Cohort 3

- Intervention Number(s): 1

Cohort 4

- Intervention Number(s): 5 , 6 , 9

Cohort 5

- Intervention Number(s): 5

Cohort 6

- Intervention Number(s): 3

T-3

A role model is someone who you admire and want to be like, someone who is like a hero to you, and who does good things for people. Please circle the ONE politician that you think is a role model for you, or the one politician that you like most. Circle your answer below. Look at the person that you circled on PAGE 2. Write a few sentences that describe why this person is important to you. Maybe this person does good things for other people. Maybe this person cares about some of the same things that you do. Maybe you admire this person, or see him or her as a hero. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 7

- Intervention Number(s): 2

Cohort 8

- Intervention Number(s): 1

T-4

A role model is someone who you admire and want to be like, someone who is like a hero to you, and who does good things for people. Think of your role model or hero. Is this person a family member, a friend, a musician, a coach or athlete, a teacher, a musician, or an artist or actor, or someone else? Circle your answer below. Look at the role model that you circled on PAGE 2. Write a few sentences that describe why this role model is important to you. Maybe this person does good things for other people. Maybe this person cares about some of the same things that you do. Maybe you admire this person, or see him or her as a hero. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 7

- Intervention Number(s): 2

T-5

A role model is someone who you admire and want to be like, someone who is like a hero to you, and who does good things for people. Think of your role model or hero. Is this person a family member, a friend, a musician, a coach or athlete, a teacher, a musician, or an artist or actor, or someone else? Circle your answer below. Look at the role model that you circled on PAGE 2. Write a few sentences that describe why this role model is important to you. Maybe this person does good things for other people. Maybe this person cares about some of the same things that you do. Maybe you admire this person, or see him or her as a hero. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 8

- Intervention Number(s): 1

T-6

In Social Studies we study things that are important to people. Put a circle around the ONE thing that is MOST important to you in your life. The things listed here might all be important to you, but pick only ONE of them to circle. Think about the one thing that you circled. Why is it important to you? What good things happen to you because of this part of your life? Give one or two examples of how what you circled helps you from day to day (for example: you made a new friend or it makes you feel good).

Cohort 1

- Intervention Number(s): 2

T-7

In this assignment you will be answering several questions about your ideas, your beliefs, and your life. It is important to remember while you are answering these questions that there are no right or wrong answers. Please read carefully over this list of personal values and think about each of the values. Then circle the one value that is MOST important to you. We understand that all of these values may be important to you. Even if you feel that two or more of the values are most important, please pick only ONE of them to circle. Look at the value you picked as most important to you. Think about times when this value was or is very important to you. In a few sentences, describe why this value is important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 6

T-8

In this assignment you will be answering several questions about your ideas, your beliefs, and your life. It is important to remember while you are answering these questions that there are no right or wrong answers. Please read carefully over this list of personal values and think about each of the values. Then circle the two or three values that are MOST important to you. We understand that all of these values may be important to you. Even if you feel that many of the values are most important, please pick only TWO or THREE of them to circle. Look at the values you picked as most important to you. Think about times when these values were or would be very important to you. In a few sentences, describe why these values are important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 7

T-9

Last fall, you chose Relationships with Friends and Family as one of your most important values. Now that it is almost springtime, please think about how your friends and family will be important to you during the spring. Write about one or two ways that you might be involved with your friends or family over the spring months. For example, what activities will you do together, or how will you help each other? Last fall, you chose Sports Ability as one of your most important values. Now that it is almost springtime, please think about how sports will be important to you during the spring. Write about one or two ways that you might be involved in athletics over the spring months. For example, what sports will you watch or play? Last fall, you chose Creativity as one of your most important values. Now that it is almost springtime, please think about how you will be creative during the spring. Write about one or two ways that you might be creative over the spring months. For example, what activities will you do that will let you be creative? Last fall, you chose Independence as one of your most important values. Now that it is almost springtime, please think about how being independent will be important to you during the spring. Write about one or two ways that you might be independent over the spring months. For example, what decisions will you make on your own, or what things will you do just for you? Last fall, you chose Being Part of a Group, as one of your most important values. Now that it is almost springtime, please think about how being part of a group will be important to you during the spring. Write about one or two ways that you might be involved with your group over the spring months. For example, what activities will you do with your group, or when will you spend time with your group? Last fall, you chose Music as one of your most important values. Now that it is almost springtime, please think about how music will be important to you during the spring. Write about one or two ways that you might be involved with music over the spring months. For example, what kind of music will you play or listen to? Last fall, you chose Religion as one of your most important values. Now that it is almost springtime, please think about how being religious will be important to you during the spring. Write about one or two ways that you might be involved with your religion over the spring months. For example, will you be part of any religious activities or celebrations? Last fall, you chose Art as one of your most important values. Now that it is almost springtime, please think about how art will be important to you during the spring. Write about one or two ways that you might be involved in art over the spring months. For example, will you work on any projects, or spend time drawing or painting?

Cohort 2

- Intervention Number(s): 7

Cohort 3

- Intervention Number(s): 4 , 8

Cohort 4

- Intervention Number(s): 4 , 8

Cohort 5

- Intervention Number(s): 4

T-10

Last month, you wrote about being humorous or funny would be important to you over your holiday break. Name one or two funny things you did or saw during vacation. Last month, you wrote about how art would be important to you over your holiday break. Name one or two artistic things you did during vacation, such as art projects or drawing for others. Last month, you wrote about how being trendy would be important to you over your holiday break. Name one or two ways you were trendy during vacation. Last month, you wrote about how religion would be important to you over your holiday break. Name one or two things you did that involved religion during vacation. Last month, you wrote about how music would be important to you over your holiday break. Name one or two things you did that involved music during vacation. Last month, you wrote about how being part of a group would be important to you over your holiday break. Name one or two things you did with your group during vacation. Last month, you wrote about how being kind to others would be important to you over your holiday break. Name one or two kind things you did for other people during vacation. Last month, you wrote about how being yourself would be important to you over your holiday break. Name one or two things you did during vacation that were true to yourself or were just for you. Last month, you wrote about how sports would be important to you over your holiday break. Name one or two athletic activities you watched or did during vacation. Last month, you wrote about how your friends and family would be important to you over your holiday break. Name one or two activities you did with your friends or family during vacation.

Cohort 1

- Intervention Number(s): 4

T-11

Last month, you wrote about how being close to your friends and family is important to you. Sage Park's holiday break starts next week. Think about how your friends and family will be important to you during your vacation. Name two or three ways that you will be involved with your friends or family over winter break (such as activities that you will do together or ways that you will help each other). Last month, you wrote about how being good at sports is important to you. Sage Park's holiday break starts next week. Think about how sports will be important to you during your vacation. Name two or three ways that you will be involved in athletics over winter break (such as athletic activities that you will watch or participate in). Last month, you wrote about how being yourself is important to you. Sage Park's holiday break starts next week. Think about how you will be yourself during your vacation. Name two or three ways that you will be true to yourself over winter break (such as activities that you will do just for you or things that show who you really are inside). Last month, you wrote about how being kind to others is important to you. Sage Park's holiday break starts next week. Think about how kindness will be important to you during your vacation. Name two or three ways that you will be kind to others over winter break (such as things that you will do for other people or ways you will act towards other people). Last month, you wrote about how being part of a group is important to you. Sage Park's holiday break starts next week. Think about how being part of a group will be important to you during your vacation. Name two or three ways that you will be involved in a group over winter break (such as activities that you will do with your group or time that you will spend with your group). Last month, you wrote about how music is important to you. Sage Park's holiday break starts next week. Think about how music will be important to you during your vacation. Name two or three ways that you will be involved with music over winter break (such as music that you will play or listen to). Last month,

you wrote about how religion is important to you. Sage Park's holiday break starts next week. Think about how being religious will be important to you during your vacation. Name two or three ways that you will be involved with your religion over winter break (such as activities or celebrations). Last month, you wrote about how being trendy is important to you. Sage Park's holiday break starts next week. Think about how keeping up with the latest trends will be important to you during your vacation. Name two or three ways that you will be involved with being trendy over winter break (such as clothes you will wear or music that you will listen to). Last month, you wrote about how being good at art is important to you. Sage Park's holiday break starts next week. Think about how art will be important to you during your vacation. Name two or three ways that you will be involved in art over winter break (such as projects you will do or time that you will spend drawing). Last month, you wrote about how being funny is important to you. Sage Park's holiday break starts next week. Think about how humor will be important to you during your vacation. Name two or three ways that you will be funny over winter break (such as telling jokes that make people laugh, or certain people who really like your sense of humor).

Cohort 1

- Intervention Number(s): 3

T-12

Last time, you chose Sense of Humor as one of your most important values. Now that it is almost springtime, please think about how a sense of humor will be important to you during the spring. Write about one or two ways that you might be funny, or enjoy laughing, over the spring months. For example, who will you laugh and tell jokes with the most, or who really likes your sense of humor? Last time, you chose Relationships with Friends and Family as one of your most important values. Now that it is almost springtime, please think about how your friends and family will be important to you during the spring. Write about one or two ways that you might be involved with your friends or family over the spring months. For example, what activities will you do together, or how will you help each other? Last time, you chose Sports Ability as one of your most important values. Now that it is almost springtime, please think about how sports will be important to you during the spring. Write about one or two ways that you might be involved in athletics over the spring months. For example, what sports will you watch or play? Last time, you chose Creativity as one of your most important values. Now that it is almost springtime, please think about how you will be creative during the spring. Write about one or two ways that you might be creative over the spring months. For example, what activities will you do that will let you be creative? Last time, you chose Independence as one of your most important values. Now that it is almost springtime, please think about how being independent will be important to you during the spring. Write about one or two ways that you might be independent over the spring months. For example, what decisions will you make on your own, or what things will you do just for you? Last time, you chose Being Part of a Group, as one of your most important values. Now that it is almost springtime, please think about how being part of a group will be important to you during the spring. Write about one or two ways that you might be involved with your group over the spring months. For example, what activities will you do with your group, or when will you spend time with your group? Last time, you chose Music as one of your most important values. Now that it is almost springtime, please think about how music will be important to you during the spring. Write about one or two ways that you might be involved with music over the spring months. For example, what kind of music will you play or listen to? Last time, you chose Living in the Moment as one of your most important values. Now that it is almost springtime, please think about how living for today will be important to you during the spring. Write about one or two ways that you might live for the moment over the spring months. For example, how will you enjoy today and not worry about yesterday or tomorrow? Last time, you chose Art as one of your most important values. Now that it is almost springtime, please think about how art will be important to you during the spring. Write about one or two ways that you might be involved in art over the spring months. For example, will you work on any projects, or spend time drawing or painting?

Cohort 6

- Intervention Number(s): 2

T-13

Look at the group you picked as most important to you. Think about time when this group was or is very important to you. In a few sentences, describe why this group is important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 5

T-14

Look at the value you have picked as most important to you. Think about times when this value was or would be very important to you. In a few sentences, describe why this value is important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 1

- Intervention Number(s): 1

T-15

Look at the values you picked as most important to you. Think about times when these values was were or is would be very important to you. In a few sentences, describe why these values is are important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 2

- Intervention Number(s): 2

T-16

Look at the values you picked as most important to you. Think about times when these values were or would be very important to you. In a few sentences, describe why these values are important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is.

Cohort 2

- Intervention Number(s): 1 , 3 , 4 , 5

Cohort 3

- Intervention Number(s): 2 , 5 , 6

Cohort 4

- Intervention Number(s): 1

Cohort 5

- Intervention Number(s): 6 , 7

T-17

Remember the two or three things from page 1 that you just circled as most important to you? What are they? (Please list them) In a few sentences, describe why the things you circled are important to you. Focus on your thoughts and feelings, and don't worry about spelling, grammar, or how well written it is. Again, write down the 2 or 3 things you circled on page 1 that are most important to you.

Cohort 4

- Intervention Number(s): 2

Cohort 5

- Intervention Number(s): 1 , 2

T-18

There are a lot of things that are important to people, things that make their lives better, happy, or special. For example, some people really enjoy and value the time they spend with their family or friends. Others think something they are good at is really important to them, like playing sports, drawing, or being a good listener for friends. In the space below, please describe something that's important to you in your life. Feel free to write as much or as little as you want.

Cohort 2

- Intervention Number(s): 6

Cohort 3

- Intervention Number(s): 3 , 7

Cohort 4

- Intervention Number(s): 3 , 7

Cohort 5

- Intervention Number(s): 3

T-19

Version A: Look at the things you picked as MOST important to you on PAGE 1. Think about times when these things were very important to you. Write a few sentences describing why these things are important to you. Focus on your thoughts and feelings, and don't worry about spelling or how well written it is. Again, think about the things you picked as most important on Page 1. List the top two reasons why these things are important to you:

Cohort 7

- Intervention Number(s): 1

Version B: Look at the things you picked as MOST important to you on PAGE 1, and CHOOSE ONE. Think about times when this one thing was very important to you. Write a few sentences describing why this one thing is important to you. Focus on your thoughts and feelings, and don't worry about spelling or how well written it is. Again, think about the thing you just wrote about. List the top two reasons why this thing is important to you:

Cohort 7

- Intervention Number(s): 1

Remaining tasks and Residual concerns

CO/CA Latino Study

The other dataset for which we have extensive longitudinal measurements is the data sometimes referred to as the CO/CA Latino data. While we have been able to include these data in our database, there are some properties that remain opaque because of the multiple files we received, each of which only contains partial data. For instance, in one excel file, the same participant ID numbers appear in both the California *and* the Colorado data. Of course, this doesn't mean that all of these individuals moved. Ordinarily, I would give them different ID numbers, but the individuals for whom this is true are missing most of their essays from the Colorado data. This is just an example of some of the problems associated with these data. **As a solution** I would suggest that whoever is responsible for the curation of these data simply reorganize them entirely in a way that is more intuitive. Kevin Binning sent me the majority of these data, and while I appreciate his prompt responses, it's difficult to make heads or tails of what he has sent me and explained over email. It would be much simpler if he could deliver a single file (or a set of relational files!) that contain the information I've described here.

Women & Math study

We are still missing the essays from this study. The last contact we had about this was with Krisitin Layous on 1/14/15. She indicated that the essays were still on hard copy and in Geoff's lab somewhere.

Affirmation choice

This refers to data from Silverman, Logel, & Cohen (2013, *JESP*). We obtained scanned copies of the essays, but have no demographic information, outcome information, nor do we have subject ID numbers for the scanned essays. Arielle Silverman has suggested that it may not be possible to obtain this information. We have made repeated requests from Christine Logel, but have not yet received anything. The last time of contact was 5/7.

Surgeons

These data (Salles, Cohen & Mueller, 2012 *Journal of the American College of Surgeons*) are largely not usable at the moment. The file we were sent is missing labels for the columns, so aside from some obvious ones (e.g. essays), it isn't clear which variable each column contains. Additionally, much of the data is missing and the essays are cut off in the middle of sentences.

Prompts

We will eventually add in the prompt text for all additional studies included

Values selected

We are just beginning to add a variable that will indicate which value the writer selected.

Annotated & Corrected essays

We will also soon add a second version of each essay that has been annotated for correct spelling and grammar. In order to best preserve the integrity of what the writer wanted to communicate, we used a limited approach, only annotating for additions (marked with `$$`), deletions (marked with `^^`), and substitutions (marked with `@@`) of words to create more grammatically well-formed sentences. Here is an example of an annotated essay:

Original:
`This morning I had no breakfast and I got to school by car.`

Annotated:
`This morning $,$ I had $did not have$ no ^no^ breakfast $,$ and I got to school by car.`

You can see that the annotated sentence not only introduces changes to create a more grammatically well-formed sentence, but *also* allows us to investigate the types of changes made.

Conclusions

This database provides a foundation on which a wide variety of analyses can be built. It is straightforward to use, simple to extend, and can be easily shared. While the data are already largely de-identified, they can be completely de-identified by simply not sharing the demographics file.

For further readings on the format and structure of relational databases such as this one, or for the process of data wrangling required to turn unruly data into something neat, refer to the following sources:

Gentzkow and Shapiro (2014) Code and Data for the Social Sciences: A Practitioner's Guide.
Lohr. (2014) For Big-Data Scientists, "Janitor Work" Is Key Hurdle to Insights. New York Times
Wickham (2014) Tidy data. Journal of Statistical Software.