



Manual for

CATSim

**Comprehensive Simulation of
Computerized Adaptive Testing**

February 2012

Version 4.0.6 and later



Table of Contents

In addition to the hyperlinked Table of Contents that follows, you may view the Table of Contents for this manual at any point in this manual by selecting the Bookmark icon or tab on the left side of the Acrobat window. The bookmark entries are hyperlinks that will take you directly to any section of the manual that you select.

Your CATSim License

CATSim is shipped in Demo mode. The demo is a fully functioning version of the software, but is limited to 50 examinees and 50 items. See the Appendix for further information about your CATSim license, unlocking your copy into a fully functioning version, and transferring your license to another computer

Technical Assistance

If you need technical assistance using CATSim, please visit the Support section of our Web site, www.assess.com. If the answer to your question is not posted, please email us at support@assess.com. Technical assistance for CATSim is provided for one year from the date you purchase or renew your license. Please provide us with the invoice number for your license purchase when you request technical assistance.

Citation

Weiss, D. J. & Guyer, R. (2012). *Manual for CATSim: Comprehensive simulation of computerized adaptive testing*. St. Paul MN: Assessment Systems Corporation.

Acknowledgments

The assistance of the following individuals in the preparation of this software is gratefully acknowledged:
Michael Finger, Benjamin Babcock, Nathan Thompson, Jeff Jones

**Copyright © 2012 by Assessment Systems Corporation
All Rights Reserved**

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written consent of the publisher.

Contents

CATSim: Comprehensive Simulation of Computerized Adaptive Testing.....	1
Post-Hoc and Hybrid Simulations	1
Monte-Carlo Simulations.....	2
1. Input Files	3
The Item Response Data File.....	3
Item Parameter File	6
Random Number Seed File.....	15
Item Selection Constraints Files	15
2. Output Files	16
Basic Output Files	16
User-Named Output Files	16
3. Options	20
The Simulation Type Tab	20
The IRT Model Tab	20
The CAT Options Tab	21
Initial θ	21
Estimation.....	21
Item Selection Options	22
The Constraints Tab.....	23
Content Balancing	23
Item Exposure.....	26
Enemy Items	27
The Termination Options Tab	28
The Monte-Carlo Options Tab.....	31
Randomly Generating Parameters	31
Appendix A. Technical Appendix.....	36
Dichotomous Model Equations	36
Response Probabilities.....	36
Item and Test Information	36
Polytomous Model Equations.....	37
Response Probabilities.....	37
Item Information.....	38
Graded Response Model and the Generalized Partial Credit Model (Difference Models).....	38
Rasch Rating Scale Model, Rasch Partial Credit Model, and Generalized Partial Credit Models (Divide-By-Total or Adjacent Category Models)	39
Equations for Estimating θ	39
Maximum Likelihood Estimation.....	39
Bayesian Estimation	39
Appendix B. Comparison of CATSim and Parscale θ Estimates	42
Index to Appendix B Tables	42
Comments on the Results in Tables B-5 through B-10	48
Appendix C: License, Unlocking, and Transferring	50
Your CATSim License and Unlocking Your Copy	50
License Transfer	53
Step 1 – Demo/Trial Program.....	54
Step 2 – Licensed Program.....	55
Step 3 – Demo/Trial Program.....	56

CATSim: Comprehensive Simulation of Computerized Adaptive Testing

CATSim implements three types of simulations for computerized adaptive testing (CAT) using both dichotomous and polytomous item response theory (IRT) models: post-hoc (real data) simulations, hybrid simulations, and monte-carlo simulations. In implementing a CAT program, all three types of simulation can be used at various stages of the CAT development process. CATsim options allow you to implement all three types of simulations varying CAT starting θ s, θ estimation methods, item selection methods, item exposure controls, and termination criteria. CATSim will implement simulations for item banks of up to 999 items, with no limit on the number of examinees for both post-hoc and hybrid simulations, and a limit of 10,000 examinees for monte-carlo simulations. However, CAT simulations can be done with as few as 200 examinees or fewer if they adequately represent the population to which the CAT will be applied. CATSim implements simulations for all three dichotomous IRT models and five polytomous models. CATSim includes all of the CAT options in version 3 of FastCAT (formerly the [FastTEST Professional Testing System](#)) so that the results of using CATSim can easily be implemented in your testing program.

Post-Hoc and Hybrid Simulations

Post-hoc simulation is an important final step prior to live implementation of a CAT. Post-hoc simulation allows you to evaluate the various CAT testing parameters prior to live testing, so that your live CAT will function optimally with the item bank that you have calibrated using an IRT model. A post-hoc simulation requires an item response matrix of real examinees responding to a CAT item bank for which item parameters have been estimated. The simulation then uses those item responses to simulate how that item bank would function if the items (for which responses are known) had been administered as a CAT. A post-hoc simulation can also be used with item response data from a conventional test to determine how much the test length could be reduced by administering the test as a CAT.

A significant problem in implementing post-hoc simulations with the relatively large item banks necessary for an adequate CAT – sometimes 250 or more items per bank – is that it is sometimes difficult or impossible to get a single group of examinees to respond to all the items in a bank. Consequently, CAT item banks are frequently constructed using linking procedures that include relatively short anchor or linking tests along with different subsets of items administered to different groups. The result is an item response matrix for an item bank that can be quite sparse, i.e., any examinee might have answered only 15% or 20% of the items in a bank – sometimes even fewer. This kind of data matrix cannot be used in a post-hoc simulation due to the large amount of missing data.

Hybrid simulations (Nydyck & Weiss, 2009) were developed to resolve this problem. A hybrid simulation is similar to a post-hoc simulation in that it uses an already calibrated bank – frequently the same sparse item response matrix used to estimate item parameters with a program

such as [Xcalibre](#) (Guyer & Thompson, 2012). To implement a hybrid simulation, the available set of item responses for each examinee are used to estimate that examinee's θ , skipping all items that were not administered to or answered by the examinee. The θ estimate is then used to impute that examinee's responses to the unadministered items using the appropriate IRT model and monte-carlo simulation methods. The result, then, is an item response matrix with complete data for each examinee (the initial real item responses supplemented by the model-fitting simulated item responses), which can be used in a post-hoc simulation. Nydick & Weiss demonstrated that the results of hybrid simulations with up to 87% imputed data yielded post-hoc simulation results that closely approximated those that were obtained from post-hoc simulations with a full matrix of real responses.

Monte-Carlo Simulations

Monte-carlo simulations are typically useful in the early stages of investigating the performance characteristics of CAT procedures when little or no data are available. A monte-carlo simulation allows you to quickly and efficiently vary different aspects of your data in conjunction with varying the parameters that control hypothetical CATs. CATSim allows you vary distributions of θ and distributions of item parameters, separately or in combination, by randomly generating these distributions using a specific IRT model. You may also fix the θ parameter and/or the item parameters, or read them from files. The result is the ability to answer a wide range of “what if” questions using assumed distributions of potential examinee θ distributions and potential item banks. Once CATSim generates a complete monte-carlo item response matrix under the conditions that you specify, the item response matrix is then analyzed by the same post-hoc simulation methods used for post-hoc and hybrid simulations.

1. Input Files

CATSim requires three input files: an item response data file, an item parameter file, and a random number seed file for implementing post-hoc and hybrid simulations. For monte-carlo simulations, the data file is generated by the monte-carlo process and an item parameter file – partial or complete – might be required, depending on the options chosen.

The Item Response Data File

The item response data file consists of item responses for your examinees preceded by four lines of control information. This file must be an ASCII/text file (*not* a word processor file) in the format required by the Assessment Systems Corporation [Item and Test Analysis Package](#) (ITAP). CATSim implements CAT for dichotomously or polytomously scored items, so the input data file can consist of item responses from multiple-choice tests or from Likert-type personality or attitude scale items (but not both types in the same file). All the item response data to be included in the analysis must be contained in a single input file. These files can have an extension of .DAT, .DATA, or any other extension that you prefer.

The file SAMPLE-DICHOT.DATA, in your CATSim installation folder includes data for 50 examinees from a 40-item multiple-choice test. The file SAMPLE-POLY.DATA includes data for 10 examinees from a 20-item rating scale.

An example of an item response data file of multiple-choice items in the proper input format is shown in Figure 1.1; these items will be scored using a dichotomous IRT model (1-, 2-, or 3-parameter). Figure 1.2 shows a portion of the input data file for items that use a polytomous IRT model.

**Figure 1.1 Example Item Response Data File
Containing Items to be Dichotomously Scored**

```
30 o N 5
143534243521132435241342351423 KEY
55555555555555555555555555555555 NO. ALTERNATIVES
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY ITEMS TO INCLUDE
EX001543542143554321542345134332413 EXAMINEE #1
EX002143534244522133002542531342513 EXAMINEE #2
EX003143534223521132435244342351233 EXAMINEE #3
EX004143534243521132435241342352NNN EXAMINEE #4
EX005143534243412132435452132341323 EXAMINEE #5
```

Figure 1.2. Example Item Response Data File Containing Items With Polytomous Responses

```

20 O N 4
+++++
55555555555555555555
YYYYYYYYYYYYYYYYYYYY
001 3232244222433333233
002 31111132112211232253
004 32232323224433244232
005 44522512112154335555
008 22334531434434233243
009 43233343333433434444
010 12131112312211132233

```

An item response file consists of five primary components:

1. A control line describing the data;
2. A line of keyed/correct responses for dichotomously scored items or a line with any characters (including blanks) for polytomously scored items (this line is ignored for polytomous items, but must be present);
3. A line with the numbers of alternatives for each of the items;
4. A line specifying which items are to be included in the analysis; and
5. The examinee data.

Comments may also be included in the item response data file. Each of these elements is described in the following sections.

The Control Line

The first line of the data file is the Control Line. It must contain the following data in the columns specified:

Column	Data
1-3	Number of items for which responses are recorded for each examinee (maximum is 999)
4	Blank
5	Alphanumeric code for omitted responses
6	Blank
7	Alphanumeric code for items not reached by the examinee
8	Blank
9-10	Number of characters of identification data recorded for each examinee (maximum is 80)

In columns 1-3, you must enter the number of items that are included in the file. *This number must be right-justified:* The “units” go into column 3, the “tens” in column 2, and the “hundreds” in column 1. Figure 1.1 shows a data file with 30 items to be analyzed; the example in Figure 1.2 includes responses to 20 multipoint (e.g., rating scale) items.

Column 5 must contain the alphanumeric code for items that the examinee has omitted. This may be a digit larger than the number of alternatives, a letter, or some other character, including a “blank.” For example, it might be “9” for a five-alternative item, an “O” for omitted, or a

period. Column 7 must contain the alphanumeric code for items that the examinee did not reach and, therefore, did not have a chance to answer. Like the omission code, it may be a digit larger than the number of alternatives or any other character. In Figures 1.1 and 1.2, the letter “o” indicates an omitted item, and “N” indicates a not-reached item.

Because operational CATs typically do not allow examinees to skip an item, for purposes of post-hoc CAT simulation CATSim considers all omitted or not reached dichotomously scored items as incorrect. However, for a hybrid simulation, responses for all omitted and not reached items are imputed. Therefore, if your item response data file includes skipped or omitted responses, you should not use post-hoc simulation – instead, use a hybrid simulation which will properly impute missing item responses. All imputation is done using the same method based on the IRT model selected.

Columns 9 and 10 contain the number of characters at the beginning of each examinee’s data record used for identification; this number must include any blank columns between examinee ID information and the beginning of the item responses. As with the number of items, *these digits must be right justified* — the “tens” must be in column 9 and the “units” in column 10. The maximum number of identification characters is 80. If columns 9 and 10 are left blank or if zero identification characters are specified, examinee identification will not be expected and the examinees’ responses must begin in column 1 on the data lines. The example in Figure 1.1 indicates that there are 5 characters of identification for each examinee; in the data lines (beginning on line 5 of the input file in Figure 1.1), you will note that examinees are identified by characters “EX001” through “EX005.” In Figure 1.2, there are four characters of examinee identification, including a space.

The Keyed Responses

The second line of the file contains the keyed/correct response for each item in the data file for items that are to be dichotomously scored. The code in column 1 corresponds to the key for Item 1, the code in column 2 corresponds to the key for Item 2, and so forth. The entire key must be contained on a single line. Thus, for the example in Figure 1.1, Item 1 is keyed “1,” Item 2 is keyed “4,” and the last item (Item 30) is keyed “3”. Note also the optional comment on the key line following item 30, which identifies the data on that line (e.g., KEY in Figure 1.1.) Optional comments on all lines must be separated by one or more spaces.

For dichotomously scored items, the key may be specified using the numerals 1 through 9 or the letters A through I. For example, a “1” means that all responses of “1” will be counted as correct. For convenience, “A” and “a” have been defined to be equivalent to “1”. Similarly, “B” = “b” = “2”. This equivalence continues through “I” = “i” = “9”. There is no letter equivalent to zero.

For polytomously scored items, the entries on this line are ignored, but the line must be present.

The Number of Alternatives

The third line of the file must specify the number of alternatives for each item; for dichotomously scored items, this is equal to the number of choices allowed for the item. In the example in [Figure 1.1](#), each of the items has five alternatives. If the item response data consists of already dichotomously scored items (i.e., scored 0, 1), the number of alternatives is “2” (and the corresponding key on the Keyed Response Line would be “1” for all items). For polytomous

items, enter the number of response alternatives for each item; this can differ among items for some polytomous models, whereas other models require that the all items have the same number of alternatives.

The Inclusion Code

The fourth line contains scale inclusion codes, which indicate whether an item should be included in the analysis. Items coded “Y” are included in the analysis; those coded “N” are not. In the example shown in Figure 1.1, all the items will be included in the analysis. Using these codes, subsets of items can easily be eliminated from an analysis.

The Examinee Data Lines

The examinee response data follow the fourth control line. The data for each examinee must be placed on a single line regardless of the number of items, and each examinee’s identification data must begin in the first column and continue through the number of characters you specified in the first line. For dichotomously scored items, any alphanumeric coding that corresponds to the omitted and not-reached codes in the first control line and to the range of legitimate responses specified in the third control line can be used to indicate the examinees’ responses. In Figure 1.1, the digits 1 through 5 were used for examinee responses. For polytomously scored items, numerical characters must be used, beginning with 1 for the first response, 2 for the second response, and so on, up to the maximum number of responses for each item.

Comments

Comments may be placed to the right of the data on any line. There must be at least one space between the data on any given line and a comment on that line. These comments are not used by the program.

Item Parameter File

For post-hoc and hybrid simulations, the item parameter files must follow the specifications below. For monte-carlo simulations, you can choose to fix or randomly generate some or all of the item parameters; in that case, parameters that are fixed or generated would not appear in the item parameter file. As a result, an item parameter file might not be required for dichotomously scored items. For polytomously scored items, however, an item parameter file is required that includes boundary locations for the items, but any parameters that are fixed or generated should not be included in the item parameter file.

Dichotomous Models

For dichotomous IRT models, CATSim assumes a 3-parameter logistic IRT model with $D = 1.7$ (the logistic approximation to the normal ogive) or $D = 1.0$ (the pure logistic model), using [Equation 1 \(Appendix A\)](#). You will select the appropriate value of D for your data on the IRT Model tab.

The item parameter file must consist of one line per item, and the number of lines in the file must equal the number of items specified in cols. 1 – 3 of the Control Line in the .DAT (or .DATA) file.

For the three-parameter model, there must be an a , b , and c parameter for each item in that order, separated by one or more spaces (except for parameters that are not read for monte-carlo simulations—see the [Monte-Carlo Options Tab](#)). For the two-parameter model, only the first two parameters (a and b) are required. If there is a third entry on the line for each item, it will be ignored. For the 1-parameter logistic (Rasch) models, specify a single value—the b parameter—for each item. Figure 1.3 shows the first ten lines of an item parameter file for dichotomous items, using a three-parameter model:

Figure 1.3. An Item Parameter File for 10 Three- or Two-Parameter Dichotomous Items

0.6891	0.6062	0.2374
0.5204	0.5360	0.2451
0.7612	-0.4503	0.2461
0.7269	-0.8308	0.2520
0.8024	-0.4112	0.2379
0.6982	-0.1783	0.2512
0.5178	-1.8573	0.2550
0.6380	0.5234	0.2424
0.6377	-0.8940	0.2555
0.6716	1.6200	0.2488

In addition to the item parameters, the item parameter file can include an item number (with no embedded spaces) before the item parameters and/or other identifying information after the item parameters, by selecting the one or both of the following options:

Item parameter file

Item numbers precede item parameters in this file

Item identifiers follow item parameters in this file (80 characters maximum)

If the item identifier option is selected, the item identifiers must follow the final parameter estimates, separated by one or more spaces. Be sure that your item parameter file is a pure ASCII text file (not a word processor file).

Two sample parameter files for dichotomously scored items are provided in your CATSim installation folder: SAMPLE 1.PAR includes only item parameters for 40 items. SAMPLE 2.PAR includes item parameters for the same 40 items, but the item parameters are preceded by item numbers and followed by other information. Sample parameter files for 20 polytomous items are also provided for each of the polytomous models.

Polytomous Models

CATSim implements CAT for five polytomous IRT models (references and equations for all polytomous models are provided in [Appendix A](#)).

Three IRT models are primarily appropriate for data collected using Likert-type and other rating scale formats that assume a set of ordered response categories:

1. Samejima's graded response model
2. Generalized rating scale model
3. Rasch rating scale model

Two additional polytomous models are generally used to analyze data that result from tests of ability, achievement, or proficiency:

4. Rasch partial credit model
5. Generalized partial credit model

Item parameter files for each of these models have different specifications. All item parameter files are simple text files (*not* word processor .DOC files) and are most easily found by CATSim if they have a .PAR extension. Item parameter files for polytomous models output by Xcalibre 4 (Guyer & Thompson, 2012) can be used in CATSim without modification.

One option for all the polytomous models will affect the data that are provided on the polytomous item parameter file. CATSim allows you to implement combined response categories as operationalized in Parscale (Muraki & Bock, 2002). If your Parscale analysis has been run with combined categories, and you have specified T weights to combine them, you will then have one or more fewer boundary locations than the usual number. You will then need to select the following option on the IRT Models tab,

Response weights (e.g., 1, 2, 3, 4, 4) are included in the parameter file for combining categories

and provide these response (T) weights to CATSim as indicated below.

These response weight are single-digit numbers beginning at 1, each separated by a one or more spaces. For example, response weights of 1 2 3 4 4 will combine the 4th and 5th response options for a 5-option items into a single category that will require 3 boundary locations. Response weights of 1 1 1 2 2 will combine a 5-option item into a 2-option item with 1 boundary location. When this is done, the number of boundary locations is reduced, but the number of response options specified below is still k . The file GPCM RECODED.PAR is an example of a parameter file for which 5-option items were recoded in Parscale to combine options 1 and 2 into a single category. Note that there are three boundary locations for the four options, but five option weights corresponding to the number of options ([see example below](#)).

Samejima's Graded Response Model (SGRM)

The SGRM is appropriate for items using Likert-type and other rating scales consisting of ordered category responses. The SGRM allows different numbers of answer categories within a set of items that measure a single construct. CATSim implements the homogeneous case of the SGRM which requires that the discriminations for each item are constant across the response options for that item, but allows the discriminations to vary across items.

The item parameter file requirements for the SGRM are as follows (Figure 1.4):

- ◆ The first line of the parameter file must have the letters SGRM beginning in column 1 to identify the file. Additional optional identifying information can appear following this identifier, provided that there is at least one blank space following SGRM.
- ◆ One line per item, in the order that the items appear in the examinee input data file, with the following information for each item with each separated by one or more spaces:
 - The number of response options for the item (k)
 - The item discrimination
 - The boundary locations for the item. For k response options for an item, there are $k - 1$ (or fewer) boundaries.
 - An optional set of response weights that can be used to combine response options ([see example below](#)).
 - An optional item description.

CATSim will read item boundary location parameters for the SGRM that have been estimated using Xcalibre 4 (Guyer & Thompson, 2012) or Parscale (Muraki & Bock, 2002). Xcalibre's boundary parameters for each item range from high negative to high positive. The boundary location parameters from Parscale range from high positive to high negative. Select the appropriate parameter scaling of your boundary parameters by checking the appropriate option shown below:

Polytomous Parameter Scaling

- Signs for the boundary parameters range from negative to positive as output by Xcalibre Version 4
- Signs for the boundary parameters range from positive to negative as output by Parscale

**Figure 1.4. Sample Item Parameter Input File for
the SGRM With 20 Five-Option Items
(Boundary Locations Range from High Negative to High Positive)**

```

SGRM GRM Parameters for 20 Items, D = 1.0
5 2.19412 -2.53855 -1.53745 -0.14199 1.26832
5 1.08005 -3.31876 -1.57415 0.47559 2.24129
5 2.5555 -1.99648 -1.12533 -0.00412 1.1258
5 1.15086 -2.50336 -0.79439 1.40688 3.09832
5 1.0222 -1.75364 0.22144 2.38014 4.29446
5 1.2666 -2.07024 -0.82333 0.486 1.55707
5 1.11364 -2.80133 -1.38031 0.87938 2.54808
5 0.64894 -1.48339 0.84443 3.05152 5.85643
5 0.60916 -2.96235 -0.16994 1.71862 5.24803
5 0.82204 -1.44754 0.4529 2.90243 6.51473
5 0.7592 -3.717 -1.05974 0.6021 2.65753
5 0.71664 -4.21596 -1.42568 0.72589 3.09217
5 2.79928 -2.23347 -1.32163 -0.06464 1.26148
5 1.08811 -1.98623 -0.58739 0.98086 2.64085
5 2.12879 -2.85833 -1.2047 0.02083 1.27427
5 1.76115 -3.33282 -1.93655 0.07621 1.62616
5 2.08687 -3.55556 -2.2058 -0.29227 0.83916
5 2.13419 -2.24118 -0.82544 0.36743 1.70146
5 1.13642 -4.33509 -3.25878 -0.18391 1.31244
5 1.61966 -2.8582 -1.66425 0.13597 1.37486

```

Generalized Rating Scale Model (GRSM)

The item parameter file requirements for the GRSM are as follows (Figure 1.5):

- ◆ The first line of the parameter file must have the letters GRSM beginning in column 1 to identify the file. This entry must be followed by
 - The number of response options (k) for all items (must be the same across items)
 - $k - 1$ (or fewer) boundary locations common to all items (decreasing from high positive to high negative)
 - An optional set of response weights that can be used to combine response options ([see example below](#)).
- ◆ One line per item, in the order that the items appear in the examinee input data file, with the following information for each item, with each entry separated by one or more spaces:
 - the item discrimination parameter (a)
 - the item location parameter (b)
 - An optional item description.

**Figure 1.5. Sample Item Parameter Input File
for the GRSM With 20 Five-Option Items
(Boundary Locations Range From High
High Positive to High Negative)**

GRSM	5	2.415	0.844	-0.867	-2.392
1.716	-1.054				
1.283	-0.539				
1.356	-0.781				
1.316	0.182				
1.291	0.952				
0.964	-0.356				
1.369	-0.068				
1.07	1.318				
0.908	0.534				
1.244	1.521				
0.941	-0.315				
1.079	-0.208				
1.842	-0.842				
0.968	0.211				
1.518	-0.839				
1.91	-0.904				
1.774	-1.158				
1.541	-0.242				
1.478	-1.266				
1.534	-0.654				

Rasch Rating Scale Model (RRSM)

The item parameter file requirements for the RRSM are as follows (Figure 1.6):

- ◆ The first line of the parameter file must have the letters RRSM beginning in column 1 to identify the file. This entry must be followed by
 - The number of response options (k) for all items (must be the same across items)
 - $k - 1$ (or fewer) boundary locations common to all items (increasing from high negative to high positive)
 - An optional set of response weights that can be used to combine response options ([see example below](#)).
- ◆ One line per item, in the order that the items appear in the examinee input data file, with the following information for each item:
 - Item location parameter (b)
 - An optional item description.

**Figure 1.6. Sample Item Parameter Input File
for the RRSM With 20 Five-Option Items
(Boundary Locations Range From High
Negative to High Positive)**

```
RRSM 5 -1.646 -0.829 0.688 1.788
-0.788
-0.33
-0.557
0.265
0.872
-0.125
-0.049
1.143
0.518
1.1
-0.097
-0.15
-0.481
0.285
-0.661
-0.576
-1.04
-0.153
-0.937
-0.63
```

Generalized Partial Credit Model (GPCM)

The item parameter file requirements for the GPCM are as follows (Figure 1.7):

- ◆ The first line of the parameter file must have the letters GPCM beginning in column 1 to identify the file. Additional optional identifying information can appear following this identifier, provided that there is at least one blank space following GPCM.
- ◆ One line per item, in the order that the items appear in the examinee input data file, with the following information for each item, with each separated by one or more spaces:
 - The number of response options for the item (k)
 - The item discrimination
 - The boundary/step locations for the item. For k response options for an item, there are $k - 1$ (or fewer) boundaries.
 - An optional set of response weights that can be used to combine response options ([see example below](#)).
 - An optional item description.

**Figure 1.7. Sample Item Parameter Input File
for the GPCM With 20 Five-Option Items
(Boundary Locations Range From High
Positive to High Negative)**

```

GPCM 20 multipoint items
5 1.536 2.263 1.526 0.141 -1.271
5 0.692 2.715 1.757 -0.681 -1.855
5 1.949 1.772 1.153 0.016 -1.084
5 0.753 2.137 1.063 -1.767 -2.557
5 0.659 1.433 -0.153 -2.543 -3.372
5 0.671 1.544 0.863 -0.754 -0.841
5 0.677 2.031 1.83 -1.319 -2.098
5 0.395 0.501 -0.907 -2.381 -4.408
5 0.348 2.393 -0.807 -0.572 -5.081
5 0.552 0.68 -0.19 -2.632 -5.892
5 0.394 3.826 0.377 -0.313 -2.078
5 0.396 4.004 1.161 -0.611 -2.581
5 2.304 2.031 1.318 0.095 -1.244
5 0.599 1.368 0.68 -1.158 -2.289
5 1.601 2.86 1.102 -0.088 -1.185
5 1.429 2.967 1.979 -0.142 -1.480
5 1.608 3.225 2.326 0.182 -0.725
5 1.585 2.134 0.719 -0.319 -1.649
5 0.754 2.298 4.375 -0.249 -0.825
5 1.054 2.374 1.955 -0.34 -1.154

```

Rasch Partial Credit Model (RPCM)

The item parameter file requirements for the RPCM are as follows (Figure 1.8):

- ◆ The first line of the parameter file must have the letters RPCM beginning in column 1 to identify the file. Additional optional identifying information can appear following this identifier, provided that there is at least one blank space following the item discrimination (e.g., “Reversed” in Figure 1.8).
- ◆ One line per item, in the order that the items appear in the examinee input data file, with the following information for each item with each separated by one or more spaces:
 - The number of response options for the item (k)
 - The boundary/step locations for the item. For k response options for an item, there are $k - 1$ (or fewer) boundaries.
 - An optional set of response weights that can be used to combine response options ([see example below](#)).
 - An optional item description.

**Figure 1.8. Sample Item Parameter Input File
for the RPCM With 20 Five-Option Items
(Boundary Locations Range From High
Negative to High Positive)**

RPCM Reversed				
5	-2.222	-1.709	-0.197	1.398
5	-2.178	-1.312	0.514	1.452
5	-1.676	-1.401	-0.074	1.212
5	-1.759	-0.821	1.398	2.078
5	-1.096	0.135	1.859	2.571
5	-1.265	-0.638	0.564	0.739
5	-1.652	-1.313	0.963	1.629
5	-0.423	0.473	1.336	2.463
5	-1.241	0.263	0.49	2.393
5	-0.55	0.183	1.748	3.821
5	-2.073	-0.315	0.257	1.238
5	-2.197	-0.65	0.369	1.451
5	-2.006	-1.686	-0.212	1.583
5	-1.079	-0.443	0.827	1.680
5	-3.195	-1.174	0.074	1.257
5	-3.032	-2.243	0.185	1.590
5	-3.37	-2.748	-0.123	0.686
5	-2.286	-0.763	0.294	1.832
5	-2.079	-3.443	0.176	0.700
5	-2.192	-1.9	0.353	1.062

A Sample Polytomous Parameter File With Response Weights

Figure 1.9 shows a sample polytomous parameter file for the GPCM in which response weights are used to combine the five response categories of each item into a smaller number of categories, by combining adjacent categories. Note that the boundary locations are in the order that they are output from Parscale (i.e., they range from high positive to high negative).

In the file in Figure 1.9, the “5” in the first column indicates that there are five options for each item, i.e., the range of item scores for each individual for each item can range from 1 to 5. The second entry on each line is the item discrimination. This is followed by the number of boundary locations estimated by Parscale for each item, after T weights were used for some items to combine categories. Thus, for items 2 and 3 there are only three boundaries since the following weights – 1 1 2 3 4 for item 2 and 1 2 3 4 4 for item 3 – were used to create four categories from the five options. Similarly, for item 14 there is only one boundary location since the weights 1 1 1 1 2 were used to combine the first four response categories into a single category, resulting in a dichotomous item. For the remaining items, no categories were combined, so there are four boundary locations followed by the response weights 1 2 3 4 5.

**Figure 1.9. Sample Item Parameter Input File for the
GPCM With 20 Five-Option Items and Response Weights
(Signs of the Boundary Locations Range From
High Positive to High Negative)**

```

GPCM
5 1.536 2.263 1.526 0.141 -1.27 1 2 3 4 5
5 0.692 2.715 1.757 -1.855 1 1 2 3 4
5 1.949 1.772 1.153 0.016 1 2 3 4 4
5 0.753 -0.2345 1 1 2 2 2
5 0.659 1.433 -0.153 -2.543 -3.372 1 2 3 4 5
5 0.671 1.544 0.863 -0.754 -0.841 1 2 3 4 5
5 0.677 2.031 1.83 -1.319 -2.098 1 2 3 4 5
5 0.395 0.501 -0.907 -2.381 -4.408 1 2 3 4 5
5 0.348 2.393 -0.807 -0.572 -5.081 1 2 3 4 5
5 0.552 0.68 -0.19 -2.632 -5.892 1 2 3 4 5
5 0.394 3.826 0.377 -0.313 -2.078 1 2 3 4 5
5 0.396 4.004 1.161 -0.611 -2.581 1 2 3 4 5
5 2.304 2.031 1.318 0.095 -1.244 1 2 3 4 5
5 0.599 1.289 1 1 1 1 2
5 1.601 2.86 1.102 -0.088 -1.185 1 2 3 4 5
5 1.429 2.967 1.979 -0.142 -1.480 1 2 3 4 5
5 1.608 3.225 2.326 0.182 -0.725 1 2 3 4 5
5 1.585 2.134 0.719 -0.319 -1.649 1 2 3 4 5
5 0.754 2.298 4.375 -0.249 -0.825 1 2 3 4 5
5 1.054 2.374 1.955 -0.34 -1.154 1 2 3 4 5

```

Random Number Seed File

The random number seed file is used by a random number routine. The file consists of a single line with three integer numbers, separated by spaces. For example,

```
15424 1113 21032
```

A file SEED.RAN is supplied as part of the CATSim installation and can be used as supplied, although you can create and use your own random number seed file.

The random number seed file is updated after each run, thus ensuring a different random sequence for each subsequent run. However, the starting values for a given run are reported on the summary output file for each run. If, for some reason, you need to exactly replicate a previous run, modify your random number seed file to use the random number seed values from the run you want to replicate. You may also specify any starting seeds that you desire for any run.

Item Selection Constraints Files

CATSim implements three types of item selection constraints that can be used in CAT administration: (1) content balancing, (2) item exposure, and (3) enemy items. Implementation of each of these constraints can require an additional input file. The structure of these optional input files is described below in the [Item Selection Constraints](#) section.

2. Output Files

CATSim creates three types of output files, some of which are optional: (1) basic output files, (2) a user-named non-optional output file, and (3) user-named optional output files.

Basic Output Files

These files are created for each run, but have the same name for each run. Two standard name output files are created for each run:

1. **INFOTBL.VAL.** This file has one row for each item and 121 columns. The columns represent θ values from -3.0 to $+3.0$ in increments of 0.05 . In each column are values of item information sorted from highest to lowest at each value of θ . This file is read-only.
2. **INFOTBL.TXT.** This file has the same structure as **INFOTABL.VAL**, but the entries are the item numbers corresponding to the sorted information values in **INFOTABL.VAL**. Thus, each column of this table identifies items in descending order of item information. This file is read-only.

Because the same names are used for these files during each run, if you want to save either of these files you (1) can rename them prior to a successive run with a different item bank, or (2) run analyses with different item banks in different folders. If you accidentally overwrite these files, simply re-run a simulation with the same item bank and the files will be re-created.

User-Named Output Files

These files all use the file name you supply for output files for a given run (e.g. FileName) on the Output Options tab, but differ in their extensions. The following files are produced for each run:

1. **FileName.summary.** This file is not optional and is the summary output file for each run. It includes the following information:
 - a. Details of all files used and all options selected.
 - b. Summary statistics for the run, including:
 - (1) Descriptive statistics for full-bank and CAT θ estimates and their SEMs
 - (2) Descriptive statistics for the differences between full-bank and CAT θ estimates and SEMs
 - (3) Correlations of full bank and CAT θ estimates and SEMs
 - (4) Descriptive statistics and frequency distribution for the number of items administered
 - c. An item exposure summary indicating for each item whether it was included in the CAT, the number and percent of uses of that item across the group of examinees, plus

- the item parameters and the scoring key for each item. If item exposure target values have been specified, they are reported as well.
- d. For monte-carlo simulations:
- (1) Descriptive statistics for true θ
 - (2) Descriptive statistics for the differences between true (generating) θ , and full-bank θ and CAT θ
 - (3) Correlations of true (generating) θ with full-bank θ and CAT θ
- e. Numerical values of the bank information function and model-predicted conditional standard errors of measurement (SEM) at values of θ from -3.0 to $+3.0$ in increments of $.05$, and the value and location of maximum information (and the associated minimum SEM) for the bank. The SEM values can be used to determine predicted SEM target values for a CAT. However, observed SEMs will differ from model-predicted SEMs to the extent that (1) θ estimates differ from true θ values and (2) examinee response patterns deviate from model-predicted response patterns (i.e., the extent to which real examinee responses do not fit the IRT model).

See [Appendix A](#) for the computational formulas for item and test information and the bank conditional SEM function.

The following user-named output files are optional:

2. ***FileName.examinee.txt or .csv***. This file contains summary data for each examinee. It is available as a .txt file in tabular (space-delimited) format, or as a .csv file that can be opened in a spreadsheet or statistical software for further analysis. It is recommended that you create this file for each run since it provides information that is useful for examining the performance of a CAT with a specific dataset and the options that you have selected. This file includes the following information for each examinee:
 - Full-bank θ estimate and its standard error (SEM). All θ estimates use the estimation method you select – maximum likelihood, EAP (expected a posteriori) MAP (maximum a posteriori or Bayesian modal), or weighted maximum likelihood. SEMs are observed SEMs computed using the second derivative of the maximum likelihood estimate or the Bayesian posterior variance. [Appendix A](#) provides equations for all θ estimation methods.
 - Number of CAT items administered.
 - CAT θ estimate and its SEM.
 - The difference between the two θ estimates and the SEMs.
 - If monte-carlo simulation has been selected, the true (generating) θ .
 - If the classification termination option has been selected, a classification is provided for each examinee.
3. ***FileName.detail***. This is the detailed output file for each run. It provides item-by-item results for each examinee, including: response; correct response; item score (1 = correct, 0 = incorrect, if dichotomous); item number; item parameters; item information value; and

full-bank and CAT θ estimates, their standard errors, and differences. If content balancing is selected, it also includes item-by-item content balancing results. With large numbers of examinees and/or items, this file can get quite large. If the classification termination option has been selected, a classification is provided for each examinee.

4. ***FileName.theta***. This file includes the final CAT θ estimates and their standard errors for each examinee, one line per examinee. If you have more than one test per examinee, these θ values (and, optionally, their standard errors) can then be used as starting θ estimates for another test in a following simulation.
5. ***FileName.info.bmp***. This file is a publication-quality graphic display of the information function for the entire set of items (i.e., the item bank information function). The values plotted are the sum of the values in each column of INFOTBL.VAL. Numerical values corresponding to the plotted points are output on the summary file (.summary) for each run. The bank information graphic is optionally displayed when each run is completed, but can be suppressed by unchecking the option on the Output Options tab. Note that this graphic will be constant for a given item bank, so it is not necessary to save it with each repeated run with a given item bank.
6. ***FileName.SEM.bmp***. This file is a publication-quality graphic display of the conditional standard error function for the entire set of items (derived from the item bank information function). Numerical values corresponding to the plotted points are output on the summary file (.summary) for each run. The bank standard error graphic is optionally displayed when each run is completed, but can be suppressed by unchecking the option on the Output Options tab. Note that it will be constant for a given item bank, so it is not necessary to save it with each repeated run with a given item bank.

The following four optional files are comma-separated-values (CSV) files with one line per examinee. They are designed to be imported into a spreadsheet or statistical software for further analysis. The entries in these files are the item-by-item values of the following variables after each item in the CAT, in the order that the items were administered, with one line of entries for each examinee:

7. ***FileName.theta.csv***. CAT θ estimates.
8. ***FileName.SEM.csv***. SEMs associated with each θ estimate.
9. ***FileName.scored.csv***. Item responses scored as correct (1) or incorrect (0).
10. ***FileName.items.csv***. The item number of the item administered.

The following file is optionally output from a hybrid simulation:

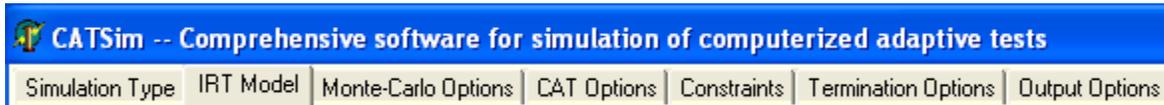
11. ***FileName.imputed.responses***. This file is the scored (for dichotomous items) or reweighted (for polytomous items) item response file after imputation of missing data based on a hybrid simulation. Following the item responses are the θ estimates (and standard errors) based on items actually answered by each examinee, which were used to impute missing data. With the addition of an appropriate ITAP header, this file then can be treated as a complete item response matrix, if desired, and run as a post-hoc simulation with a different set of CAT options from the hybrid simulation run that generated it, thereby eliminating the random effects from the hybrid imputation process.

The following files are optionally output from a monte-carlo simulation:

12. ***FileName.simulated.responses.*** This file is a completely formatted item response file resulting from a monte-carlo simulation run. Following the ITAP header, it has one line per simulee with item response scores: 1,0 for dichotomous items and 1 through the number of response options for polytomous models. The item responses are followed by the θ used for that simulee to generate them in conjunction with the item parameters for the specified model. This file can then be used as input for a subsequent post-hoc simulation run with different CAT options, if you do not want to introduce additional randomness into a simulation analysis.
13. ***FileName.simulated.thetas.*** This file contains the θ s simulated for a monte-carlo simulation run, one line per simulee.
14. ***FileName.simulated.parameters.*** This file contains the item parameters generated and/or used in a monte-carlo simulated run, fully formatted for input into another monte-carlo or post-hoc simulation. If used in another monte-carlo simulation run, it will allow generation of a new randomly generated set of item responses from a new set of θ s, or a previous set, with either the same set of CAT options or different CAT options.

3. Options

CATSim options are presented on six standard tabs, and a seventh Monte-Carlo Options tab which is activated if a monte-carlo simulation is selected. When the program begins, only the Simulation Type tab is active. Once you select a simulation type, the IRT Models tab will activate. After you select an IRT model, the remainder of the tabs will be active. The Run button below the tabs will activate after you select a file name for your output files. It is best to complete the options on each tab in the order that the tabs are presented.



The Simulation Type Tab

The Simulation Type tab provides for a choice among the three simulation types and allows you to specify the input files for the type of simulation that you have selected, as described in [Chapter 1](#).

The IRT Model Tab

As indicated above, CATSim implements simulations of all types for all three dichotomous IRT models and five polytomous models. You select your model on the IRT Models tab. For both types of models, you will need to select either $D = 1.0$ or $D = 1.7$ (for the Rasch-based models, this choice will be fixed at $D = 1.0$):

Model Constant

- $D = 1.0$ (pure logistic model)
- $D = 1.7$ (logistic approximation to the normal ogive)

For all polytomous models, there are two other options: Check the box shown below if your analysis that estimated the item parameters included T option weights for purposes of combining response options and reducing the range of weights assigned. If so, your item parameter file will need to have the option weights as part of the parameter input ([see Chapter 1](#)).

Response weights (e.g., 1, 2, 3, 4, 4) are included in the parameter file for combining categories

The second option concerns the scaling of your polytomous boundary location parameters. You need to inform the program as to whether your boundary location parameters range from positive to negative or vice-versa (see Chapter 1):

Polytomous Parameter Scaling

- Signs for the boundary parameters range from negative to positive as output by Xcalibre Version 4
- Signs for the boundary parameters range from positive to negative as output by Parscale

The CAT Options Tab

Initial θ

CATSim provides three options for beginning your CAT:

Initial Theta

Initial theta for all examinees is

Initial theta is random in the interval to

Read initial thetas and SEMs for each examinee from a file

Use individual SEMs as Bayesian prior SDs when Bayesian scoring is selected

1. Using the first option, all examinees will begin the CAT with the θ value specified. The valid range is -4.0 to $+4.0$.
2. The second option allows you to randomly start each examinee's CAT with a different θ value in the specified interval. The valid range is -4.0 to $+4.0$. This option can be used to reduce item exposure for the first few items in a CAT.
3. The third option allows θ values (and optionally, their SEMs) to be read from a file for each examinee, in the order the examinees appear in the input data file. This option is particularly useful in a situation in which you have more than one test for each examinee and want to use the final CAT θ estimate from one test as an entry point (initial θ) into the next test, in a following simulation run. In this application, you should use the *.theta* file output from the first test as input to the second.

If you use this option to input variable starting θ s for your examinees and do not select the variable SEM option, the standard deviation of the Bayesian prior θ s will use the value you specify as the Bayesian standard deviation (see below).

θ Estimation

CATSim provides three ways to estimate θ : Maximum likelihood, Bayesian, and weighted maximum likelihood.

Estimate Theta By

Maximum likelihood: Estimate theta for initial items for non-mixed response patterns using

A step size of theta/difficulty =

Bayesian estimation with a mean of and a standard deviation of

Bayesian estimation with a mean of and a standard deviation of

For Bayesian estimation, estimate theta by EAP (expected a posteriori) MAP (maximum a posteriori)

Weighted maximum likelihood

1. **Maximum likelihood.** When using maximum likelihood estimation, estimates cannot be obtained for single items or for item response strings that are all correct or all incorrect. In these circumstances in the administration of a CAT, you have two options:
 - Attempting to “force” a mixed response pattern (at least one correct and at least one incorrect) by selecting the next item to be more difficult for a correct response or less difficult for an incorrect response, using a specified step size on difficulty to select the next item. This arbitrary process is used until a mixed response pattern is obtained, at which point maximum likelihood estimation is used. The valid range of step size is .01 to 4.0. A larger step size will force a mixed response pattern more quickly than a small step size.
 - Use Bayesian estimation (see below) until maximum likelihood estimation can be used (i.e., until there is a mixed response pattern). For Bayesian estimation you will need to specify a mean and standard deviation of the Bayesian prior distribution on a scale with mean of 0.0 and SD of 1.0. The mean of the prior distribution can range from -4.0 to +4.0; the SD can range from 0.0 to 100.0. If you select Bayesian estimation, you will then need to choose between Bayesian modal (or MAP) and EAP estimation (see [Appendix A](#) for an explanation of the difference between these two Bayesian methods).
2. **Bayesian estimation.** For Bayesian estimation you will need to specify a mean and standard deviation of the Bayesian prior distribution on a scale with mean of 0.0 and SD of 1.0. The mean of the prior distribution can range from -4.0 to +4.0; the SD can range from 0.0 to 100.0. You will also need to choose between Bayesian modal (or MAP) and EAP estimation (see [Appendix A](#) for an explanation of the difference between these two Bayesian methods).
3. **Weighted maximum likelihood (WML).** WML is a variation of maximum likelihood in which the likelihood function is weighted by a function of the test information function. Because the likelihood function is weighted, it is able to provide a θ estimate based on a single item response or a non-mixed vector of item responses, similar to Bayesian methods. WML θ estimates, however, are not as biased as Bayesian estimates.

Item Selection Options

CATSim provides three item selection options:

Item Selection Option

Select all items by maximum information at the current theta estimate

Randomly from the items with maximum item information for the first items in the test

Select items in order of maximum information at theta =

1. **Maximum information.** This option is the classical CAT item selection option, using Fisher information. Items are selected at each stage of a CAT based only on the amount of information that they provide, excluding items that have already been used for a given examinee. This item selection method provides the fastest reduction in the standard error of measurement, resulting in the most efficient CAT.

2. *Exposure control maximum information.* The second item selection option provides one means of controlling the exposure of early items in a CAT (other item exposure controls can be implemented on the [Item Selection Constraints](#) tab). It uses maximum information item selection, but instead of selecting the unused item at a current θ estimate that provides maximum information, it allows you to have items selected randomly among a specified number of unused items with maximum information at the current θ estimate. For example, if you specify 10 as the first value for this option, items will be selected randomly from among the 10 (unused) items with maximum information at each current θ estimate. Furthermore, if you specify 10 also for the second value, this procedure will continue for the first 10 items administered; thereafter, items will be selected only by maximum information.
3. *Sequential testing.* Option 3 provides the capability of implementing sequential (rather than adaptive) testing. In sequential testing, items are ordered by maximum information at the specified θ value and are administered in that order. Sequential testing differs from adaptive testing in that the same θ value is used throughout a sequential test, whereas in adaptive testing the θ value from which items are selected is updated after each item is administered and items are selected by maximum information at each new θ estimate as it is calculated. Sequential testing using this approach has been used primarily in mastery/classification testing when a cutoff value on the θ scale has been specified (Spray & Reckase, 1994, 1996).

The Constraints Tab

CATSim implements three types of item selection constraints that can be used in CAT administration: (1) content balancing, (2) item exposure, and (3) enemy items. These options are selected on the Constraints tab. Note that use of any item constraints will reduce the efficiency of a CAT, with greater impact for smaller item banks. CATSim allows you to investigate the impact that these constraints will have on CAT using a specific item bank.

In implementing these constraints, content balancing is considered first to identify an item with the appropriate content classification. Once identified, the item is checked to see if it is in an enemy items set (if that option is chosen); if so, and an enemy item from that set has already been administered, it is not used further in the examinee's CAT. If it passes the enemy items constraint, it then is evaluated against the item exposure criterion (if selected).

Content Balancing

Content balancing is used when an otherwise unidimensional test is comprised of multiple content areas and it is desired that each examinee's CAT contain approximately similar proportions of items from each content area. The content balancing procedure used in CATSim is based on a procedure proposed by Kingsbury and Zara (1989). In this procedure, you first specify the desired target proportions for each content area in each examinee's CAT. As the CAT proceeds, the observed proportion of items in each content area is calculated and compared with the specified target proportions. The content area with the largest difference between the observed and target proportions is identified as the next content area for administration. The item selection algorithm then selects the item of that type with the maximum information at the current θ estimate.

To implement content balancing, first check the Content Balancing box on the Constraints tab. This will activate the Number of Content Categories box. Use the arrows to specify the number of content categories (up to 15) in your item bank and the appropriate number of Code and Prop. boxes will then activate. Specify a unique single alphanumeric character for each content category and the target proportions that you would like to approximate in each examinee's CAT. Note that the sum of the proportions must be 1.0 (± 0.01).

The screenshot shows a software interface with a yellow background. At the top, there is a checked checkbox labeled "Content Balancing: Number of Content Categories" with a dropdown menu set to "2". Below this, a text box explains: "For each category, enter a single-character alphanumeric code used on the input file to identify each item's content classification and the target proportion for that category. The sum of the proportions must equal 1.0." Below the text is a table with 10 columns: "Code" and "Prop." repeated five times. Each "Code" cell contains a small square input box, and each "Prop." cell contains a text box with "0.00". Below the table is a label "Select content balancing item input file:" followed by a file selection icon and an empty text box. Below that is a checked checkbox labeled "Exposure Control". Under "Exposure Control", there are two radio buttons. The first is selected and labeled "Use constant item exposure value for all items:" followed by a text box containing "0.000". The second is unselected and labeled "Use item-specific exposure values read from the following file:". Below this is a file selection icon and an empty text box. At the bottom, there is a checked checkbox labeled "Enemy Items:" followed by the text "Read list of enemy items from the following file:". Below this is a file selection icon and an empty text box.

Next, select a file that has content codes, using the same alphanumeric characters, for each of your items. This file must be an ASCII text file (you may use an extension of .CON) with one line per item. Each line must contain a (case-sensitive) alphanumeric code that matches one of the alphanumeric content codes specified on the Constraints tab, followed by a space, and then followed by an item number. Item numbers may appear in any order within the file. Any items for which there is not a content code entry will not be used in the content balancing process.

Figure 3.1 shows a sample of a portion of .CON file:

Figure 3.1 A Portion of a Content Balancing Input File

```

A 40
A 39
A 38
A 37
A 36
B 35
B 34
B 33
B 32
B 31
B 30
B 29
C 1
C 2
C 3
C 4
C 5
C 6
C 7
C 8
D 28
D 27
D 26
.
.
.

```

The results of content balancing are shown for each examinee at each stage of the CAT on the *.detail* output file, as shown in Figure3.2:

Figure 3.2. A Portion of the .detail Output File Showing Item-By-Item Content Balancing Results for a Single Examinee

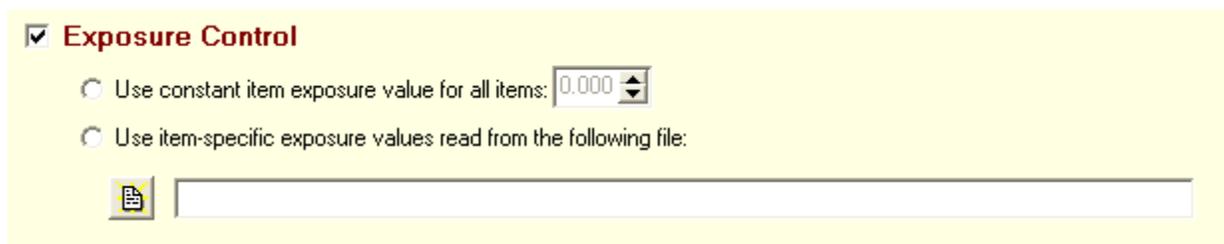
Seq #	Item #	...	Code	Content	Observed	Content	Proportions.....
1	3	. . .	C	A=0.000	B=0.000	C=1.000	D=0.000
2	17		D	A=0.000	B=0.000	C=0.500	D=0.500
3	30		B	A=0.000	B=0.333	C=0.333	D=0.333
4	36		A	A=0.250	B=0.250	C=0.250	D=0.250
5	21		D	A=0.200	B=0.200	C=0.200	D=0.400
6	16		C	A=0.167	B=0.167	C=0.333	D=0.333
7	19		B	A=0.143	B=0.286	C=0.286	D=0.286
8	18		A	A=0.250	B=0.250	C=0.250	D=0.250
9	28		D	A=0.222	B=0.222	C=0.222	D=0.333
10	5		C	A=0.200	B=0.200	C=0.300	D=0.300
11	31		B	A=0.182	B=0.273	C=0.273	D=0.273
	.						
	.						
	.						
	.						

Item Exposure

Item exposure controls are designed to limit, across a group of examinees, the proportion of times that each item is used in a CAT. This can be important in a “high-stakes” test used to make important decisions about examinees. In this type of testing situation, examinees might remember item content and pass it along to friends or distribute them in other ways, thus compromising item content.

CATSim implements item exposure controls using a probabilistic process, partially based on the work of Sympson and Hetter (Hetter & Sympson, 1999). CATSim implements two options for controlling item exposure—the target maximum exposure rate for items can be (1) the same for all items or (2) it can vary among items.

To use a constant maximum exposure rate for each item, select that option and specify the maximum proportion of times you would like each item to be used in a CAT across a group of examinees. For example, if you specify a constant rate of 0.25, any item that is selected by the CAT algorithm will, on average, be used in approximately one in four CATs.



Exposure Control

Use constant item exposure value for all items: 0.000

Use item-specific exposure values read from the following file:



To use item-specific exposure rates, create an item exposure data file and select it using the second Exposure Control option. This file contains one line per item with two entries: the first entry is an item number and the second is the desired item exposure proportion for that item. You need not include all items in this file—exposure values for any item not included will be set to 1.0, thus making it available without consideration of its exposure rate. This allows you to control exposure only for items that are over-exposed based on prior simulation results. Figure 3.3 shows an example of a portion of an item exposure file (the default extension is .TXT):

Figure 3.3. A Portion of an Item Exposure Input File

```
1 .40
2 .41
3 .42
4 .43
5 .44
6 .45
7 .46
8 .47
9 .48
10 .49
11 .50
```

The item exposure parameters for each item, in the Sympson-Hetter approach, are developed from monte-carlo simulations. They can, however, be specified based on other considerations.

To implement item exposure constraints, CATSim selects an item based on other item selection options. If the item is not eliminated by other constraints, and item exposure control has been selected, a random number is generated from a uniform distribution between 0.0 and 1.0. If the random number is greater than that item’s exposure control parameter, the item is not administered and will not be further considered for that examinee. If the random number is equal to or less than the item’s exposure control parameter, the item is administered. Thus, by this procedure the maximum exposure rate for any item will be approximately the specified rate, and frequently lower since not every item will be selected to be administered to each examinee.

The item exposure target for each item, and the actual number and proportion of times the item was selected in a simulation run is reported on the summary output file. Figure 3.4 shows a portion of that report using a bank of 40 items.

Figure 3.4. A Portion of the Item Exposure Output Report

Item Exposure Summary, Item Parameters, and Scoring Key				
Item Number		Admin. Freq.	Admin. Prop.	Target Prop.
-----		-----	-----	-----
1	...	266	0.404	0.400
2	...	287	0.436	0.410
3	...	279	0.424	0.420
4	...	317	0.482	0.430
5	...	282	0.429	0.440
6	...	294	0.447	0.450
7	...	357	0.543	0.460
8	...	287	0.436	0.470
9	...	383	0.582	0.480
10	...	316	0.480	0.490
11	...	334	0.508	0.500
12	...	327	0.497	0.510
.		.	.	.
.		.	.	.
.		.	.	.

Enemy Items

Enemy items are subsets of items that you specify that should not be administered to the same examinee. These might be items that provide clues from one item to another which might affect an examinee’s answers to an item in the set, or items that have very similar content and therefore might be redundant.

Sets of enemy items are specified in a text file, one line per set. Enter on each line the number of items in the enemy set followed by the item numbers of items in that set, with each entry separated by one or more spaces. Figure 3.5 shows an example of an enemy items set file:

**Figure 3.5. An Enemy Items Input File
With Three Sets of Enemy Items**

2 1 5
3 4 6 10
4 20 18 30 40

Three enemy item sets are specified in 3.5. The first set has two items—numbers 1 and 5. If either item is administered to an examinee, the other item will not be considered for that examinee. The second enemy item set has three items— numbers 2, 6, and 10. Administration of any of those items will cause the other items in that set to be skipped. The third enemy item set has four items: 20, 18, 30, and 40.

The Termination Options Tab

CATSim provides a number of different options for terminating a CAT. These include both variable-length and fixed-length termination.

Variable-Length Termination

Variable termination of a CAT allows the test length to vary across examinees. This is a major advantage of CAT over conventional tests because it allows the test developer to continue testing for each examinee until a pre-defined criterion of precision, operationalized by a number of termination options, has been reached. Which termination criterion, or combinations of termination criteria, should be used in a particular CAT depends on the purposes of the CAT and the characteristics of the item bank from which the CAT will be administered.

CATSim allows you to select multiple variable termination criteria or a single termination option. When multiple termination criteria are selected, an examinee's test will be ended when any of the multiple criteria has been met. This can be particularly useful for CATs from item banks that do not have horizontal information functions. In these banks, the standard errors will differ across θ levels. If a fixed standard error termination is used, test length will likely vary considerably across θ . However, using a fixed standard error termination in conjunction with another termination criterion (e.g., minimum information) will allow the termination criterion to vary with θ level, thus potentially avoiding very long CATs when a region of the item bank cannot support a given standard error termination criterion.

The termination criterion that is first satisfied will be recorded on the output file for each examinee, and a count of the number of times each termination criterion was used in a group of examines will be provided on the summary (*.summary*) output file. This information is also provided for each examinee on the detail (*.detail*) output file.

As shown below, there are six variable termination options provided in CATSim. All can also be used with a fixed minimum and/or maximum number of items to ensure that CATs for a given examinee are neither unusually short nor long.

Termination Options

- Variable Termination**
 - Terminate when the standard error of the theta estimate is less than or equal to
 - Terminate when the change in successive standard errors is less than or equal to
 - Terminate when the absolute change in successive theta estimates is less than or equal to
 - Terminate when the standard error of the theta estimate increases by or more
 - Terminate when the item information in an administered item is less than or equal to
 - Terminate when the theta estimate + or standard errors is above or below a theta cutoff value of
- Number of Items Constraints**
 - Require a minimum of items before terminating a test
 - Terminate when items have been administered

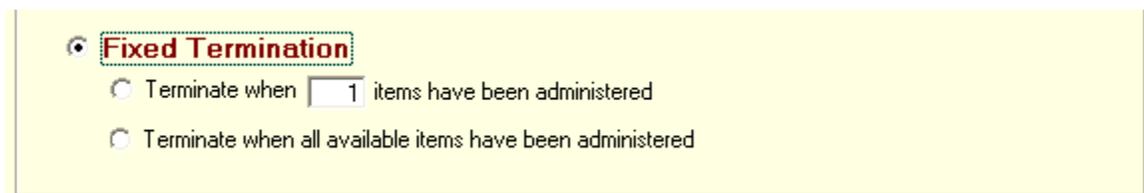
1. *Fixed standard error of the θ estimate.* This option allows you to control the standard error of the θ estimate (observed SEM), resulting in CATs that measure each examinee to a prespecified SEM, or “equiprecise” measurements. This is the variable termination option most usually applied in CATs and is most appropriate when the CAT item bank has a flat information function. When the bank information deviates substantially from being flat, fixed standard error termination should be combined with other termination criteria to ensure that CATs for examinees whose θ estimates are in regions of the item bank where there is less information do not exhaust the item bank in that region of θ .
2. *Change in standard errors.* A major characteristic of a CAT is that generally the standard error (SEM) of θ estimates decrease as each item is administered. Thus, a CAT can be terminated when the SEMs for an examinee fail to decrease by some small amount. Decreases in SEMs as a CAT converges tend to occur in the second or third decimal place with items with moderate discriminations, so termination values such as .01 or .005 might be appropriate as trial values for terminating CATs. Note, however, that there has been no research on using changes in SEMs as CAT termination values.
3. *Change in θ estimates.* Similar to the SEMs, a characteristic of a well implemented CAT is that the θ estimates for a given examinee tend to stabilize as a CAT progresses. Therefore, it might be appropriate to terminate a CAT when the absolute difference between successive θ estimates for an examinee is less than some value (such as .01 or .005). Babcock and Weiss (2009) report results from research using this termination criterion.
4. *Increase in the standard error of θ .* Occasionally an examinee’s CAT shows an increase in the SEM as the test progresses. This usually occurs if the examinee’s responses do not

fit the IRT model being used to estimate θ . Lack of fit can result from idiosyncratic examinee characteristics (e.g., an examinee whose first language is not English taking a test that is heavily English-based), inattentiveness, distraction, cheating, faking, or lack of cooperation. In these cases it might be appropriate to use this termination criterion in conjunction with others to terminate CATs for these examinees.

5. *Minimum item information.* CATs can be terminated when the information in the next item to be administered falls below a value you specify. Because of the relationship between item information and the model-predicted SEM (as determined from the inverse of the item information function), this approach is similar to using a SEM termination criterion, but it is not directly affected by examinee deviations from model fit as is the observed SEM termination criterion. Minimum information termination is particularly useful for tests that have information functions that are not approximately horizontal. In these cases it can be used alone or in conjunction with other termination criteria.
6. *Classification termination.* The last termination option is used with adaptive mastery/classification testing in which a cutoff value on θ is specified and the CAT is designed to classify individuals as above or below the cutoff value (Weiss & Kingsbury, 1984). This option allows you to vary the width of the confidence interval (in SEMs) around the estimated θ for each examinee. This confidence interval is used in the process of determining whether the examinee's θ estimate plus or minus the confidence interval is above or below the θ cutoff value.

Fixed-Length Termination

Two fixed-length termination options are available:



The screenshot shows a software interface for setting termination criteria. A radio button labeled "Fixed Termination" is selected. Below it, there are two options, each with a radio button. The first option is "Terminate when [1] items have been administered", where "1" is in a text input field. The second option is "Terminate when all available items have been administered".

1. *Administer a constant number of items to all examinees.* The first option allows you to administer a fixed-length CAT. When a fixed-length CAT is used, SEMs will likely vary across examinees and the θ estimates will not be equiprecise.
2. *Administer all the items in the bank to all examinees.* The second fixed-length termination criterion will administer all the items in the item bank as a CAT. Obviously, under these circumstances, the results for the CAT will be the same as for the entire item bank administered as a conventional test. This termination option might be useful if you output the item-by-item results files, import them into data analysis software, and examine the relationship between CAT results and full bank results on an item-by-item basis.

The Monte-Carlo Options Tab

The monte-carlo options tab appears only when a monte-carlo simulation is selected on the Simulation Type tab. It is activated when either a dichotomous or polytomous model is selected. This tab provides options for creating a pure monte-carlo dataset in which all parameters (θ and the item parameters) are randomly generated to your specifications, then a model-fitting item response matrix is generated from those parameters. It also allows you to fix some of the parameters and/or to read them from a file, then generate a model-fitting item response matrix. The model-fitting item response matrix is then used in a post-hoc simulation using the options you select on the other tabs.

You can fix any parameter by selecting the “Fix” button and specifying the constant value in the active box provided:



The screenshot shows the 'Theta' section of the Monte-Carlo Options Tab. It features a title 'Theta' in red. Below the title are three radio buttons: 'Fix' (selected), 'Read from file', and 'Generate'. To the right are five input fields with dropdown arrows: 'Alpha' (1.0), 'Beta' (1.0), 'Minimum' (-3.00), 'Maximum' (3.00), and 'Mean' (0.00). A 'Fix' checkbox is also present, which is currently unchecked.

You can read appropriate parameters from an input file by selecting that option and then selecting the file with the parameter values.

The “Theta” options are active for all models. For other models, only the appropriate item parameters will be active. The example below is for the three-parameter dichotomous model, so options are active for all three item parameters.



The screenshot shows the 'Item Parameters' section. It has a title 'Item Parameters' in red. Below the title are three rows of controls for 'Discrimination (a)', 'Location (b)', and 'Guessing (c)'. Each row has three radio buttons: 'Fix', 'Read from file', and 'Generate'. To the right of each row are five input fields with dropdown arrows: 'Alpha', 'Beta', 'Minimum', 'Maximum', and 'Mean'. A 'Generate' button is located at the end of each row.

Randomly Generating Parameters

For randomly generating parameters, CATSim uses the beta family of distributions. By specifying the appropriate values of the two parameters of the beta distribution—alpha and beta—you can generate a distribution of virtually any shape, from uniform/rectangular through normal, to very peaked, and virtually any kind of skewed distribution.. The text at the top of the Monte-Carlo Options tab provides information on how to use these two parameters to specify the distribution that you desire. Alpha and beta can be whole numbers or decimal numbers (e.g., 1.0 or 1.5).

Alpha and beta control the shape of the beta distribution. Setting alpha and beta equal creates a symmetric distribution. 5 creates a normal distribution. Setting alpha and beta to 1 creates a uniform distribution. The difference between alpha and beta controls the skew. Negative skew is increased as the difference between alpha and beta becomes more positive, whereas positive skew increases as the difference becomes more negative. The sum of alpha and beta controls the flatness. The distribution becomes more peaked (narrower tails) as the sum increases; the distribution becomes flatter (larger tails) as the sum decreases.

For example, the monte-carlo alpha and beta options selected below will generate a normal distribution for the a parameter in the range .50 to 1.50, a uniform distribution for the b parameter between -3.0 and 3.0 , and a negatively skewed distribution for c with a mean of approximately .13.

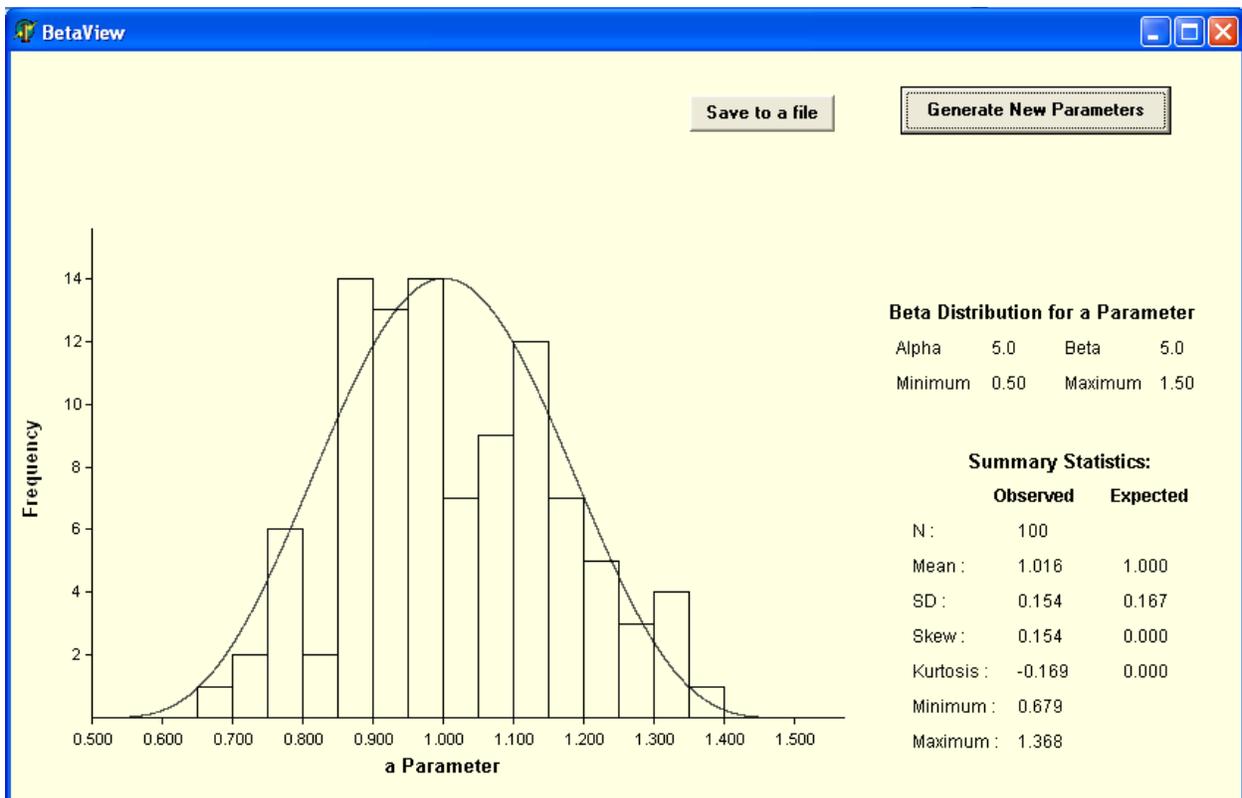
Item Parameters

Discrimination (a) Fix Read from file Generate 5.0 5.0 0.50 1.50 1.00 **Generate**

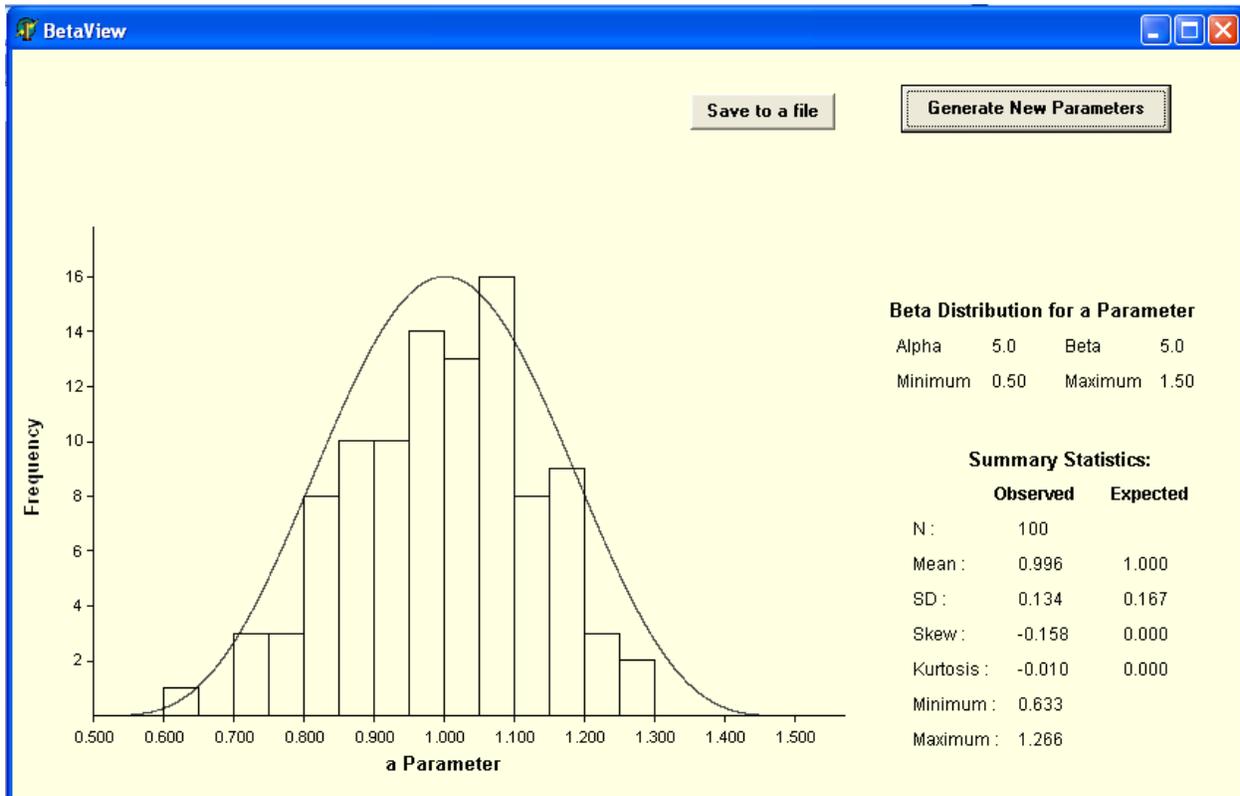
Location (b) Fix Read from file Generate 1.0 1.0 -3.00 3.00 0.00 **Generate**

Guessing (c) Fix Read from file Generate 4.0 1.0 0.00 0.25 0.13 **Generate**

Once you specify your beta distribution parameter values, you must click the **Generate** button to view the random set of values for that parameter. A graphic like that below will appear:



This graphic shows the theoretical expected beta distribution as a solid line and the observed generated distribution of the parameter as a bar graph. It also provides descriptive statistics for both the observed and expected distributions. If you are satisfied with the generated distribution, you might want to save it as a file for future reference before you close the graphic window. If you want a slightly different random distribution with the same specifications, click the **Generate New Parameters** button. For example, the following is another generated random distribution of the a parameter using the same specifications:



The last distribution that you view will be the distribution of the parameter used in your monte-carlo simulation (you cannot go back to a previous generated set of parameters). You need to click the Generate button for each parameter for which you have selected the Generate option.

If you are planning to run a number of monte-carlo simulations with the same beta specifications, you can save your alpha and beta selections in a file by selecting

Save Monte-Carlo Simulation Defaults

Then for subsequent simulations, select

Load Saved Monte-Carlo Simulation Defaults

The defaults are saved in a file with the name MonteCarlo.defaults.txt in the same folder as your input and output files. Thus, if you keep different datasets in different folders, you can have different defaults for different types of datasets.

Note that for polytomous models, although you can generate the relevant a and b parameters, you must read all boundary parameters from a file.

References

- Babcock, B. & Weiss, D. J. (2009). Termination criteria in computerized adaptive tests: Variable-length CATs are not biased. In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*. Available from www.psych.umn.edu/psylabs/CATCentral/
- Baker, F. B. & Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques* (Second Edition). New York: Marcel Dekker, Inc.
- Choi, S. W. (2007). *FIRESTAR: Computerized adaptive testing (CAT) simulation program for polytomous IRT models* (Version 1.2.2). Evanston IL: Evanston Northwestern Healthcare Research Institute.
- Guyer, R. D. (2009). *Comparison of a numerical derivative to the exact value for weighted maximum likelihood estimation*. Unpublished Manuscript.
- Guyer, R., & Thompson, N.A. (2012). *User's Manual for Xcalibre item response theory calibration software, version 4.1.6*. St. Paul MN: Assessment Systems Corporation. Available from <http://www.assess.com/>
- Hetter, R. D., & Sympson, J. B. (1997). Item exposure control in CAT-ASVAB. In W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (pp. 141-144). Washington DC: American Psychological Association.
- Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education*, 2, 359-375.
- Lord, F. (1983). Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. *Psychometrika*, 48, 233-246.
- Muraki, E. & Bock, R. D. (2002). *Parscale (Version 4)* [Computer software]. Lincolnwood IL: Scientific Software International.
- Nydick, S., & Weiss, D. (2009). A hybrid simulation procedure for the development of CATs. In D. J. Weiss (Ed.), *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*. Available from www.psych.umn.edu/psylabs/CATCentral/
- Ostini, R. & Nering, M. L. (2006). *Polytomous item response theory*. Thousand Oaks CA: Sage Publications.
- Samejima, F. (1993). An approximation for the bias function of the maximum likelihood estimate of a latent variable for the general case where the item responses are discrete. *Psychometrika*, 58, 119-138.
- Spray, J. A., & Reckase, M. D. (1994, April). *The selection of test items for decision making with a computer adaptive test*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans LA.

Spray, J. A., & Reckase, M. D. (1996). Comparison of SPRT and sequential Bayes procedures for classifying examinees into two categories using a computerized test. *Journal of Educational and Behavioral Statistics, 21*, 405-414.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*, 427-450.

Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement, 21*, 361-375.

Appendix A.

Technical Appendix

This Appendix includes response probability equations and information equations for the IRT models used in CATSim, and equations used for estimating θ and its standard error.

Dichotomous Model Equations

Response Probabilities

CATSim uses the following three-parameter logistic equation. For the two-parameter model, $c_i = 0.0$. For the one-parameter (Rasch) model, $c_i = 0.0$ and $a_i = 1.0$.

$$P_{ij} = c_i + (1 - c_i) \frac{\exp[Da_i(\theta_j - b_i)]}{1 + \exp[Da_i(\theta_j - b_i)]}, \quad (1)$$

where P_{ij} is the probability of a correct response to item i by person j

θ_j is the trait level for person j ,

a_i is the discrimination parameter for item i ,

b_i is the difficulty or location parameter for item i ,

c_i is the lower asymptote or “pseudo-guessing” parameter for item i , and

$D = 1.7$ or 1.0 .

Item and Test Information

Item information for the dichotomous IRT models for item i is defined as

$$I_i = \frac{P_i'}{P_i Q_i} \quad (2)$$

where P_i' is the first derivative of the IRF with respect to θ and

$$Q_i = 1 - P_i. \quad (3)$$

Item information is then computed by

$$I_i = D^2 a_i^2 \frac{Q_i}{P_i} \left(\frac{P_i - c_i}{1 - c_i} \right)^2, \quad (4)$$

and test information is

$$I \theta = \sum_{i=1}^n I_i \theta . \quad (5)$$

The conditional model-predicted standard error of measurement (SEM) is computed from the equation

$$\text{SEM } \theta = 1 / [I \theta]^{1/2} . \quad (6)$$

Polytomous Model Equations

Response Probabilities

Samejima's Graded Response Model

The following equations are for the homogenous case of the graded response model, which assumes that within each item the discriminations of the options are equal (i.e., there is a single discrimination for each item), but it allows discriminations to vary across items.

The boundary response function (BRF) is defined as

$$P_{i_g}^* \theta_j = \frac{\exp [a_i \theta_j - b_{i_g}]}{1 + \exp [a_i \theta_j - b_{i_g}]} , \quad (7)$$

where a_i is the item discrimination parameter,

b_{i_g} is the boundary location parameter for boundary g , and

$$P_0^* \theta_j = 1 \text{ and } P_{m+1}^* \theta_j = 0 , \quad (8)$$

where $g = m - 1$ and m is the number of response options.

Then the option response function (ORF) is defined as

$$P_{i_g} \theta_j = P_{i_g}^* \theta_j - P_{i_{g+1}}^* \theta_j . \quad (9)$$

Thus, the probability of responding by selecting a given response option is equal to the probability of responding above the category's lower boundary (i_g) minus the probability of responding above the category's upper boundary (i_{g+1})

Generalized Rating Scale Model

This model is a variation of the SGRM in which there is a single set of boundary locations, c_g , that is constant for all items, and a single location parameter, b_i , for each item. The boundary response functions then become

$$P_{i_g}^* \theta_j = \frac{\exp\left[a_i \theta_j - b_i + c_g\right]}{1 + \exp\left[a_i \theta_j - b_i + c_g\right]}, \quad (10)$$

and the option response functions are then computed from Equation 9.

Rasch Rating Scale Model

$$P_{i_g} \theta_j = \frac{\exp\sum_{n=0}^g \left[\theta_j - b_i + c_n\right]}{\sum_{h=0}^m \exp\left\{\sum_{x=0}^h \left[\theta_j - b_i + c_x\right]\right\}}. \quad (11)$$

Generalized Partial Credit Model and the Rasch Partial Credit Model

In the partial credit models, the probability of responding by selecting a particular response option, g , is computed directly from

$$P_{i_g|g-1,g} \theta_j = \frac{\exp\left[\sum_{n=0}^g a_i \theta_j - b_{i_g}\right]}{\sum_{h=0}^m \exp\left[\sum_{x=0}^h a_i \theta_j - b_{i_g}\right]}, \quad (12)$$

where b_{i_g} is the boundary (or “step”) location parameter. For the Rasch partial credit model, $a_i = 1.0$ for all items.

Item Information

Graded Response Model and the Generalized Partial Credit Model (Difference Models)

In this class of models, option information is defined as

$$I_{i_g}(\theta) = \frac{P'_{i_g}(\theta)^2}{P_{i_g}(\theta)} = \frac{\left[P'_{i_g}(\theta) - P'_{i_{g+1}}(\theta)\right]^2}{\left[P_{i_g}(\theta) - P_{i_{g+1}}(\theta)\right]}, \quad (13)$$

where P' is the first derivative of the given function

Total item information, then, is the sum of option information:

$$I_i \theta_j = \sum_{g=1}^m \frac{P'_{i_g} \theta_j^2}{P_{i_g} \theta_j}. \quad (14)$$

Rasch Rating Scale Model, Rasch Partial Credit Model, and Generalized Partial Credit Models (Divide-By-Total or Adjacent Category Models)

For these models, item information is calculated by

$$I_i(\theta) = D^2 a_i^2 \left\{ \left[\sum_{n=0}^{m_i} T_i^2 P_{i_g} \theta \right] - \left[\sum_{n=0}^{m_i} T_i P_{i_g} \theta \right]^2 \right\}, \quad (15)$$

where T is the scoring function, typically consisting of equally spaced positive integers corresponding to the response options (for example 1, 2, 3, 4).

Equations for Estimating θ

Maximum Likelihood Estimation

Maximum likelihood estimation is implemented by finding the maximum of the likelihood function, defined for dichotomously scored items by

$$L(\mathbf{u} | \theta_j) = \prod_{i=1}^n P_{ij}^{u_{ij}} Q_{ij}^{1-u_{ij}}. \quad (16)$$

For polytomously scored items, the likelihood function is

$$L(\mathbf{u}_j | \theta, \xi) = \prod_{i=1}^n P_{i_g}(\theta). \quad (17)$$

Bayesian Estimation

Bayesian modal estimation is implemented by estimating the Bayesian posterior distribution, defined by

$$f(\theta | \mathbf{u}) = L(\mathbf{u} | \theta_j) f(\theta_j), \quad (18)$$

where

$f(\theta | \mathbf{u})$ is the posterior distribution function,

$L(\mathbf{u} | \theta_j)$ is the likelihood defined by Equation 16 or 17, and

$f(\theta_j)$ is the prior distribution, which usually is assumed to be normal with a user-specified mean and standard deviation.

As Equation 18 indicates, the Bayesian posterior distribution is the product of the likelihood function (computed across all items administered at any point in the test) and the Bayesian prior distribution. Bayesian modal estimation—or maximum a posterior (MAP)—estimates θ by

evaluating the mode (or maximum) of the posterior distribution. Newton-Raphson iterations are used to find the maximum of the function. EAP (expected a posteriori) estimation estimates θ by determining the mean of the posterior distribution. These two estimates will be the same if the posterior distribution is symmetrical (and has a maximum) and will differ when the posterior distribution is skewed.

The standard error of each $\hat{\theta}$ (observed SEM) is determined from the variance of the likelihood function for ML estimation (Baker, 2004, pp. 64 – 67) and from the Bayesian posterior variance for Bayesian estimation (i.e., the likelihood function multiplied by the Bayesian prior distribution). In both cases, this SEM is determined from the second derivative of the log-likelihood function,

$$SEM \mid \hat{\theta}_j = \sqrt{Var \hat{\theta}_j \mid \theta} , \quad (19)$$

where

$$Var \hat{\theta}_j \mid \theta = \frac{1}{I \hat{\theta}_j} \quad (20)$$

and

$$I \hat{\theta}_j = \left(\frac{\partial^2 \ln L}{\partial \theta_j^2} \right) . \quad (21)$$

Weighted Maximum Likelihood

The first order bias of MLE for dichotomously scored items was derived by Lord (1983) as

$$BIAS_1(\hat{\theta}_{MLE}) = \frac{1}{I(\theta)^2} \sum_{i=1}^n A_i I_i(\theta) \varphi_i - .5 , \quad (22)$$

where

$$\varphi_i = \frac{P_i - c_i}{1 - c_i} . \quad (23)$$

Warm (1989) proposed a weighted maximum likelihood (WML) estimator that corrects for the bias of MLE. The weighted first derivative (*WFD*) of the log of the likelihood (*LL*) function is

$$WFD = \frac{\partial(LL)}{\partial(\theta)} - BIAS_1(\hat{\theta}_{MLE})I(\theta) . \quad (24)$$

The derivative of the *WFD*, which serves as the second derivative for the Newton-Raphson procedure, is

$$\frac{\partial(WFD)}{\partial(\theta^2)} = \frac{\partial^2(LL)}{\partial(\theta^2)} - \left\{ \frac{-I'}{I^2} \left[\sum_{i=1}^n A_i I_i (\varphi_i - .5) \right] + \frac{1}{I} \sum_{i=1}^n A_i I'_i (\varphi_i - .5) + A_i I_i \frac{P'_i}{(1-c_i)} \right\}. \quad (25)$$

For polytomous items, Samejima (1993) derived the formula for the MLE bias function when the responses are discrete. She showed that the first order bias is

$$BIAS_1 = \frac{-1}{2 I(\theta)^2} \sum_{i=1}^n \sum_{h=1}^m \frac{\frac{\partial(P_{i_h})}{\partial(\theta)} \left[\frac{\partial^2(P_{i_h})}{\partial(\theta^2)} \right]}{P_{i_h}}. \quad (26)$$

The summand in Equation 26 is performed for all categories across all items. The WFD can be obtained by substituting Equation 26 into Equation 24. It was shown by Samejima (1993) that Equations 26 and 24 are equivalent when the responses are dichotomous. The Newton-Raphson procedure is used by CATSim to obtain the WML θ estimate.

A numerical derivative of the WFD is used in CATSim to obtain the second derivative for the Newton-Raphson procedure. The numerical derivative is

$$\frac{\partial(WFD)}{\partial(\theta)} = \frac{WFD(\theta + \delta) - WFD(\theta)}{\delta}, \quad (26)$$

where $\delta = 1E-9$. Delta was chosen to minimize the difference between the exact SE computed for the dichotomous WML and the approximated SE computed for the SGRM/GPCM WML when the 2PL model was used. A δ of $1E-9$ was shown by Guyer (2009) to result in θ estimates precise to at least 17 decimal places and SE values precise to at least 7 decimal places.

Appendix B. Comparison of CATSim and Parscale θ Estimates

To confirm the calculations for the θ estimates in CATSim, item parameters and θ estimates for a set of polytomous items were run in Parscale for each polytomous model. The dataset consisted of item responses from 200 examinees on 20 five-alternative Likert scale type of items. Item parameters estimated by Parscale are those shown in Figures 1.4 –1. 8. θ estimates were compared for the response vectors for the 10 examinees shown in Figure B-1.

**Figure B-1. Item Responses for
10 Examinees Used to Compare
CATSim and Parscale
 θ Estimates and Their SEs**

3 2 3 2 2 4 4 2 2 2 4 3 3 3 3 3 3 2 3 3
3 1 1 1 1 1 3 2 1 1 2 2 1 1 2 3 2 2 5 3
3 2 2 3 2 3 2 3 2 2 4 4 3 3 2 4 4 2 3 2
4 4 5 2 2 5 1 2 1 1 2 1 5 4 3 3 5 5 5 5
2 2 3 3 4 5 3 1 4 3 4 4 3 4 2 3 3 2 4 3
4 3 2 3 3 3 4 3 3 3 3 4 3 3 4 3 4 4 4 4
4 4 4 4 2 4 4 2 2 3 3 4 4 4 4 4 5 4 4 4
4 2 3 3 3 3 2 3 2 4 3 4 3 2 3 3 3 3 4
4 3 4 2 2 3 2 2 3 2 3 4 3 2 3 4 5 4 4 3
3 3 3 3 2 1 3 2 1 2 1 2 3 2 4 3 4 3 3 3

Index to Appendix B Tables

Samejima’s Graded Response Model (SGRM): Tables B-1 and B-2

Generalized Rating Scale Model (GRSM): Tables B-3 and B-4

Rasch Rating Scale Model (RRSM): Tables B-5 and B-6

Rasch Partial Credit Model (RPCM): Tables B-7 and B-8

Generalized Partial Credit Model (GPCM): Tables B-9 and B-10

Comments on the Results in Tables B-7 through B-10

**Table B-1. EAP θ Estimates and SEs From Parscale
and EAP and MAP θ Estimates and SEs From CATSim, for the SGRM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	EAP θ	EAP θ	MAP θ	EAP SE	EAP SE	MAP SE
1	-0.6521	-0.6522	-0.6758	0.2471	0.2470	0.2551
2	-1.9723	-1.9724	-2.0288	0.2711	0.2709	0.2743
3	-0.7899	-0.7900	-0.8235	0.2583	0.2581	0.2658
5	0.8382	0.8382	0.8736	0.2958	0.2956	0.3001
8	-0.5993	-0.5993	-0.6298	0.2702	0.2700	0.2787
9	0.1594	0.1595	0.1561	0.2537	0.2535	0.2594
12	0.7889	0.789	0.8107	0.2383	0.2381	0.2452
13	-0.2266	-0.2266	-0.2361	0.2519	0.2517	0.2570
16	0.0429	0.0428	0.0431	0.2507	0.2504	0.2552
17	-0.5616	-0.5616	-0.5724	0.2496	0.2494	0.2572

Note. Parscale does not provide MAP estimates.

**Table B-2. MLE θ Estimates and SEs From Parscale
and MLE and WML θ Estimates and SEs From CATSim for the SGRM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	MLE θ	MLE θ	WML θ	MLE SE	MLE SE	WML SE
1	-0.7229	-0.7229	-0.7219	0.2640	0.2640	0.2670
2	-2.1948	-2.1948	-2.1892	0.2875	0.2875	0.2827
3	-0.8858	-0.8858	-0.8881	0.2742	0.2742	0.2762
4	0.9591	0.9591	0.9625	0.3112	0.3112	0.3084
5	-0.6827	-0.6827	-0.6803	0.2892	0.2892	0.2934
6	0.1674	0.1674	0.1632	0.2687	0.2687	0.2666
7	0.8624	0.8624	0.8652	0.2524	0.2524	0.2524
8	-0.2528	-0.2528	-0.2475	0.2662	0.2662	0.2640
9	0.0461	0.0461	0.0444	0.2639	0.2639	0.2608
10	-0.6131	-0.6131	-0.6095	0.2672	0.2672	0.2700

Note. Parscale does not provide WML estimates.

**Table B-3. EAP θ Estimates and SEs From Parscale and
EAP and MAP θ Estimates and SEs From CATSim, for the GRSM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	EAP θ	EAP θ	MAP θ	EAP SE	EAP SE	MAP SE
1	-0.6680	-0.6680	-0.6681	0.2779	0.2779	0.2771
2	-1.9590	-1.9590	-1.9564	0.3187	0.3187	0.3158
3	-0.6827	-0.6826	-0.6825	0.2948	0.2948	0.2938
4	0.4612	0.4611	0.4634	0.3503	0.3502	0.3496
5	-0.4277	-0.4277	-0.4286	0.3111	0.3111	0.3100
6	0.3292	0.3292	0.3294	0.2783	0.2782	0.2765
7	0.8539	0.8537	0.8552	0.2697	0.2697	0.2684
8	-0.2240	-0.2240	-0.2223	0.2814	0.2814	0.2793
9	0.0113	0.0113	0.0118	0.2829	0.2829	0.2808
10	-0.6439	-0.6439	-0.6409	0.2797	0.2796	0.2786

Note. Parscale does not provide MAP estimates.

**Table B-4. MLE θ Estimates and SEs From Parscale
and MLE and WML θ Estimates and SEs From CATSim, for the GRSM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	MLE θ	MLE θ	WML θ	MLE SE	MLE SE	WML SE
1	-0.7237	-0.7237	-0.7213	0.2886	0.2885	0.2896
2	-2.1754	-2.1754	-2.1681	0.3368	0.3368	0.3355
3	-0.7470	-0.7470	-0.7446	0.3076	0.3076	0.3090
4	0.5278	0.5278	0.5262	0.3727	0.3727	0.3739
5	-0.4742	-0.4741	-0.4700	0.3261	0.326	0.3259
6	0.3567	0.3566	0.3553	0.2877	0.2877	0.2877
7	0.9216	0.9214	0.9220	0.2783	0.2783	0.2783
8	-0.2412	-0.2411	-0.2388	0.2911	0.291	0.2900
9	0.0128	0.0128	0.0130	0.2926	0.2926	0.2913
10	-0.6950	-0.6950	-0.6924	0.2909	0.2908	0.2918

Note. Parscale does not provide WML estimates.

**Table B-5. EAP θ Estimates and SEs From Parscale
and EAP and MAP θ Estimates and SEs From CATSim, for the RRSM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	θ Estimate	EAP θ	MAP θ	EAP SE	EAP SE	MAP SE
1	-0.5267	-0.4109	-0.3959	0.3098	0.2468	0.2503
2	-2.0029	-1.5610	-1.5814	0.3361	0.2640	0.2702
3	-0.6045	-0.4719	-0.4586	0.3101	0.2471	0.2505
4	0.175	0.1801	0.1697	0.312	0.2476	0.2520
5	-0.0602	-0.0899	-0.0198	0.3104	0.2464	0.2508
6	0.3334	0.3032	0.2973	0.3136	0.2486	0.2532
7	0.8221	0.6819	0.6904	0.3216	0.2545	0.2594
8	-0.2159	-0.1244	-0.1454	0.3099	0.2464	0.2504
9	-0.0602	-0.0029	-0.0198	0.3104	0.2466	0.2508
10	-0.9176	-0.7176	-0.7108	0.3120	0.2488	0.2519

Note. See the discussion concerning these results following Table B-10. Parscale does not provide MAP θ estimates.

**Table B-6. MLE θ Estimates and SEs From Parscale and MLE and WML
 θ estimates and SEs From CATSim for the RRSM With D = 1.0**

Person	Parscale	CATSim		Parscale	CATSim	
	MLE θ	MLE θ	WML θ	MLE SE	MLE SE	WML SE
1	-0.5196	-0.4223	-0.4216	0.2952	0.2586	0.2580
2	-1.9875	-1.7085	-1.6953	0.327	0.2865	0.2839
3	-0.5960	-0.4893	-0.4882	0.2955	0.2589	0.2582
4	0.1692	0.1812	0.1789	0.2973	0.2605	0.2598
5	-0.0617	-0.0211	-0.0224	0.2957	0.2591	0.2584
6	0.3250	0.3177	0.3145	0.2990	0.2620	0.2612
7	0.8075	0.7404	0.7341	0.3079	0.2698	0.2686
8	-0.2145	-0.1551	-0.1557	0.2951	0.2586	0.2579
9	-0.0617	-0.0211	-0.0224	0.2957	0.2591	0.2584
10	-0.9039	-0.7590	-0.7565	0.2976	0.2608	0.2600

Note. See the discussion concerning these results following Table B-10. Parscale does not provide WML θ estimates.

Table B-7. EAP θ Estimates and SEs From Parscale, FIRESTAR, and CATSim for the RPCM With D = 1.0

Person	θ Estimate			Standard Error		
	Parscale	FIRESTAR	CATSim	Parscale	FIRESTAR	CATSim
1	-0.5043	-0.41607	-0.4161	0.3088	0.255563	0.2556
2	-2.0620	-1.70498	-1.7050	0.3446	0.285143	0.2851
3	-0.5835	-0.4816	-0.4816	0.3099	0.256432	0.2564
4	0.1922	0.160251	0.1603	0.3044	0.251917	0.2519
5	-0.0380	-0.03021	-0.0302	0.3048	0.252188	0.2522
6	0.3458	0.287358	0.2874	0.3050	0.252371	0.2524
7	0.8150	0.675647	0.6756	0.3111	0.257393	0.2574
8	-0.1921	-0.15774	-0.1577	0.3057	0.252931	0.2529
9	-0.0380	-0.03021	-0.0302	0.3048	0.252188	0.2522
10	-0.9064	-0.74879	-0.7488	0.3150	0.260664	0.2607

Note. See the discussion concerning these results following Table B-10.

Table B-8. MLE θ Estimates and SEs From Parscale, FIRESTAR, and CATSim for the RPCM

Person	θ Estimate			Standard Error		
	Parscale	FIRESTAR	CATSim	Parscale	FIRESTAR	CATSim
1	-0.4925	-0.44150	-0.4415	0.2941	0.264136	0.2641
2	-2.0572	-1.84656	-1.8466	0.3367	0.302378	0.3024
3	-0.5706	-0.51156	-0.5116	0.2954	0.265248	0.2652
4	0.1892	0.17073	0.1707	0.2886	0.259196	0.2592
5	-0.0355	-0.03101	-0.0310	0.2892	0.259668	0.2597
6	0.3390	0.30529	0.3053	0.2892	0.259691	0.2597
7	0.7982	0.71768	0.7177	0.2961	0.265925	0.2659
8	-0.1862	-0.16635	-0.1664	0.2903	0.260690	0.2607
9	-0.0355	-0.03101	-0.0310	0.2892	0.259668	0.2597
10	-0.8900	-0.79843	-0.7984	0.3013	0.270568	0.2706

Note. See the discussion concerning these results following Table B-10.

Table B-9. EAP θ Estimates and SEs From Parscale, FIRESTAR, and CATSim for the GPCM With D = 1.0

Person	θ Estimate			Standard Error		
	Parscale	FIRESTAR	CATSim	Parscale	FIRESTAR	CATSim
1	-0.6726	-0.64276	-0.6428	0.2753	0.263454	0.2635
2	-2.0014	-1.91453	-1.9145	0.283	0.270854	0.2709
3	-0.8211	-0.78500	-0.7850	0.2754	0.263557	0.2636
4	0.7540	0.72258	0.7226	0.2807	0.268639	0.2686
5	-0.5615	-0.53669	-0.5367	0.2749	0.263123	0.2631
6	0.0808	0.07819	0.0782	0.2727	0.261030	0.2610
7	0.8517	0.81603	0.8160	0.2831	0.270991	0.2710
8	-0.2910	-0.27760	-0.2776	0.2736	0.261833	0.2618
9	0.0785	0.07602	0.0760	0.2727	0.261026	0.2610
10	-0.5983	-0.57174	-0.5717	0.2751	0.263253	0.2633

Note. See the discussion concerning these results following Table B-10.

Table B-10. MLE θ Estimates and SEs From Parscale, FIRESTAR, and CATSim for the GPCM, With D = 1.0

Person	θ Estimate			Standard Error		
	Parscale	FIRESTAR	CATSim	Parscale	FIRESTAR	CATSim
1	-0.6556	-0.69027	-0.6903	0.2567	0.27429	0.2743
2	-1.9344	-2.05647	-2.0565	0.2667	0.28488	0.2843
3	-0.7998	-0.84436	-0.8444	0.2566	0.27415	0.2741
4	0.7132	0.77211	0.7721	0.2608	0.27861	0.2786
5	-0.54780	-0.57534	-0.5753	0.2562	0.27376	0.2738
6	0.0687	0.08345	0.0834	0.252	0.26923	0.2692
7	0.8075	0.87274	0.8273	0.2634	0.28138	0.2814
8	-0.2866	-0.29609	-0.2961	0.2538	0.27114	0.2711
9	0.0665	0.08113	0.0811	0.252	0.26922	0.2692
10	-0.5835	-0.61330	-0.6133	0.2565	0.27399	0.2740

Note. See the discussion concerning these results following Table B-10.

Comments on the Results in Tables B-5 through B-10

The results in tables B-1 through B-7 show that the θ estimates and their standard errors computed by Parscale and CATSim, using both MLE and EAP, agreed (in most cases to three decimal places) for the SGRM and the GRSM. For the other three models—the RRSM, RPCM, and GPCM—however, the results for the two programs did not agree.

Because of this disagreement, two type of additional information were used to determine which program was giving correct results:

1. For the RRSM, RPCM, and GPCM, MLE θ estimates were approximated using a discrete arithmetic estimation procedure. This procedure, accurate to .01, estimated θ by multiplying the option response functions for the observed response pattern at 601 values of θ from $\theta = -3.00$ to $+3.00$. The result of this multiplication is a discrete approximation to the likelihood function for the response pattern. The maximum of the function was determined by finding the highest likelihood among the 601 values. The θ estimate for the response pattern is the θ value associated with that maximum. The estimated standard error of the θ estimate (SEM) was computed by summing the values of item information, given the θ estimate, for all items administered to obtain $I_{\hat{\theta}_j}$. This value was then used in Equations 20 and 21 to obtain the estimated SEM. These estimates were compared with those from CATSim, and they agreed with the tabled results to two decimal places (the limit of accuracy of the discrete arithmetic approach).
2. For the RPCM and GPCM, the θ estimates obtained from CATSim were compared with those obtained from FIRESTAR (Choi, 2007), a public domain CAT simulation program that operates using the R computing language (FIRESTAR does not implement the RRSM). Results from FIRESTAR are shown in Tables B-7 through B-10. Note that CATSim and FIRESTAR θ estimates and their SEMs—both MAP and EAP—agreed with each other, but neither agreed with Parscale.
3. For the RRSM, the differences between CATSim and Parscale are similar in direction and magnitude to those for the RPCM and GPCM. In addition, to determine whether the differences in these θ estimates were due to differences in the option response functions, item and test information functions (which are based on the option response functions) were compared between CATSim and Parscale. The results of this comparison indicated that both programs obtained identical information functions; hence, it can be concluded that they were using identical option response functions. Therefore, differences in the θ estimation results had to derive from differences in the estimation procedures. Since the estimation methods in CATSim for the RRSM used the same computational procedures as did the RPCM and GPCM, it is assumed that the RRSM results for CATSim are correct.
4. Generally, when MLE θ estimates are compared with Bayesian θ estimates, the Bayesian estimates are somewhat regressed toward the prior mean, and the standard errors of the Bayesian estimates are smaller than those of the MLEs. For example, Table B-1 shows EAP θ estimates and SEMs for the SGRM (for which CATSim and Parscale agreed). All EAP θ estimates are regressed toward the prior mean of 0.0, compared to their MLE counterparts

(Table B-2), and all EAP SEMs were smaller than the MLE SEMs. The same pattern was observed for the RPCM (Tables B-7 and B-8) and the GPCM (Tables B-9 and B-10): For CATSim and FIRESTAR, EAP θ estimates were more regressed than MLE θ estimates and their SEMs were smaller. For Parscale, however, an opposite pattern was observed: MLE θ estimates were slightly more regressed than the EAP estimates and their SEMs were generally smaller.

Appendix C: License, Unlocking, and Transferring

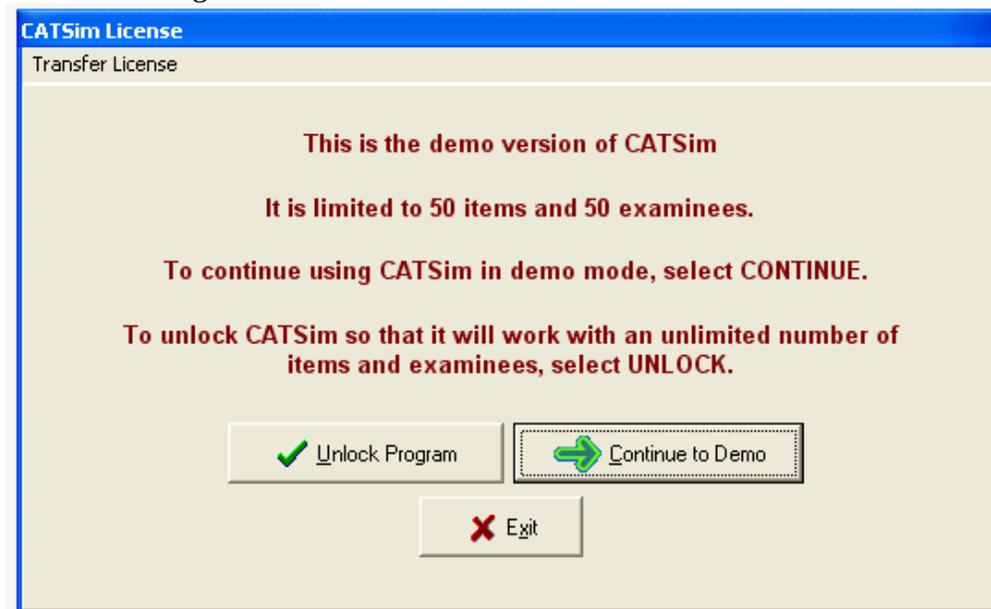
Your CATSim License and Unlocking Your Copy

Unless you have purchased a network or multiple-computer license, your license for *CATSim* is a single-user license. Under this license you may install *CATSim* on two computers (e.g., a desktop and a laptop) so long as there is no possibility that the two copies of the software will be in use simultaneously. If you would like to use *CATSim* on a network or by more than one user, please contact us to arrange for the appropriate number of additional licenses.

CATSim is shipped as a functionally-limited demonstration copy. It is limited to no more than 50 items and 50 examinees, but has no expiration date. We can permanently convert your demo copy to the fully functioning software by email, phone, or fax once you have completed the license purchase. To unlock *CATSim*, please email, phone, or fax to ASC:

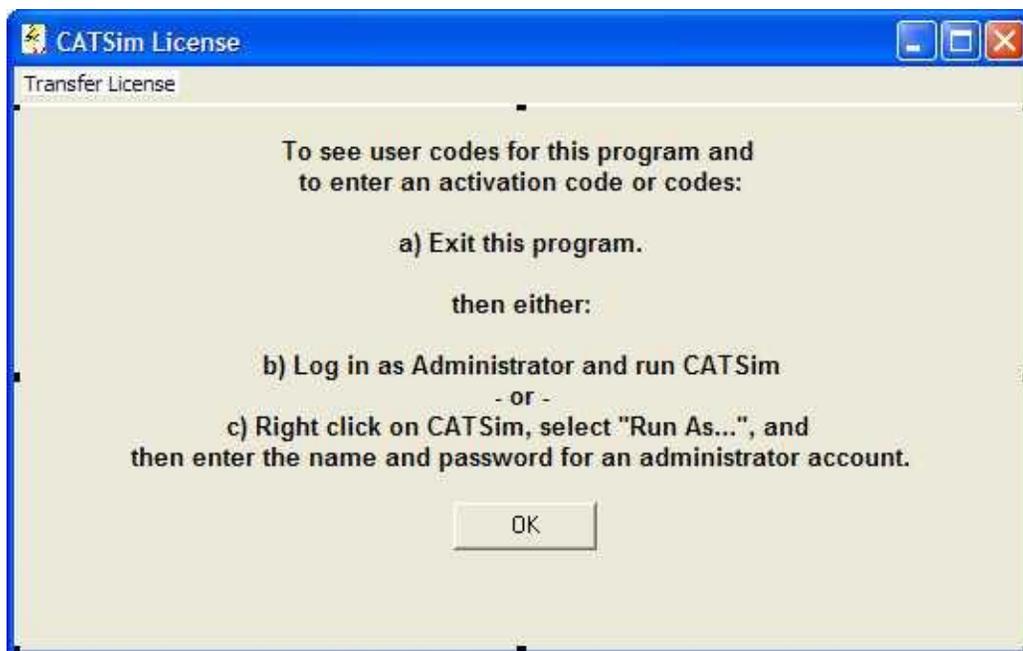
1. Your name and email address.
2. Your organization or affiliation.
3. Your invoice number (in the top right corner of your invoice). You should make a record of your invoice number since you might be asked for it if you request technical support.
4. The “unlock codes,” which are two numeric codes that are unique to the installation of *CATSim* on any given computer. To obtain these two codes, click on the “Unlock Program” button when *CATSim* starts (Figure C.1) This license window can also be reached by clicking on the License button and selecting “Unlock” when *CATSim* is running in demo mode.

Figure C.1. Screen Visible When CATSim is Locked

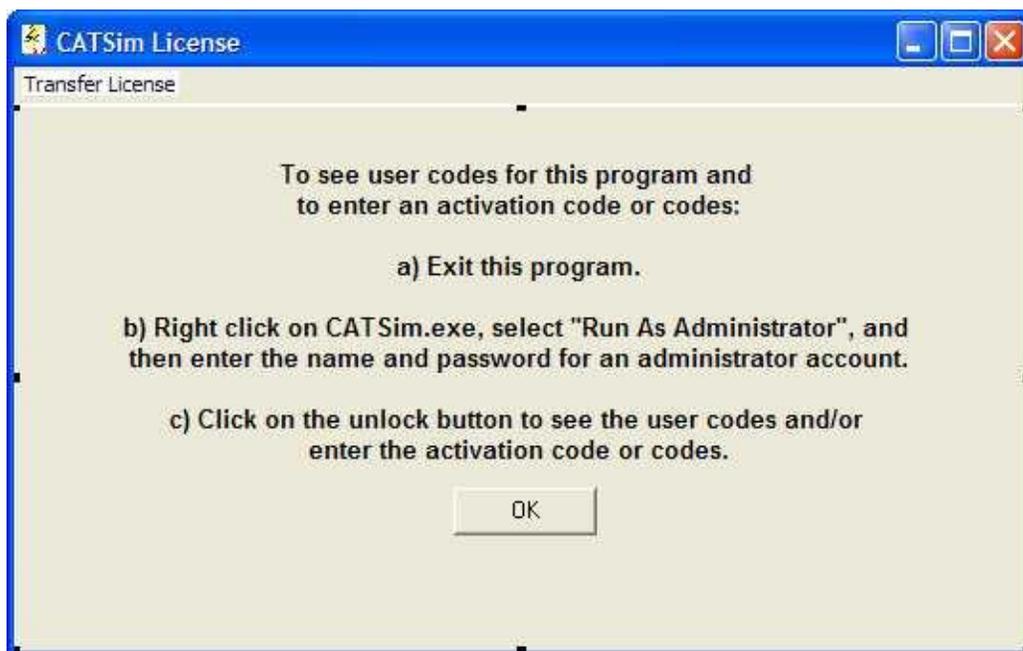


If the program has not been run in administrator mode, you may see one of the following windows depending on if you are an XP user with non-administrator rights, a Vista or Windows 7 user with non-administrator rights, or a Vista or Windows 7 administrator:

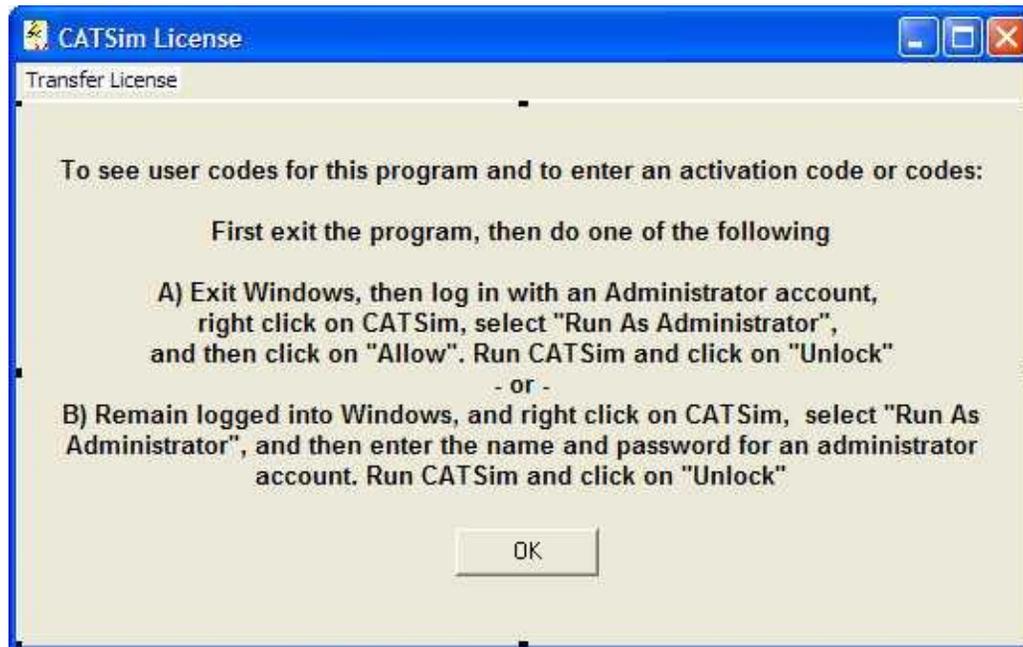
XP user with non-administration rights:



Vista or Windows 7 user with non-administration rights:



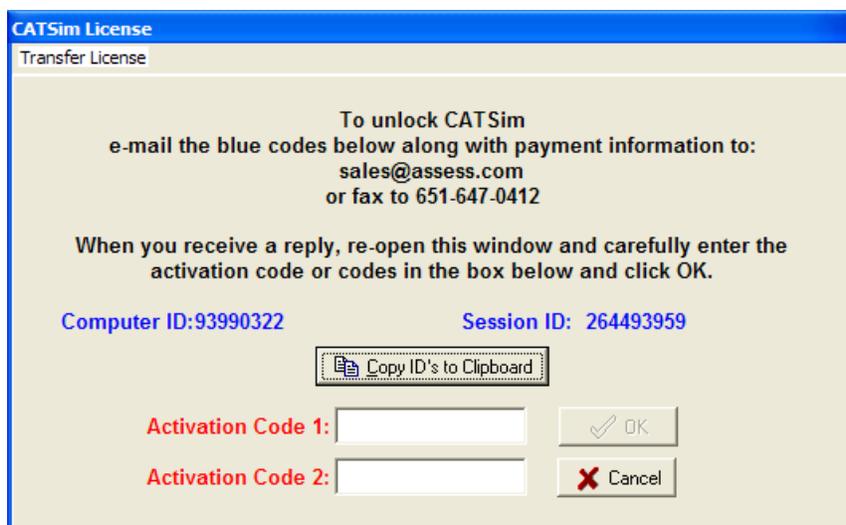
Vista or Windows 7 user with administrator rights:



From the unlock screen you will need to send us the two blue Computer ID and Session ID numbers (Figure C.2). For your convenience, we have provided a “Copy IDs to Clipboard” button. This will copy both IDs to the Windows clipboard along with a brief message and the email address to which to send your payment information. This can then be pasted into an email message, filled in, and sent to sales@assess.com. If you have already paid for your CATSim license, be sure to add your invoice number to this message.

When we receive these codes from you, we will respond with a single numeric Activation Code (if you have purchased a permanent license) or two codes (if you have purchased an annual subscription license) that you will need to enter into this same window from which you obtained your Activation Codes (the red labels in Figure C.2). Once you enter the code(s) that we send you, your copy will be unlocked and fully functional.

Figure C.2. The Unlock Screen



Note that if you install *CATSim* on a second computer, you will need to repeat this process for that computer since the unlock codes are specific to a given computer.

CATSim is permanently unlocked for academic use, but is an annual subscription for non-academic use. The license status box in the lower right-hand corner of the *CATSim* window will display the current license status, including the number of days remaining for your subscription. As the subscription nears the end, the background color of the box will change to alert you to the need to renew your subscription for another year (red if you have less than 30 days remaining, yellow if 30–90 days, and green if more than 90 days).

License Transfer

License transferring is a 3-step process that takes a license from a licensed program on one computer, and gives it to a program already installed in demo mode on another computer. The original demo program (new computer) becomes a licensed program, and the original licensed program (old computer) reverts to a demo. This process can transfer a license between PCs running the same program on different versions of Windows such as XP and Vista.

This process starts with two computers, one that has an unlicensed program (original demo computer), and one that has an already licensed program (original licensed computer). It starts on the original demo computer, where the program creates a transfer file. This transfer file is taken to the original licensed computer, where the program there puts its license in the transfer file. The transfer file, now containing the license, is carried back to the original demo computer. The program on the demo computer takes the license out of the transfer file, becoming licensed. The program on the original license computer becomes a demo after it puts its license in the transfer file.

This process requires the use of a separate drive, such as:

- An external removable drive such as a USB flash/thumb drive.
- Blank formatted floppy disk.

- Other connected or networked drives.

This transfer drive will carry the transfer file from the (new) original demo computer to the (old) original licensed computer to get the license from the licensed program and back to the (new) original demo computer to give the license to the demo program.

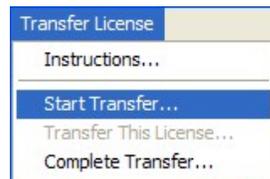
Step 1 – Demo/Trial Program

Start with the unlicensed demo program on the original demo computer. Run the program in Administrative mode, logging in as administrator if necessary. Click on the License button (Figure C.3; marked as ‘Demo’) to bring up a dialog with transfer license menu in upper left corner (Figure C.4).

Figure C.3. License Button



Figure C.4. Transfer License Menu and “Start Transfer” Option



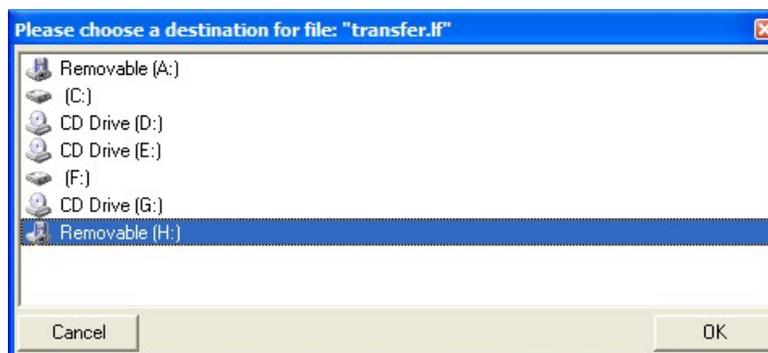
Select “Start Transfer” and follow the prompts. Be sure to connect the appropriate drive for use as the transfer drive when prompted, if it isn’t already connected (Figure C.5). Remember the drive letter assignment for this drive.

Figure C.5. Final Prompt to Connect Drive or Insert Disk



Once OK is clicked, the drive dialog is displayed (Figure C.6). “Removable (A:)” will always be the floppy drive. Internal hard drives are marked by their drive letter only. USB flash/thumb drives and other externally connected drives will be marked as “Removable”.

Figure C.6. Choose a Drive



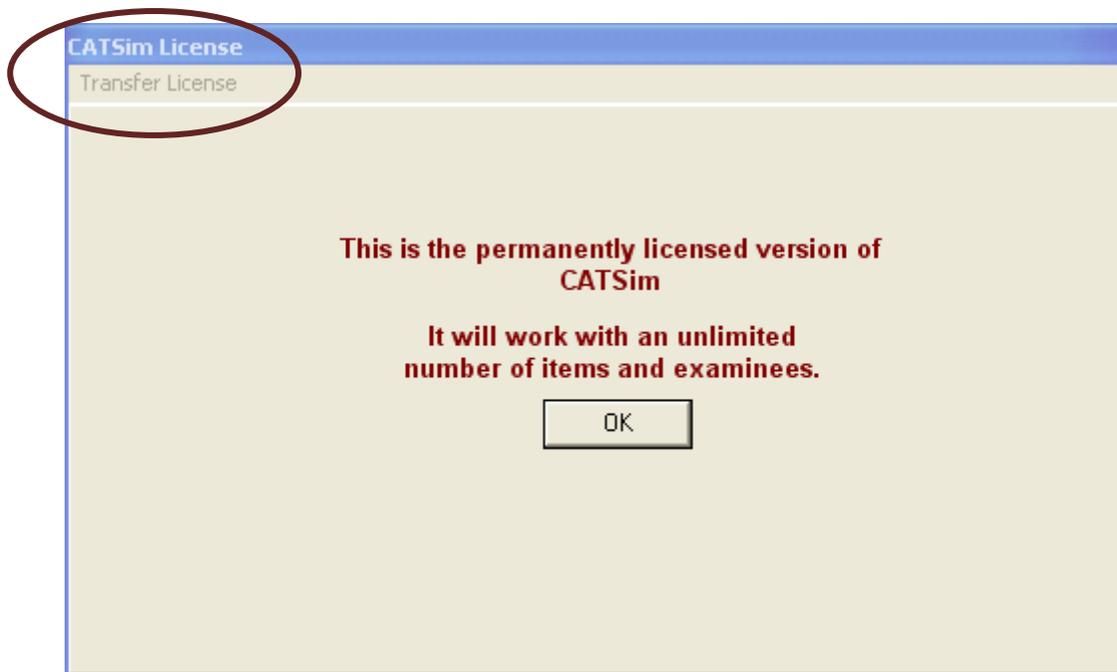
Select the drive to carry the transfer file. Once the process is complete, if a USB flash/thumb drive or external hard drive is used, carefully disconnect it. If there is a problem during this step, an error message will be shown. Please note any error codes and report the error to Assessment Systems at support@assess.com.

Step 2 – Licensed Program

If a USB flash/thumb drive or external hard drive is carrying the transfer file, connect it to the original licensed computer. If a networked hard drive is carrying the transfer file, make sure it can be reached on the original licensed computer. Regardless of which type of drive is used for the transfer, it might have a different drive letter assignment on the original licensed computer than on the original demo computer.

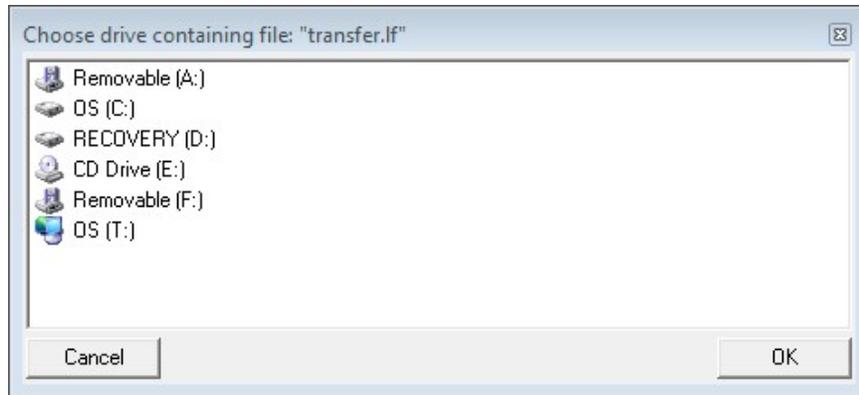
Run the program on the original licensed computer in Administrative mode, logging in as administrator if necessary. Click on the License button to bring up the license window, and click on the transfer license menu in the upper left again (Figure C.7). Select the “Transfer This License” option.

Figure C.7. Transfer This License



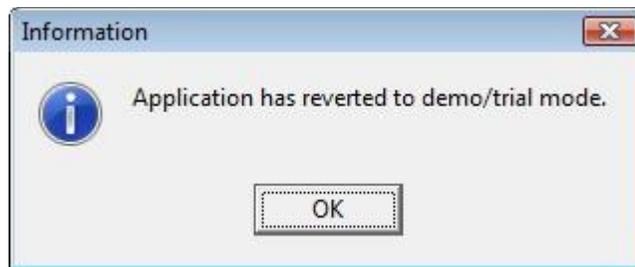
The program will ask for confirmation, then prompt once again to connect the drive or diskette carrying the transfer file (Figure C.8). If this has not been done already, please do so, and remember which drive letter Windows assigns to it.

Figure C.8. Drive Dialog



Follow the prompts to the drive dialog (Figure C.8), and select the appropriate drive, which might have a different drive letter on the original licensed computer than on the original demo computer. The program will transfer the license to the transfer file and will indicate that it is now in demo/trial mode (Figure C.9).

Figure C.9. Notification of Change in Mode

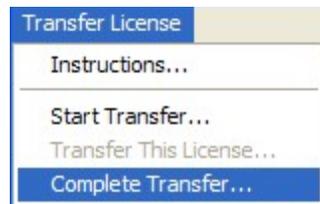


Carefully disconnect the drive once this step is complete. If there have been any errors, please note them along with any specific codes and report them to Assessment Systems at support@assess.com.

Step 3 – Demo/Trial Program

Connect the transfer drive to the original demo computer. Run the demo/trial program, in Administrative mode, logging in as administrator if necessary. and click on the License button to bring up the license window, then click on the transfer license menu in the upper left again. Select the “Complete Transfer” option (Figure C.10).

Figure C.10. Complete Transfer



Follow the prompts to connect the transfer drive if this hasn't already been done, and to select the drive. If the license transfer was successful, a message will appear.

Figure C.11. Successful Transfer



If there have been any errors, please note them along with any specific codes and report them to Assessment Systems at support@assess.com.