

Client Resource

A Guide to Understanding Categorical Growth Analysis

by
Sarah M. Callahan, Ph.D.



**Assessment
Technology
Incorporated**

Assessment Technology, Incorporated

6700 E. Speedway Boulevard
Tucson, Arizona 85710

Phone: 520.323.9033 • Fax: 520.323.9139

Copyright © Assessment Technology, Incorporated 2013. All rights reserved.
"Galileo" and the Galileo logos are trademarks or registered trademarks of Assessment Technology, Incorporated.

Copyright © 2013 by Assessment Technology, Incorporated All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the publisher.

“Galileo” and the Galileo logos are trademarks or registered trademarks of Assessment Technology, Incorporated.

Assessment Technology, Incorporated, Publishers Tucson, Arizona, U.S.A.
Printed in the United States of America.
V1-120613

A Guide to Understanding Categorical Growth Analysis

Sarah M. Callahan, Ph.D.
Assessment Technology, Incorporated

Table of Contents

Table of Contents.....	i
I. Overview and Goals of Categorical Growth Analysis.....	1
II. Explanation of Underlying Methods	1
A. Calculating Observed Growth	1
B. Determining Expected Growth	1
C. Adjusting for Ceiling Effects	2
D. Evaluating Effectiveness Using the Repeated-Measures t-Test.....	2
E. Assigning a Categorization	3
F. Sample Illustrations including Calculations	3



I. Overview and Goals of Categorical Growth Analysis

Categorical Growth Analysis (CGA) is a statistical approach developed by Assessment Technology Incorporated (ATI) to assist districts and charters in assessing the growth of the students in a class or school. CGA employs a common, well-established statistical method, the repeated-measures t-test, to evaluate the observed growth for the students in a given class or school relative to an expectation for growth. Based on this evaluation, classes and schools are categorized as “Expected Growth Exceeded,” “Expected Growth Maintained,” or “Expected Growth Not Maintained.” The information provided by CGA is displayed on the Galileo® K-12 Online *Dashboard* as part of the *Student Growth and Achievement* widget and the *Categorical Growth Summary* widget and, if desired, can be designated to contribute to an educator’s overall evaluation within the *Evaluation Score Compiler*. This resource document is designed to help teachers and administrators understand the methods underlying CGA including the repeated-measures t-test. A thorough understanding provides the foundation for productive use of the information provided by the analysis to inform educator evaluations, guide decisions regarding professional development, and identify instructional practices resulting in exceptional or inadequate growth.

II. Explanation of Underlying Methods

CGA evaluates whether the observed growth for a class or school is significantly different from a growth expectation. First, the relevant group of students is identified, observed growth is calculated for each student, and the expected growth for the time between the two assessments is determined. If needed, observed growth scores for students who experienced ceiling effects are adjusted so that the results of the CGA analysis are not negatively impacted. Next, a repeated measures t-test is used to evaluate whether the average observed growth for the class or school is significantly different from the growth expectation. The results of the t-test are then translated into a categorization indicating whether the students in the class or school, as a group, exceeded, maintained, or failed to maintain expected growth.

A. Calculating Observed Growth

For purposes of CGA, observed growth is defined as the change in a student’s ability between two assessments (e.g., an instructional effectiveness pretest and posttest). Galileo provides an Item-Response-Theory-based estimate of student ability for each district-wide assessment (i.e., the Developmental Level [DL] score). Unlike raw scores, DL scores can be placed on a common scale across assessments in order to support the evaluation of growth. Provided that the DL scores are on a common scale, the student’s DL score for the first assessment can simply be subtracted from their DL score for the second assessment, yielding an estimate of observed growth. For assessments in state-tested content areas containing sufficient numbers of items with established Item Response Theory item parameters, ATI automatically places DL scores on a common scale across assessments within each grade. For assessments without sufficient item parameter coverage (e.g., assessments in non-state-tested content areas), ATI can also place DL scores on a common scale when the two tests contain a sufficient number of overlapping items.

B. Determining Expected Growth

ATI conducts annual research that employs regression analyses to model student growth patterns throughout the previous year for each grade and content area for which

sufficient data are available. Each regression analysis provides an estimate of the slope of the line that best describes the daily change in student DL scores for a grade and content area (i.e., a growth constant). ATI research has demonstrated that growth patterns differ based on whether the assessments are comprehensive or curriculum-aligned. Therefore, ATI provides a growth constant for use between two comprehensive assessments as well as a growth constant for use between two curriculum-aligned assessments. To accommodate testing strategies that employ a mix of the two assessment types, ATI also provides a “hybrid” growth constant that can be used between two assessments of different types. Finally, in order to accommodate the heightened interest in growth expectations from a pretest administered at the beginning of the school year to a posttest administered at the end of the school year, ATI also provides a growth constant that is based specifically on student performance from a pretest administered at the beginning of the year (i.e., July 1 through October 31) to a posttest administered at the end of the year (i.e., March 1 through June 30).

For purposes of CGA, growth expectations in state-tested content areas (i.e., math, English language arts, science, writing) are determined by multiplying the appropriate model-based growth constant by the number of days between the two assessments. This yields an expected growth value for the time period between the two assessments. Since sufficient data has not yet been collected to conduct regression analyses for non-state-tested content areas (e.g., music, physical education), growth expectations in non-state-tested content areas are determined based on the average observed growth for the students who took the tests within the district/charter school. This approach provides a growth expectation that supports the categorization of a class or school as exceeding, maintaining, or failing to maintain the average growth demonstrated by the broader group of students. This approach can also sometimes be applied in state-tested content areas (e.g., when highly atypical growth patterns are observed).

C. Adjusting for Ceiling Effects

When a student receives the maximum possible score on an assessment, they have “hit the ceiling” of the assessment (i.e., experienced a ceiling effect). Although the student receives the highest possible DL score for the assessment, their true ability may be even higher. If a student experiences a ceiling effect on the second of the two assessments used to estimate growth, it is also possible that the student’s true growth may be underestimated. In addition, the constraints of a ceiling on the second test may make it impossible for the student to achieve a score that meets the growth expectation.

ATI adjusts the observed growth value in circumstances where the ceiling effect may have a negative impact on the CGA. A student experiencing a ceiling effect receives the highest of the following values: their observed growth, expected growth, and the median observed growth for the class. Note that under this approach the student’s observed growth is never adjusted to be a lower value. Currently, the adjustment for a ceiling effect is applied at the class level. The observed growth values for students adjusted, if necessary, for ceiling effects are then included in the corresponding school-level CGA.

D. Evaluating Significance Using the Repeated-Measures t-Test

The repeated measures t-test is a statistical test that is used when the data represent a pair of measurements collected from participants at two time points. For purposes of the t-test, the two measurements are converted to a difference score for each participant. In the repeated measures t-test conducted as part of the CGA for a class or school, the two measurements represent the DL scores for a student on two assessments (e.g., an instructional effectiveness

pretest and posttest). The two DL scores are used to calculate an observed growth score for each student. CGA is applied when the class or school contains a minimum of ten students with scores at both time points.

The t-test evaluates the difference between the average observed growth and the growth expectation. A difference that is very large relative to the variability in student growth scores is unlikely to occur by random chance. In other words, it is a significant difference. The complete set of calculations involved in the repeated-measures t-test is provided as part of the sample illustration in Section F.

E. Assigning a Categorization

Based on the results of the repeated measures t-test, CGA provides a categorization of the growth observed for the students, as a whole, in the class or school. A categorization of “Expected Growth Exceeded” means that the average observed growth was significantly larger than the growth expectation. In contrast, a categorization of “Expected Growth Not Maintained” means the average observed growth was significantly smaller than the growth expectation. A categorization of “Expected Growth Maintained” means that the average observed growth was comparable to the growth expectation (i.e., not significantly different).

F. Sample Illustrations including Calculations

This section is designed to provide a deeper understanding of CGA by walking through the process for a sample class step by step. In this sample illustration, Mrs. Smith teaches third-grade math in Canyon School District (CSD). There are 25 students in her class. CSD administered a third math instructional effectiveness (IE) pretest at the beginning of September and a third math IE posttest at the end of March.

Step 1: Identifying the Group of Students

Mrs. Smith has 25 students in her class, but only 21 of these students have scores for both the pretest and posttest. These 21 students will be included in the CGA.

Step 2: Determining the Growth Expectation

The third math IE pretest and third math IE posttest represent a pretest administered at the beginning of the year and a posttest administered at the end of the year. As described in Section II.C, the daily growth constant for pretest-posttest pairs in third grade math established by ATI research is 0.33. By looking at the dates on which scores came into Galileo[®], the time period between the pretest and posttest is established as 270 days (i.e., nine months). Multiplying the number of days by the daily growth constant yields a growth expectation of 89 points for that time period.

$$\text{Growth Expectation} = (\text{Daily Growth Constant}) * (\# \text{ Days}) = 0.33 * 270 = 89.1$$

Step 3: Calculating Observed Growth and Adjusting for Ceiling Effects

For each student, observed growth is calculated by subtracting the student’s DL score on the IE pretest from their DL score on the IE posttest.

TABLE 1
DL scores and calculated observed growth values for a sample class

Mrs. Smith's 5th Grade Math Class				
Student	DL Score 2 (IE Posttest)	DL Score 1 (IE Pretest)	Observed Growth (DL2-DL1)	Observed Growth (adjusted if necessary)
1	851	828	23	23
2	1024	983	41	41
3	1024	955	69	69
4	903	828	75	75
5	1112	1032	80	80
6	978	886	92	92
7	978	886	92	92
8	1060	966	94	94
9	1203*	1101	102	112
10	940	835	105	105
11	1121	1009	112**	112
12	952	837	115	115
13	978	860	118	118
14	1121	1001	120	120
15	1203*	1078	125	125
16	1142	1009	133	133
17	972	837	135	135
18	952	815	137	137
19	1024	886	138	138
20	1094	955	139	139
21	1142	1001	141	141

*Ceiling Effect **Class Median

Students who received the maximum possible score on the IE posttest (1203 in this case) are identified as having experienced a ceiling effect. To determine whether the student's observed growth should be adjusted, three values are considered: the student's observed growth, the growth expectation, and the median observed growth for the class. If necessary, the student's observed growth is adjusted to reflect the highest of the three values.

In the sample below, two students experienced a ceiling effect on the IE posttest (highlighted in blue). For Student 9, the student's observed growth (102) was higher than the growth expectation (89), but lower than the class median (112). Since the student's observed growth may not reflect the student's true growth and could negatively impact the CGA, the student's observed growth is adjusted to the class median (112). For Student 15, the student's observed growth (125) is higher than the growth expectation (89) and the class median (112).

This student already has a positive impact on the CGA, so the student's observed growth score is not adjusted.

Step 4: Performing the Repeated-Measures t-Test and Assigning a Categorization

To perform the repeated measures t-test, an observed t-statistic (t_{obs}) is calculated using the following equation. This observed t-statistic is then compared to a critical t-statistic (t_{crit}) that indicates the value that represents a desired level of significance for the sample. In keeping with convention, the required level of significance for the t-test used as part of CGA is set at $\alpha=0.05$. Since the CGA is designed to detect a significant positive difference as well as a significant negative difference, a two-tailed test is performed.

$$t_{obs} = \frac{\overline{OG} - EG}{\frac{s_{OG}}{\sqrt{n}}}$$

At this point, it is useful to put together a set of descriptive statistics for the observed growth scores including the mean, standard deviation, and count of observations. These descriptive statistics will be used along with the growth expectation and the critical t-statistic to perform the repeated measures t-test.

TABLE 2
Values for use in calculating the t-test

Values to be Used in Calculations for the t-Test	
Mean Observed Growth (\overline{OG})	104.6
Standard Deviation of Observed Growth (s_{OG})	32.7
Count of Observations (n)	21
Growth Expectation (EG)	89
Critical t-statistic (t_{crit})	2.09

Plugging the values into the equation as illustrated below yields an observed t-statistic of 2.18 for Mrs. Smith's class.

$$t_{obs} = \frac{\overline{OG} - EG}{\frac{s_{OG}}{\sqrt{n}}} = \frac{104.6 - 89}{\frac{32.7}{\sqrt{21}}} = 2.18$$

Since the observed t-statistic is greater than the critical t-statistic (2.09), the t-test indicates that there is a significant difference between the observed growth for the class and the growth expectation. Note that the observed t-statistic is positive indicating that the average observed growth (104.6) was higher than the growth expectation (89). Therefore, Mrs. Smith's fifth grade math class receives a categorization of "Expected Growth Exceeded."

