

## **ActSds 4.0: Convert INTERACT Data File to SDS files for GSEQ**

Roger Bakeman  
Georgia State University

Vicenç Quera  
Universitat de Barcelona

### **ActSds version 4.0 for .act, .iact, and .xiact files**

November 15, 2016

This is an update to the Appendix for: Bakeman, R., & Quera, V. (2008). ActSds and OdfSds: Programs for Converting INTERACT and The Observer Data Files into SDIS Timed-Event Sequential Data Files. *Behavior Research Methods*, 40, 869–872.

Correspondence should be addressed to Roger Bakeman, Department of Psychology, Georgia State University, Atlanta, GA 30303, USA, or Vicenç Quera, Departamento de Psicología Social y Psicología Cuantitativa, Sección de Psicología Cuantitativa, Facultad de Psicología, Universidad de Barcelona, Campus Mundet, Paseo Valle de Hebrón 171, 08035 Barcelona, Spain.

Electronic mail may be sent via Internet to bakeman@gsu.edu or vquera@ub.edu.

Mirror web sites: [www2.gsu.edu/~psyra/gseq/](http://www2.gsu.edu/~psyra/gseq/) and [www.ub.edu/gcai/gseq/](http://www.ub.edu/gcai/gseq/)

## ActSds 4.0: Convert INTERACT Data File to SDS files for GSEQ

This write-up is intended for users who wish to convert data collected with Mangold International's INTERACT into SDS files (Bakeman & Quera, 2009)—files formatted according to Sequential Data Interchange Standard conventions (SDIS; Bakeman & Quera, 1992, 1995). The resulting SDS files can then be used with our Generalized Sequential Querier analysis program (GSEQ; Bakeman & Quera, 1995, 2009, 2011).

INTERACT data files are comma-separated values (CSV) or tab-delimited text files. They can be read by spreadsheet programs such as Excel, statistical packages such as SPSS, and other programs such as this one. ActSds can read data files in either format.

The conversion program converts files with the extension .ACT, .IACT, or .XIACT and lets you read in as many ACT files as desired (all with the same extension). Each file may contain one or more *groups* (level 1), each group may contain one or more sessions (*sets*; level 2), and each session contains coded events (level 3). You can then define factor and factor levels for each observation session, if you wish, and can edit the names you gave to your initial class and behavior codes, shortening them so that the resulting SDIS codes are more readable.

Technical: Beginning with ActSds 3.0, text files can be coded following ANSI (Windows default) or UNICODE conventions, that is, can read standard 8-bit ANSI, UTF-8 (used by later versions of INTERACT), UTF-16, but not UTF-32.

### 1 INTERACT text file format

Each INTERACT file defines a rows by columns grid and contains event data for one or more sessions (sets) nested in one or more groups. The first line contains the word "SYSTEM" and is ignored.

- 1.1 For ACT and IACT files, the second line contains names for column headings and begins with "Type <tab> Entry <tab> Exit <tab> Memo <tab> ..." followed by tab-delimited class names.

For XIACT files, a subsequent line begins with "Type <tab> Onset <tab> Offset <tab> Memo <tab> ..." followed by tab-delimited class names.  
([INDVAR] lines are ignored for now).

- 1.2 For ACT and IACT files, the following lines contain data and begin with "S" for *group* (formerly scene, name in memo column), "T" for *set* (formerly take, name in memo column), or "E" for *event*.

For XIACT files, the following lines contain data and begin with "1" for *group* (formerly scene, name in memo column), "2" for *set* (formerly take, name in memo column), or "3" for *event*.

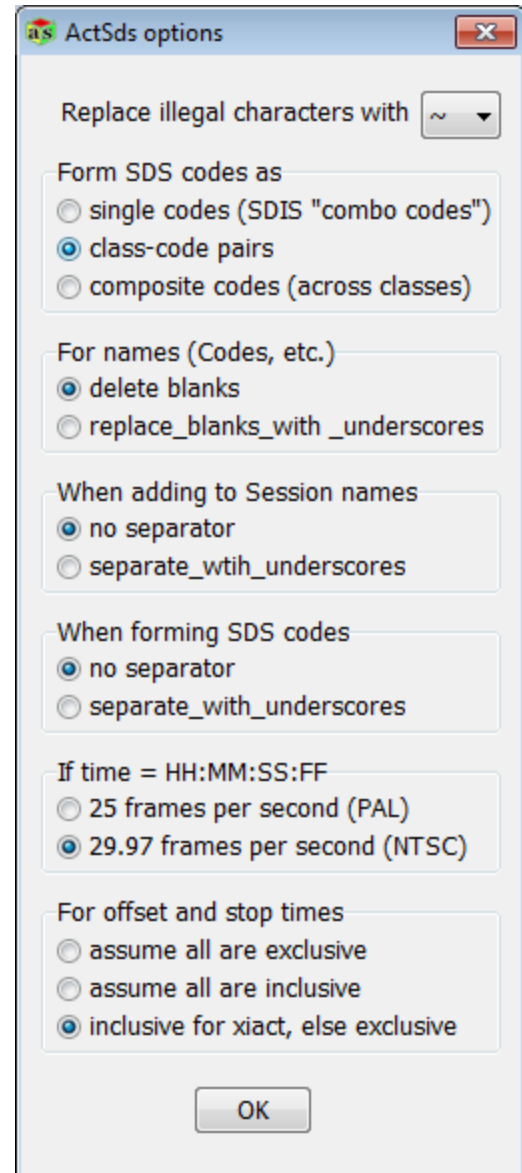
- 1.3 The one or more codes for each event are listed in the appropriate column, that is, in the column headed by the class to which that code belongs. (According to INTERACT conventions, each code belongs to a class.)

## 2 ActSds Options

Before opening or converting ACT files, first check the ActSds options (**File > Options**) in the main **ActSds** window.

- 2.1 By default, characters not allowed in SDIS codes are replaced with a tilde (~). The first option lets you change this to another character.
- 2.2 The second option let you specify how SDS codes are formed. As noted earlier, all codes are assigned to a particular *class* (*category* in older INTERACT documentation), and specific columns in the ACT file are reserved for each class. Each event line in an ACT file contains an onset time, an offset time, and the code or codes assigned that event—listed in the appropriate column. Which of the three options you chose depends on how you use classes and define codes in INTERACT.

- 2.2.1 If all codes are unique, no matter their class—and no matter whether events are often assigned just one or more than one code—it probably works best to select the *single codes* (SDIS "combo codes") option. Any codes assigned the same event are separated with plus signs. (Such codes—several codes with the same onset and offset times—are called *combo codes* in GSEQ.) With this option, a separate SDS code is formed for each of the codes assigned the event (although superordinate codes could be formed later with GSEQ's data modification capability).
- 2.2.2 If events are usually assigned a code from only one class, but some classes contain the same code, or if different classes often contain the same codes (e.g., classes = *Mom* and *Kid*, codes for both = *Look*, *Vocalize*, etc.), then it probably works best to select the *class-code pairs* option. With this option, a separate SDS code, whose name includes the class name, is formed for each of the codes assigned to an event.
- 2.2.3 If events are often assigned codes from several classes (in effect, events are cross-classified on several dimensions), then it probably works best to select the *composite*



*codes (across classes)* option. With this option, a single SDS code is formed for each event (although subordinate codes could be formed later with GSEQ's data modification capability).

- 2.3 The third option lets you specify whether any blanks in names (e.g., for Codes or Sessions) are deleted or replaced with an underscore.
- 2.4 If you opt to form new sessions names (e.g., by combining the file name with the initial session name), the fourth option lets you specify whether the names are separated with nothing or an underscore.
- 2.5 When names are combined to form codes (*class-code pairs* or *composite codes* option), the fifth option lets you specify whether SDS codes are formed with *no separator* or *separate\_with\_underscore*.
- 2.6 Prior to 2011 (.act and .iact files) times in ACT data files included a frame number. The sixth option indicates either 25 (generally Europe) or ~30 (generally US) frames per second.
- 2.7 Generally offset times in .act and .iact files are exclusive (if onset is 2 and offset is 5, duration is 3), whereas offset times in .xiact files are inclusive (if onset is 2 and offset is 5, duration is 4). If your files are an exception, you can select whether offset are all inclusive or exclusive. For further details, see Section 7.

### 3 Selecting Files to Convert

GSEQ expects that data for all sessions you wish to analyze will be in a single file. Thus the first task is to select the files you want to convert into an SDS file. To open one or more CSV files, select **File > Open** in the main window (or select the open icon on the tool bar).

### 4 Sessions and Research Factors

- 4.1 Once you have selected your INTERACT files, a second window opens. It lists the file names, the group name or names in each file, and the session (set) names in each group, along with the number of events in each session. The session names are taken from the data file (the 5th column in the set line). You can add the file name to the session name (useful if each ACT file represents a separate session) or the group name to the session name (useful if you want the group identifier to be part of the session name), or edit the session name as you wish. The session names are used to label session in the SDS file.
- 4.2 Often sessions are embedded in a research design consisting of factors (either between- or within-subjects), each with two or more levels (e.g., sex with two levels, male and female, or age with three levels, 3-, 4- and 5-years). SDIS conventions let you declare up to 7 factors. For example, the SDIS declaration, which gives the data type and names of codes, could end with:

... \* sex (male female) age (1year 2year 3year);

indicating two factors, one with two and one with three levels. Then, the following SDIS representation for a session:

<subject #24> (female 2year) ... /

would indicate that the session with the identifier *subject #24* was for a 2-year-old female.

- 4.3 If you want sessions to be identified with factor levels, indicate the number of factors using the up-down buttons. A new row opens at the top of the grid with default factor names (Factor#1, etc.) and space is made in the table for each session's levels names. You can edit factor names and add level names for each session. The declaration in the SDS file will then contain the factor names and their levels and the sessions in the SDS file will indicate their levels.

If several sessions have the same factor level, you need only enter the level name for the first. ActSds will automatically fill in subsequent level names for that factor. This can be useful when sessions are ordered by factor levels.

You can also add the file name or the group name to the factor level name. If you select any level name cell for a factor and then right click—and after confirming the factor number—you can select menu items that let you replace level names for that factor with the session's file name or the group name. This can be useful if file names or group names reflect factor level names. You can also clear any level names already given for that factor.

## 5 Forming SDIS Codes

- 5.1 Selecting *Next* at the bottom of the window displays a two column grid. The first column contains the codes in your data file and a second column displays those same codes, but in a form you can edit. If you have selected *class-code pairs*, first class names are displayed and selecting *Next* will display behavior codes. If you have selected *composites across classes*, behavior codes are displayed immediately. You can then select *Convert*.

If you asked to form SDS codes with *no separator* or *separate\_with\_underscore*, **ActSds** forms SDIS codes by concatenating your class-code pairs or codes within an event. For example, if classes = *Mom* and *Kid* and codes for both = *Look*, *Vocalize*, etc., then an SDIS code could be *MomLook* or *Mom\_Look*, depending on which separator option for forming SDS codes you chose.

Some of your initial codes may be quite lengthy and concatenating makes the SDIS codes even more so. For that reason, it is often advantageous to edit your initial codes (the second column) before converting. Remember that by default SDIS codes are case sensitive, so *mother* and *Mother* are two different codes.

To avoid ambiguous SDIS codes, the codes in the edited list usually are, but are not required to be, unique; if they are not unique, you are warned and given the option of changing the edited list before proceeding.

- 5.2 Later, if you wish, you can form additional codes using GSEQ's flexible data modification commands. For example, the following GSEQ command:

OR MomActs = MotherSoothesVocalizes MotherIgnoresQuiet ... ;

would let you create a new code representing several different mother behaviors. For further details see Bakeman and Quera (2009, forthcoming).

- 5.3 SDIS codes are formed from letters, digits, and special characters, for example # ^ \_ [ ] ? and |. Specifically, 19 characters used for SDIS syntax may not be used; they are ! " \$ % & ' ( ) \* + , - . / : ; < = >. By default, **ActSds** replaces any of these 19 characters with a tilde (~) but you can change this to \_ # ? or | if you wish. **ActSds** also replaces any blanks with either nothing or an underscore (to change these defaults see #2).

## 6 Saving and Restoring Edited Versions of Codes

- 6.1 Like data cleaning in general, data conversion is almost always an iterative process. Likely you will run the conversion program not once and be done, but multiple times as you find and fix errors and other anomalies. You will edit most of the lists the conversion program prepares and understandably you won't want to redo the same edits every time you run the conversion program with the same or similar INTERACT data files.
- 6.2 If you want to save the edited versions of your class and behavior codes, select **File > Save** in the code editing window. Next time you run **ActSds** you can select **File > Open** to restore your previous edits. The extension used for the save file is *tdf*, for tab-delimited file. This tab-delimited text file can be edited with any word processor or Excel. In fact, you may find it easier to save and then edit this file with a word processor than to edit the tables within the conversion programs.

## 7 Time Rounding and Other Conversion Options

Before converting your file, you can specify whether to round time to the second or to the tenth, hundredth, or thousandth (i.e., millisecond) of a second, the line length for the SDS file produced, and whether to display the converted lines in the main window.

- 7.1 All times in the SDS file are represented as MM:SS, MM:SS.d, MM:SS.dd, or MM:SS.ddd, depending on which rounding option you selected, where *d* is a decimal digit and *MM* can be any number, including numbers larger than 59 (i.e., hours are not represented separately).
- 7.2 Your data files may contain times represented as HH:MM:SS:FF (.act and .iact files in which FF indicates PAL's 25 or NTSC's 29.97 frames per second). Data files produced by

more recent version of INTERACT (.xiact files) contain times represented as a string of unpunctuated digits—these are the number of seconds  $\times 10^{-7}$  (i.e., 100-nanoseconds), so an implied decimal point is located before the 7th digit from the end of the string. In any case, select the rounding option you think most reasonable.

- 7.3 As noted earlier (Section 2.7), generally offset times in .act and .iact files are exclusive (if onset is 2 and offset is 5, duration is 3), whereas offset times in .xiact files are inclusive (if onset is 2 and offset is 5, duration is 4). If you have .xiact files and accepted the default, or if you selected the inclusive-offset-time option, ActSds adds a frame to offset times to make them exclusive (1/25 second for PAL, 1/29.97 second for NTSC).
- 7.4 In many cases, rounding to seconds may be reasonable. However, if similar events occur only tenths of a second apart, the SDIS compiler may issue an error message like "code precedes or overlaps itself" when rounding with second precision. In such cases, select a rounding option with greater precision—one that doesn't result in SDIS compiler errors. After compiling the SDS file produced by **ActSds**, you may find it convenient to round times in that file to lesser (and more reasonable) precision using the rounding utility in GSEQ. See the GSEQ help file for details.

The advantage of rounding to a second is that the printer-style plots in the GSEQ Plot and Kappa routines are easier to read. The disadvantage of rounding to a second or a tenth of a second—which is less precise than the 1/25 or  $\sim 1/30$  second precision that video recordings provide—is that rounding errors may be introduced. Percentages for events that are mutually exclusive and exhaustive may not sum exactly to 100%. If you want to maintain the full precision provided by video recording, round to a hundredth of a second, but rounding to a millisecond does not seem justified by the precision afforded by video recordings.

## 9 SDIS Declaration

- 9.1 The first line of an SDS file begins with a type declaration, which in this case is *Timed* (SDIS allows for other possibilities). It is followed by a list of codes.
- 9.2 The conversion program lists codes alphabetically. GSEQ output will order analyses for codes as they appear in the declaration. If you want a different order, edit the declaration in the SDS file to reflect the order you wish.

## 10 Compile Errors

The SDIS format is precise and attempting to compile in GSEQ an SDS file produced by the conversion program may reveal anomalies and errors that previously went undetected. SDIS compile error message that can occur include:

- a. Error 51: Wrong time. In SDIS events must be time-ordered. This error occurs when a code begins earlier than a preceding code.
- b. Error 56: Wrong session offset time. The last code cannot have the same time as the

offset time. To fix, increment the offset time.

c. Error 60: Code precedes or overlaps itself. In SDIS the times for the same code may not overlap.

In particular, the SDIS compiler expects events to be time ordered. If events are not ordered by onset time in your ACT data file, sort them by time before conversion.

## 11 Problems

Please let us know of any errors or problems you encounter. If you encounter a problem, first try to get help locally; often a second pair of eyes proves beneficial. Otherwise, try to create a very small test file that illustrates the program and e-mail it to us, along with a description of the problem—we will try to find a solution.

## References

Bakeman, R., & Quera, V. (1992). SDIS: A sequential data interchange standard. *Behavior Research Methods, Instruments, and Computers*, 24, 554–559.

Bakeman, R., & Quera, V. (1995). *Analyzing Interaction: Sequential Analysis with SDIS and GSEQ*. Cambridge, UK: Cambridge University Press

Bakeman, R., & Quera, V. (2008). ActSds and OdfSds: Programs for Converting INTERACT and The Observer Data Files into SDIS Timed-Event Sequential Data Files. *Behavior Research Methods*, 40, 869–872.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge, UK: Cambridge University Press.

Perhaps the best resource for SDIS syntax and information about the capabilities of the GSEQ program is the help file that comes with GSEQ: **Help > Show GSEQ Help** in GSEQ.

Visit our websites for more information and to download GSEQ:

<http://www2.gsu.edu/~psyab/gseq/Download.html>

<http://www.ub.edu/gcai/gseq/Download.html>



## Examples

### 1. This ACT data file:

***** SYSTEM: Mom-Kid Interact *****  COLUMN: Mom COLUMN: Kid COLUMN: Exp 						
Type	Entry	Exit	Memo	Mom	Kid	Exp
S	00:00:00:00	00:00:00:00		male		
T	00:00:00:00	00:02:59:00		N27		
E	00:00:00:00	00:01:03:22				Present
E	00:00:12:08	00:00:42:12			KCry	
E	00:00:15:14	00:00:27:20		MLook		
E	00:00:17:15	00:00:21:29		Sooth		
E	00:01:53:16	00:02:12:08			Look	
E	00:02:09:09	00:02:12:21			KVoc	
E	00:02:11:19	00:02:19:13		MVoc		

With *single codes*, rounding to a second, and one item/line specified, produces this SDS file:

```
% File converted by ActSds...
Timed <seconds>
  Cry
  KLook
  KVoc
  MLook
  Present
  Sooth
  VMoc
* sex (male ) ;

<N27> (male )
, 0:0
Present, 0:0-1:4
Cry, 0:12-0:42
MLook, 0:15-0:28
Sooth, 0:18-0:22
KLook, 1:54-2:12
KVoc, 2:9-2:13
VMoc, 2:12-2:19
, 2:59
/
```

## 2. This ACT data file:

***** SYSTEM: Mom-Kid Interact *****  COLUMN: Mom COLUMN: Kid COLUMN: Exp 						
Type	Entry	Exit	Memo	Mom	Kid	Exp
S	00:00:00:00	00:00:00:00		male		
T	00:00:00:00	00:02:59:00		N27		
E	00:00:00:00	00:01:03:22				Present
E	00:00:12:08	00:00:42:12			Cry	
E	00:00:15:14	00:00:27:20		Look		
E	00:00:17:15	00:00:21:29		Sooth		
E	00:01:53:16	00:02:12:08			Look	
E	00:02:09:09	00:02:12:21			Voc	
E	00:02:11:19	00:02:19:13		Voc		

With *class-code pairs*, no separator when forming SDS codes, rounding to a second, and one item/line specified, produces this SDS file:

```
% File converted by ActSds ...
Timed <seconds>
  ExpPresent
  KidCry
  KidLook
  KidVoc
  MomLook
  MomSooth
  MomVoc
* sex (male female) ;

<N27> (male )
, 0:0
MomLook, 0:15-0:28
MomSooth, 0:18-0:22
MomVoc, 2:12-2:19
&
KidCry, 0:12-0:42
KidLook, 1:54-2:12
KidVoc, 2:9-2:13
&
ExpPresent, 0:0-1:4
, 2:59
/
```

SDIS requires that events be time ordered (i.e., ordered by their onset times), although separate streams, each time ordered, may be specified.

With the *class-code pairs* option, events for each class are placed in separate streams, each separated per SDIS convention with an ampersand (&).

SDIS streams can be a convenient representational option, but the data—whether in one time ordered stream or several streams—remain the same.

## 3. This ACT data file:

***** SYSTEM: Mom-Kid						
Interact *****  COLUMN:						
Init COLUMN: Focus COLUMN: Resp 						
Type	Entry	Exit	Memo	Init	Focus	Resp
S	00:00:00:00	00:00:00:00		male		
T	00:00:00:00	00:02:59:00		N27		
E	00:00:22:25	00:00:46:22		mom	new	none
E	00:00:57:08	00:01:04:12		mom	old	attend
E	00:01:33:15	00:01:41:29		kid	old	attend
E	00:02:09:09	00:02:17:21		mom	new	reject

With *composite codes*, separate with underscores when forming SDS codes, rounding to a second, and one item/line specified, produces this SDS file:

```
% File converted by ActSds ...
Timed <seconds>
  kid_old_attend
  mom_new_none
  mom_new_reject
  mom_old_attend
* sex (male female) ;

<N27> (Level#1 )
,0:0
mom_new_none,0:23-0:47
mom_old_attend,0:57-1:4
kid_old_attend,1:34-1:42
mom_new_reject,2:9-2:18
,2:59
/
```

With the composite codes option, each unique combination is a separate code.

With *single codes*, rounding to a second, and one item/line specified, produces this SDS file:

```
% File converted by ActSds ...
Timed <seconds>
  attend
  kid
  mom
  new
  none
  old
  reject
* sex (male female) ;

<N27> (Level#1 )
,0:0
mom+new+none,0:23-0:47
mom+old+attend,0:57-1:4
kid+old+attend,1:34-1:42
mom+new+reject,2:9-2:18
,2:59
/
```

With the *single codes* option, codes that characterize the same event remain separate codes, separated by plus signs (called *combo codes* in SDIS), but share the same onset and offset times.