

GroupSize

The GroupSize program was designed and written for the training and testing of a marine mammal observer's propensity for estimating school sizes. This serves two useful purposes: to train a new observer to estimate a total school size from various shifting patterns, and to test the observer to determine the bias the observer introduces into a school estimate.



GroupSize presents the viewer with two species represented by oval dots moving across the screen in randomized groups. At the finish of the display, the viewer can enter the estimates on each school size.

GroupSize was written in Microsoft¹ Visual Basic 5.0 and uses the WinG OCX module by TegoSoft². GroupSize will run on Microsoft Windows 95/98/ME/NT/2000 operating systems. It will also run on a Macintosh in a Virtual PC³ environment.

¹ Visual Basic and Windows are registered trademarks of Microsoft Corporation.

² WinG is a registered trademark of TegoSoft Incorporated.

³ Virtual PC is a registered trademark of Connectix Corporation.

Software

There are two programs necessary for GroupSize:

Groupsize.exe – The main program for displaying group data.

Unlock.exe – Masker program for group data.

Group Data File

A group data file contains the school size information necessary for GroupSize to run. A group data file needs to have a .grp extension, such as Data.grp. The format of the group data file is as follows:

A1
B1
C1
D1

A2
B2
C2
D2

A3
B3
C3
D3

.
. .
. .
. .
. .

where A is the total number of dots for species #1, B is the total number of dots for species #2, C is the average size for a subgroup (a small subset for each species that is distinct, but part of the whole), and D is the \pm random variation for each subgroup. A blank line must separate each set. There can be up to 100 sets. For example, here is a short group data file with 3 sets:

19
33
5
4

270

0

13

7

58

44

7

4

In the first set, species #1 would have 19 total dots and species #2 would have 33 total dots. The average school size is 5 and the \pm variation for the school size is 4. So each subgroup can range from a size of 1 ($5 - 4$) to 9 ($5 + 4$) where the variation is the random variable. Each subgroup is randomly placed on the screen and each dot within the subgroup is randomly placed within a cohesive grid.

In the second set, species #1 would have 270 total dots and species #2 would have not dots. The average school size is 13 and the \pm variation for the school size is 7. Therefore, each subgroup can range from 6 to 20.

The program can handle a maximum of 5000 dots and 3000 schools. The variation must be less than the average school size. One, but not both, species counts can be zero.

When the group data file is created, it must be masked using the Unlock program. The reason for this is so that a user can't look at the group data file to find the answers to the group sizes. Run Unlock.exe, click on the .grp file you wish to mask, enter the mask value, which is **99** and click the "Unlock" button. When you want to edit the group data file, run Unlock again and enter the same mask value. The program uses an XOR (exclusive-or) of the mask value with each byte of the group data file. Doing the XOR again on the same file will return the file to its original state. This mask value is hard coded into both programs so it currently cannot be changed.

So in order for GroupSize to run properly, the group data files must first be masked; otherwise the program will fail. There can also be multiple group data files. Each can be chosen when GroupSize first starts running. When the program is in training mode, the program will choose a set at random and continue picking random sets until the program is terminated. When in testing mode, the program will go through each set in order and will end when the estimates for the last set have been entered.

Spec.dat File

The Spec.dat file is a program specific file containing the dot size, color and timing values. This file is also masked using the above mask value. An example follows:

8	#Ellipse width in pixels
4	#Ellipse height in pixels
3	#Jitter in pixels
404040	#Species 1 color in hex
C0C0C0	#Species 2 color in hex
1	#Time On (Visible) minimum in .5 second increments
3	#Time On (Variable) additional time in .5 second increments
5	#Time Off (Invisible) minimum in .5 second increments
4	#Time Off (Variable) additional time in .5 second increments
8	#Speed of motion in pixels. 8 is default. Lower numbers mean slower motion

There is only one Spec.dat file and must be named as such. For high-resolution screens, a higher number for the height and width may be necessary so that the dots don't appear too small. Jitter is a random up/down movement of each subgroup when the dots reappear on the screen. So each time a subgroup reappears, it can be ± 3 pixels up or down from the last position (or any random number in between). The color is in hexadecimal RGB (red-green-blue) format. The first two digits are for red, the second two digits for green and the last two for blue. The value for each color can range from 00 to FF.

The Time On (Visible) value is the minimum time each subgroup will be visible (in 0.5 second increments). The Time On (Variable) is the amount of random additional time each subgroup can remain visible (in 0.5 second increments). Since the Variable value is the random part, a random value of 0 to n-1 is possible. Thus, for the above example, the Time On (Variable) value is 3, so the program can choose a random value between 0 and 2 (inclusive). The Time Off values work in the same fashion as the Time On values.

So for the above example, each subgroup will be visible for a minimum of 0.5 seconds before fading out. The maximum visible time will be 1.5 seconds:

$$(1 \text{ (minimum)} + 2 \text{ (highest variable value)}) * 0.5 = 1.5$$

Also, the minimum time each subgroup will not be visible will be 2 seconds. The maximum time each subgroup will not be visible will be 5 seconds:

$$(4 \text{ (minimum)} + 6 \text{ (highest variable value)}) * 0.5 = 5$$

The motion speed is in pixels and is the distance each subgroup will move in 0.5 seconds.

Operation

When you run GroupSize, you are presented with a file directory window. Locate and click on the group data file you want the program to use and click the Ok button. You will then see the introduction window explaining the program and waiting for an Observer Identification Number. When the number is entered, you have the choice of running the program in either Training mode or Testing mode.

GroupSize

Welcome to GroupSize. The purposes of this program are (1) to improve your estimates of group size and (2) to measure your tendency to under- or overestimate group size.

You will observe two distinct species move across the screen, disappearing and reappearing at random intervals.

Below are two buttons: Training Mode and Testing Mode. Training Mode will pick school sizes randomly from a file and show you a pattern. At the end of each pattern, you will be asked for your school size estimates. You will then be given feedback on your estimate. In Testing Mode, you will not be given any feedback. To activate the buttons, enter your Observer Identification Number.

Observer Identification Number

Training ModeTesting Mode

Training mode runs the same as Testing mode except at the end when a window asks for your group size estimates. When all numbers are entered, you can click the “Compare” button to view the actual values. You can’t repeat the previous screen, but you can click “Ok” to go on to the next set, which is selected randomly. Clicking the “Exit” button from the Results window can stop the program.

Results

Species 1 = Species 2 =

Estimates		Actual (not active in Testing Mode)	
Best <input style="width: 40px;" type="text"/>	Species 1 % <input style="width: 40px;" type="text"/>	Total <input style="width: 40px;" type="text"/>	Species 1 % <input style="width: 40px;" type="text"/>
High <input style="width: 40px;" type="text"/>	Species 2 % <input style="width: 40px;" type="text"/>	Best Error <input style="width: 40px;" type="text"/>	Species 2 % <input style="width: 40px;" type="text"/>
Low <input style="width: 40px;" type="text"/>		High Error <input style="width: 40px;" type="text"/>	Species 1 % Error <input style="width: 40px;" type="text"/>
		Low Error <input style="width: 40px;" type="text"/>	Species 2 % Error <input style="width: 40px;" type="text"/>

Compare

OkExitRepeat

Testing mode goes through each set (from the Group Data File) in succession. At the end of a set display, you can repeat the previous screen display up to 2 times. Then the group size estimates must be entered and the next set is displayed when the “Ok” button is pressed. When all sets have been viewed and estimated, the program ends.

Based on the mode (training/testing) the user clicked at the beginning of the program will determine the extension for the output data file. Say for observer number 34, the output file can be either named 34.1 or 34.2. The .1 signifies the training mode was selected and .2 signifies the testing mode was selected. The main difference is that the 34.2 file is masked, and therefore needs Unlock to unmask the file. The 34.1 file is not masked and can be viewed by the user for review of the results.

In both training and testing modes, each set estimate is appended to the output data file. So if the same observer runs GroupSize multiple times in testing mode, all set estimates will be in the observer’s output data file.