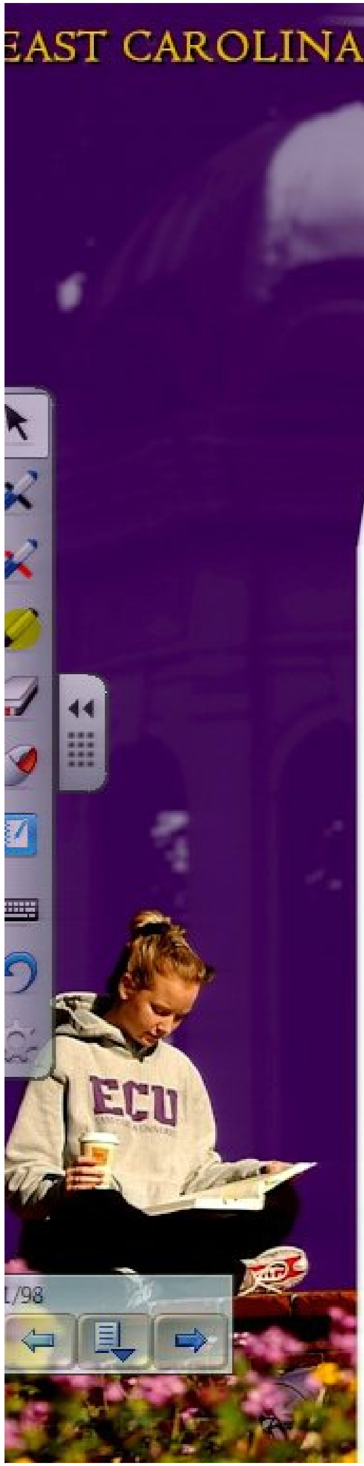


Improving Statistical Methods for Better Quantitative Research and Evaluation

ECU Five-Year Research Achievement Award Seminar

Guili Zhang, PhD
Research and Evaluation Methods
Department of Curriculum and Instruction
College of Education
East Carolina University
zhangg@ecu.edu

My Research Focuses

2/98

My Research Focuses

Advancing

statistical & measurement methods

EAST CAROLINA UNIVERSITY

My Research Focuses

Advancing

statistical & measurement methods

Applying

Outline

Outline

- Advancing statistical and measurement methods.

Outline

- Advancing statistical and measurement methods.
- Applying statistical and measurement methods in research and evaluation (7 example studies)

EAST CAROLINA UNIVERSITY

Outline

- Advancing statistical and measurement methods.
- Applying statistical and measurement methods in research and evaluation (7 example studies)
- Grants

Outline

- Advancing statistical and measurement methods.
- Applying statistical and measurement methods in research and evaluation (7 example studies)
- Grants
- National leadership services

Outline

- Advancing statistical and measurement methods.
- Applying statistical and measurement methods in research and evaluation (7 example studies)
- Grants
- National leadership services
- National awards and honors

3/98

EAST CAROLINA UNIVERSITY

Improving
Statistical & Measurement Methods

4/98

Improving
Statistical & Measurement Methods

- ***t*-test**

  What is wrong and how to improve it?

4/98

EAST CAROLINA UNIVERSITY

Improving
Statistical & Measurement Methods

- ***t*-test**

  What is wrong and how to improve it?

- ***F*-test**

  What is wrong and how to improve it?

4/98

EAST CAROLINA UNIVERSITY

Improving
Statistical & Measurement Methods

- ***t*-test**

  What is wrong and how to improve it?

- ***F*-test**

  What is wrong and how to improve it?

- **Kenward-Roger *F* test**

  Does it work well?

Improving
Statistical & Measurement Methods

- ***t*-test**

  What is wrong and how to improve it?

- ***F*-test**

  What is wrong and how to improve it?

- **Kenward-Roger *F* test**

  Does it work well?

- **Cronbach's α for Internal Consistency**

  Can we do better?

4/98

EAST CAROLINA UNIVERSITY

Improving
Statistical & Measurement Methods

- ***t*-test**

  What is wrong and how to improve it?

- ***F*-test**

  What is wrong and how to improve it?

- **Kenward-Roger *F* test**

  Does it work well?

- **Cronbach's α for Internal Consistency**

  Can we do better?

- **Formative and Summative Evaluation**

  Is it systematic enough?

What is *t*-test?

What is *t*-test?

- The ***t*-test assesses whether the means of two groups are different** from each other.

- **Introduced in 1908** in *Biometrika* **by William Gosset**, a chemist working for the Guinnes Brewery in Dublin, Ireland.

"Student" was Gosset's pen name, that's where the term "Student's *t* Distribution" came from.

- Gosset devised the *t*-test as a cheap way to monitor the quality of stout.

EAST CAROLINA UNIVERSITY

What is Wrong with *t*-test?

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

EAST CAROLINA UNIVERSITY

What is Wrong with *t*-test?

- **It depends** (on sample size)**!**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}}$$

What is Wrong with *t*-test?

- **It depends** (on sample size)**!**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- **It doesn't tell us what we want to know.**
  - What we want to know: Given these data, what is the probability that the null hypothesis is true?
  - What it tells us: Given the null hypothesis is true, what is the probability of these (or more extreme) data?

What is Wrong with *t*-test?

- **It depends** (on sample size)**!**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}}$$

- **It doesn't tell us what we want to know.**
  - What we want to know: Given these data, what is the probability that the null hypothesis is true?
  - What it tells us: Given the null hypothesis is true, what is the probability of these (or more extreme) data?
- **It tells us the direction, not the amount.**
  - P-value does not tell you the amount.
  - The ritual dichotomous reject-accept decision is not the way any science is done. Physical scientists have learned much by storing up the amounts, not just directions. Measuring things on a communicable scale let us stockpile information about amounts.

What is Wrong with *t*-test?

- **It depends** (on sample size)**!**

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- **It doesn't tell us what we want to know.**
  - What we want to know: Given these data, what is the probability that the null hypothesis is true?
  - What it tells us: Given the null hypothesis is true, what is the probability of these (or more extreme) data?
- **It tells us the direction, not the amount.**
  - P-value does not tell you the amount.
  - The ritual dichotomous reject-accept decision is not the way any science is done. Physical scientists have learned much by storing up the amounts, not just directions. Measuring things on a communicable scale let us stockpile information about amounts.
- **The null hypothesis is always false** (given big enough sample size)!

5/98

EAST CAROLINA UNIVERSITY

What is Wrong with *t*-test?

- **It depends** (on sample size)!

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- **It doesn't tell us what we want to know.**
  - What we want to know: Given these data, what is the probability that the null hypothesis is true?
  - What it tells us: Given the null hypothesis is true, what is the probability of these (or more extreme) data?
- **It tells us the direction, not the amount.**
  - P-value does not tell you the amount.
  - The ritual dichotomous reject-accept decision is not the way any science is done. Physical scientists have learned much by storing up the amounts, not just directions. Measuring things on a communicable scale let us stockpile information about amounts.
- **The null hypothesis is always false** (given big enough sample size)!
- **So, essentially, it's testing whether there are a lot of subjects!**

5/98

EAST CAROLINA UNIVERSITY

How to Improve *t*-test?

$$d = \frac{\bar{Y}_2 - \bar{Y}_1}{S}$$

$$\delta = \frac{\mu_1 - \mu_1}{\sigma}$$

EAST CAROLINA UNIVERSITY

How to Improve *t*-test?

- **Use Effect Size (ES)** **(one of the most commonly used is Cohen's *d*):**

$$d = \frac{\bar{Y}_2 - \bar{Y}_1}{S} \quad \text{estimates} \quad \delta = \frac{\mu_1 - \mu_1}{\sigma}$$

Sizes of the effect:

- **Small**: 0.2-0.3
- **Medium:** around 0.5
- **Large**: 0.8 to infinity

How to Improve *t*-test?

- **Use Effect Size (ES)** **(one of the most commonly used is Cohen's *d*):**

$$d = \frac{\bar{Y}_2 - \bar{Y}_1}{S} \quad \text{estimates} \quad \delta = \frac{\mu_1 - \mu_1}{\sigma}$$

Sizes of the effect:

- **Small**: 0.2-0.3
- **Medium:** around 0.5
- **Large**: 0.8 to infinity

- **Use Confidence Interval (CI) for Effect Size**

How to Improve *t*-test? cont'd

How to Improve *t*-test? cont'd

Why Confidence Interval (CI) for Effect Size?

How to Improve *t*-test? cont'd

Why Confidence Interval (CI) for Effect Size?

- A **CI contains all the information found in the significance tests**.

How to Improve *t*-test? cont'd

Why Confidence Interval (CI) for Effect Size?

- A **CI contains all the information found in the significance tests**.
- **A CI contains vital information** not provided by the significance tests: **magnitude of effects and precision of estimates.**

3/98

EAST CAROLINA UNIVERSITY

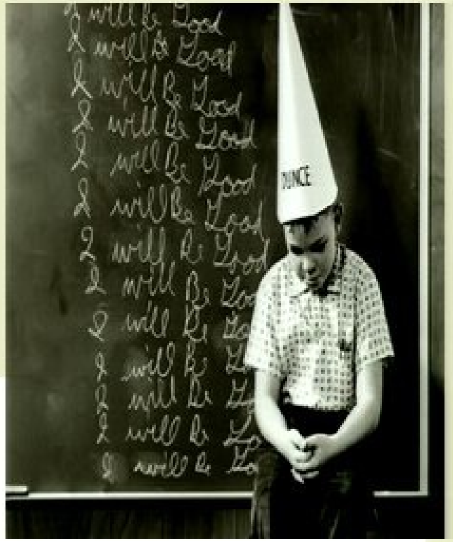
How to Improve *t*-test? cont'd

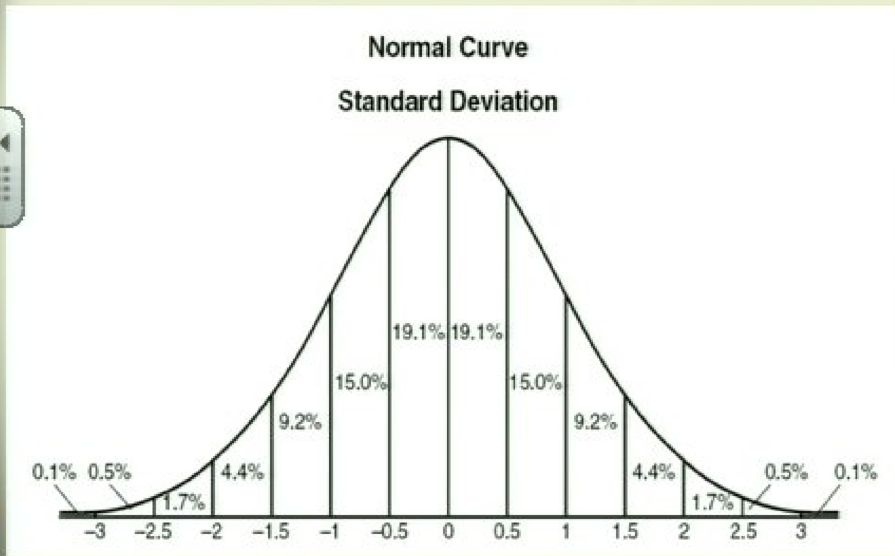
Why Confidence Interval (CI) for Effect Size?

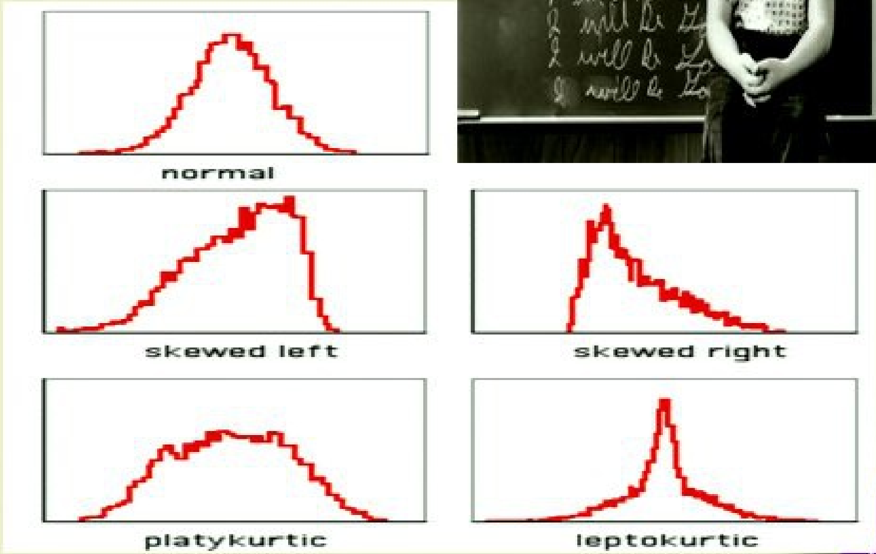
- A **CI contains all the information found in the significance tests**.
- **A CI contains vital information** not provided by the significance tests: **magnitude of effects and precision of estimates.**
- **A CI indicates the range of population ESs** with which the data are consistent. A significance test merely indicates whether the data are consistent with a population ES of zero.

How to Improve *t*-test? cont'd

What happens when data "misbehave"?



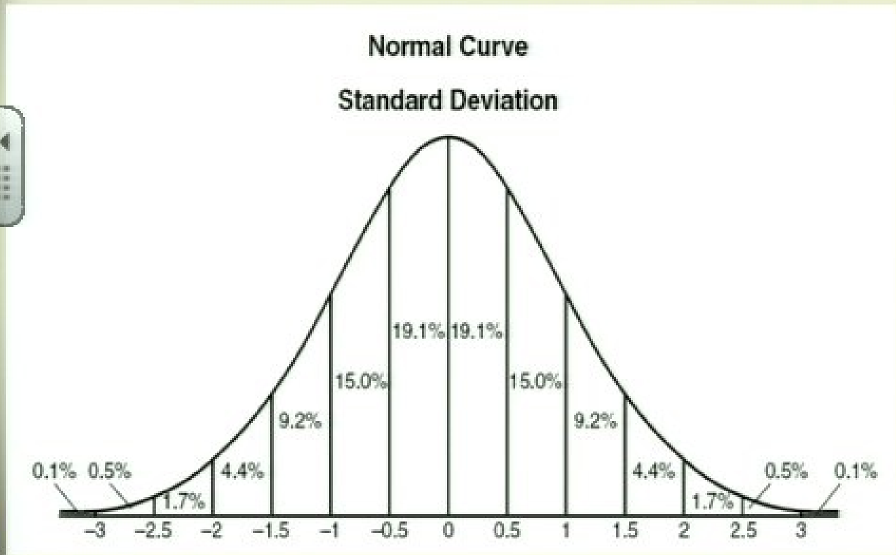


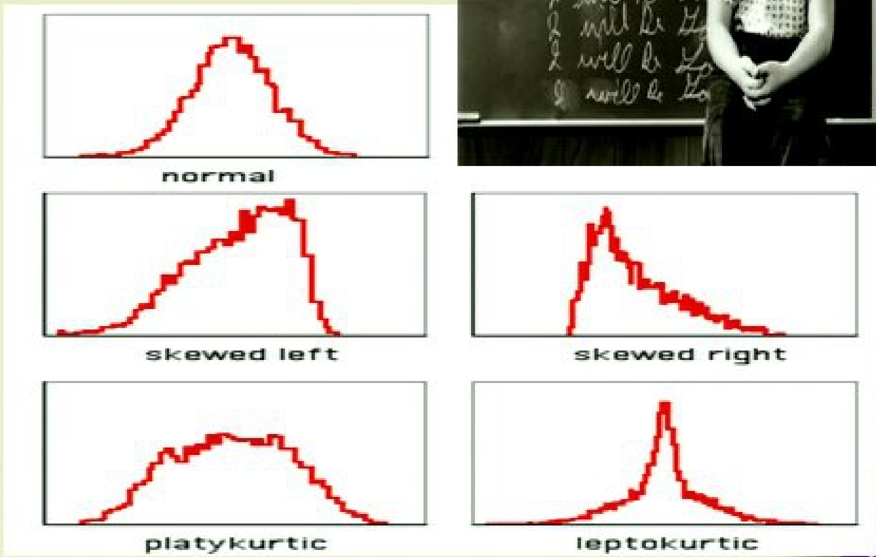


EAST CAROLINA UNIVERSITY

How to Improve *t*-test? cont'd

What happens when data "misbehave"?

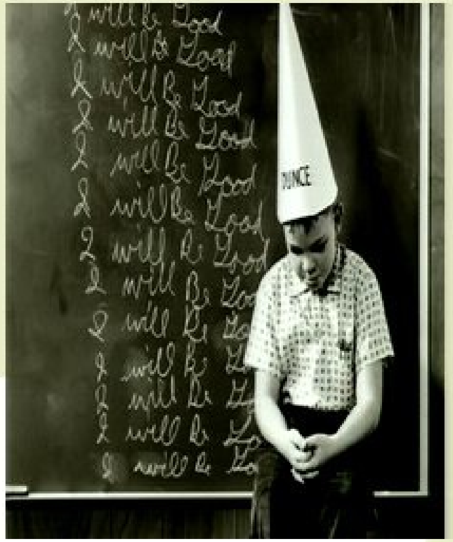


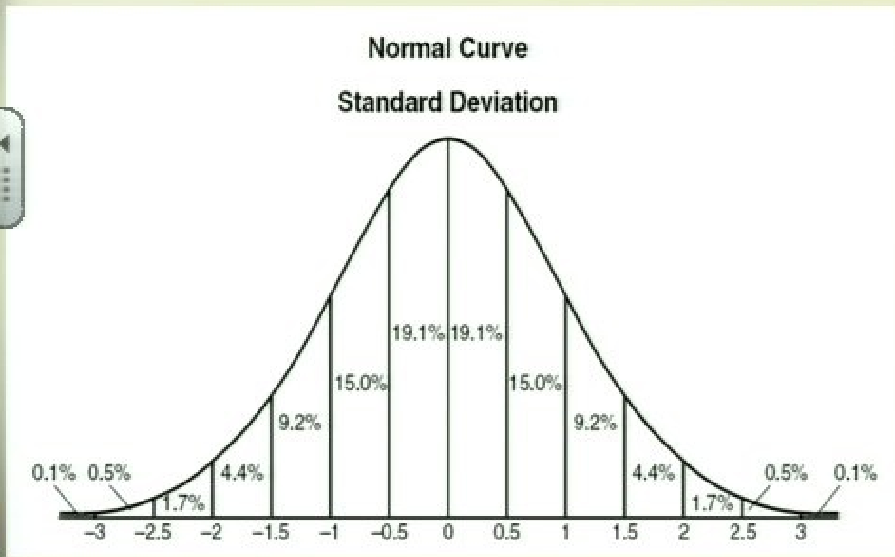


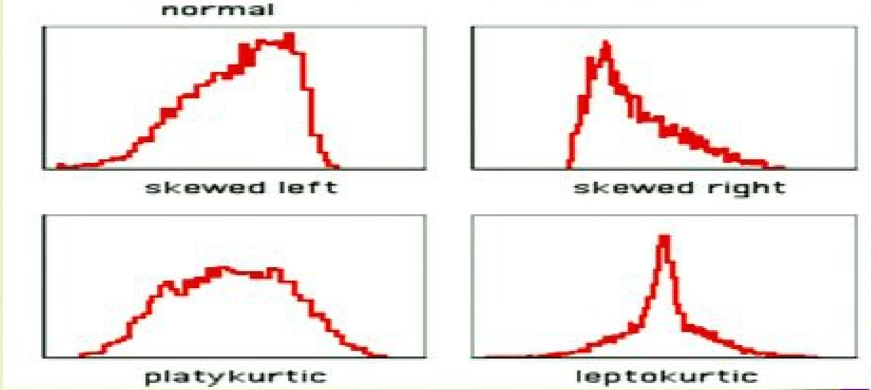
Effect Size can become an **inadequate measure of group separation**.

How to Improve *t*-test? cont'd

What happens when data "misbehave"?







Effect Size can become an **inadequate measure of group separation.**

Then what do we do?

How to Improve *t*-test? cont'd

- **Use Robust Effect Size** (Algina et al., 2005)

$$d_R = .642\left(\frac{\bar{Y}_{t2} - \bar{Y}_{t1}}{S_W}\right) \textbf{ estimates } \delta_R = .642\left(\frac{\mu_{t2} - \mu_{t1}}{\sigma_W}\right)$$

*The multiplier .642 is the Winsorized standard deviation of a standard normal distribution, and is used to ensure that when both samples being compared are drawn from normal distributions with equal variances.

- **Use Confidence Intervals for Robust Effect Size**
  - Noncentral *t*-distribution based CI
  - Percentile bootstrap CI

What is F-test? What is Wrong with Significance Testing?

What is *F*-test? What is Wrong with Significance Testing?

- The *F*-test is one of the most commonly used significance tests for comparing three or more groups' means. Developed by Sir Ronald Fisher in the 1920s as the variance ratio.

- Significance testing – the mechanical dichotomous decision around a sacred .05 criterion, has been severely criticized for the last five decades.

- A great deal of **mischief** has been associated with the test of significance (Bakan, 1966).

- Significance testing has not only failed to support the advance of social science as a science, but also has seriously impeded it (Cohen, 1994).

How to Improve *F*-test?
Root Mean Square Standardized Effect Size (RMSSE)?

In a balanced, one-way, between-subjects, fixed-effects design, Root Mean Square Standard Effect Size (RMSSE), denoted by f^*, is defined by Steiger & Fouladi (1997) as follows:

$$f^* = \sqrt{\frac{\sum_{j=1}^{J} (\mu_j - \mu)^2}{(J-1)\sigma^2}}$$

where μ_j is the mean for the j th level, μ is the grand mean, and σ^2 is the within-level variance, which is assumed to be constant across levels.

12/98

EAST CAROLINA UNIVERSITY

Estimate RMSSE

Based on expected mean squares in a balanced design, f^* can be estimated by using

$$\hat{f}^* = \sqrt{\frac{MS_B - MS_W}{nMS_W}} = \sqrt{\frac{1}{n}(F-1)} \qquad \text{(A)}$$

if $F \geq 1$ and by using $\hat{f}^* = 0$, otherwise. Alternatively, based on the expected value of F under normality f^* can be estimated by using

$$\hat{f}^* = \sqrt{\frac{(N-J-2)}{n(N-J)}(F-1)} \qquad \text{(B)}$$

if $F \geq 1$ and $\hat{f}^* = 0$ otherwise. Both estimates are very similar but the estimate in Equation A was used in our study because it does not require the normality assumption.

EAST CAROLINA UNIVERSITY

Noncentral *F* Distribution-Based CI for RMSSE

The CIs for f^* can be constructed based on the noncentral *F* distribution. In a one-way, between-subjects, fixed-effects ANOVA, the *F* statistic with $J-1$ and $N-J$ degrees of freedom has noncentrality parameter

$$\lambda = \frac{\sum_{j=1}^{J} n_j(\mu_j - \mu)^2}{\sigma^2}.$$

Clearly in a balanced design

$$f^* = \sqrt{\frac{\lambda}{n(J-1)}}.$$

To find a 95% CI for f^*, we first use the noncentral *F* distribution to find a 95% CI for λ. We then transform the endpoints of the CI for λ by dividing λ by $(J-1)n$ and then take the square root. The result is an exact CI for f^*.

14/98

EAST CAROLINA UNIVERSITY

How well does the CI for RMSSE Perform? The Investigation

- The noncentral F distribution-based and the percentile bootstrap CIs were implemented for all combinations of the following five factors:
  - five population distributions including normal distribution and 4 nonnormal distributions from the family of the g and h distributions;
  - two numbers of levels for treatment groups: $J = 3$ and $J = 6$;
  - three cell sample sizes in each treatment: 20, 35, and 50;
  - six values of population RMSSEs: 0, .1, .25, .40, .55, and .70;
  - two mean configurations: the equally spaced mean configuration and the one extreme mean configuration.

- The nominal confidence level for all intervals investigated was .95 and each condition was replicated 2500 times.
- The number of bootstrap replications in the bootstrap procedure was 1000.

15/98

EAST CAROLINA UNIVERSITY

Coverage performance of CIs for RMSSE under Normality

- When sampling from a normal distribution, the coverage probability of the noncentral *F* distribution-based CI should be .975 when $f^* = 0$, and the results in are consistent with the theory.

- When $f^* > 0$, the coverage probability of the noncentral *F* distribution-based CI is expected to be .95 under normality and the results are consistent with this expectation.

- The percentile bootstrap CI also work adequately under normality.

What Happens when Data "Misbehave"?

- Zhang and Algina (2011) investigated the coverage performance of the noncentral *F* distribution-based CI and the percentile bootstrap CI for RMSSE in a one-way, fixed-effects, between-subjects ANOVA.

- We found that both the noncentral *F* distribution-based CI and the percentile bootstrap CI for RMSSE yielded inadequate coverage probabilities under data nonnormality.

So what do we do now?

EAST CAROLINA UNIVERSITY

Development of the Robust Root Mean Square Standardized Effect Size ($RMSSE_R$) and its CIs

To overcome the weaknesses in f^*, we developed a robust version of the generalized effect size, the Robust Root Mean Square Standardized Effect Size ($RMSSE_R$), denoted by f_R^* in our study.

In a balanced one-way between-subjects ANOVA design, f_R^* is defined as

$$f_R^* = .642\sqrt{\frac{\sum_{j=1}^{J}\left(\mu_{Tj} - \mu_T\right)^2}{(J-1)\sigma_W^2}},$$

where μ_{Tj} is the trimmed mean for the j th level, μ_T is the grand mean based on the trimmed means, and $\sigma_W^{\;2}$ is the within-level Winsorized variance, which is assumed to be constant across levels. The quantity .642 is the square root of the population Winsorized variance for a standard normal distribution. Therefore, including .642 in the definition of the robust effect ensures that $f_R^* = f^*$ when the data are drawn from normal distributions with equal variances.

Estimate RMSSE$_R$

An estimate of f_R^* can be attained from sample statistics by applying the following

formula:

$$\hat{f}_R^* = .642 \times \sqrt{\frac{\sum_{j=1}^{J}\left(\bar{Y}_{Tj} - \bar{Y}_T\right)^2}{(J-1)S_{Wp}^2}}, \qquad (2)$$

where $\bar{Y}_{Tj}$ is the trimmed sample mean for the j th level, $\bar{Y}_T$ is the sample grand trimmed mean, and $S_{Wp}{}^2$ is the sample pooled within-level Winsorized variance.

Noncentral *F* distribution-based CI for RMSSE_R

In a balanced one-factor between-subject design with equal *n*s, f_R^* can be written as a function of λ_R:

$$f_R^* = \sqrt{\frac{.4129 \times \left(\sum_{j=1}^{J} n_j - J\right)}{\left(\sum_{j=1}^{J} h_j - J\right) \times (J-1)h} \lambda_R} . \qquad (9)$$

To find a (1 - α)% (95% in this study) CI for f_R^*, we first use the noncentral *F* distribution to find a 95% CI for λ_R. Once the CI on λ_R is found, we then apply Equation 9 to transform the endpoints of the CI for λ_R to obtain the endpoints for the CI for f_R^*.

Percentile Bootstrap CI for RMSSE$_R$

To apply the percentile bootstrap method, the following steps are completed 1000 times within each replication of a condition.

1. A sample of size n_j is randomly selected with replacement from the scores for the group j, $j = 1, \ldots, J$. These J samples are combined to form a bootstrap sample.
2. The parameter f_R^{*2} is estimated by using

$$\widehat{f}_R^{*2} = \frac{.642^2}{n} \frac{\sum_{j=1}^{J} n_j - 3}{\sum_{j=1}^{J} h_j - 3} (F_R - 1). \tag{15}$$

3. The 1000 f_R^{*2} estimates are ranked from low to high. The lower limit of the CI for f_R^{*2} is determined by finding the 26th estimate in the rank order [i.e., the (.025x1000+1)th estimate]; and the 975th estimate is the upper limit of the CI for f_R^{*2} (i.e. the (.975x1000)th estimate].
4. The lower limit of the CI for f_R^{*} is equal to the square root of the lower limit of the CI for f_R^{*2} if the latter lower limit is larger than zero and is zero otherwise. The upper limit of the CI for f_R^{*} is equal to the square root of the upper limit of the CI for f_R^{*2}.

Investigating the Performance of $RMSSE_R$

- The noncentral *F* distribution-based and the percentile bootstrap CIs were implemented for all combinations of the following five factors:
  - five population distributions including the normal distribution and four nonnormal distributions from the family of the *g* and *h* distributions;
  - two numbers of levels for treatment groups: $J = 3$ and $J = 6$;
  - three cell sample sizes in each treatment: 20, 35, and 50;
  - six values of population $RMSSE_R$: 0, .1, .25, .40, .55, and .70;
  - two mean configurations: the equally spaced mean configuration and the one extreme mean configuration.

- The nominal confidence level for all intervals investigated was .95 and each condition was replicated 2500 times.

- The number of bootstrap replications in the bootstrap procedure was 1000.

EAST CAROLINA UNIVERSITY

Conclusions on $RMSSE_R$

- The noncentral F distribution-based CIs for f_R^*, which was proposed in the current study and was formulated with the robust parameters including the trimmed means and Winsorized variances, yielded fairly adequate coverage probabilities and better coverage probability than the percentile bootstrap CI.

- Accordingly, researchers who want to set a CI for f_R^* can use the CI constructed by using the noncentral F distribution. These researchers will enjoy the additional benefit of using a robust measure of effect size, that is, a measure that is not likely to be strongly affected by outlying data points.

Conclusions on *t*-test, *F*-test, Effect Size, and Robust Effect Size

EAST CAROLINA UNIVERSITY

Conclusions on *t*-test, *F*-test, Effect Size, and Robust Effect Size

- ***t*-test *and F*-test depend on sample size, and fail to provide the magnitude of effects and precision of estimation.**

24/98

EAST CAROLINA UNIVERSITY

Conclusions on *t*-test, *F*-test, Effect Size, and Robust Effect Size

- ***t*-test *and F*-test depend on sample size, and fail to provide the magnitude of effects and precision of estimation.**

- CI for RMSSE in ANOVA provides the magnitude of group difference and precision of estimation. However, it yields inadequate coverage probabilities under data non-normality.

24/98

EAST CAROLINA UNIVERSITY

Conclusions on *t*-test, *F*-test, Effect Size, and Robust Effect Size

- ***t*-test *and F*-test depend on sample size, and fail to provide the magnitude of effects and precision of estimation.**

- CI for RMSSE in ANOVA provides the magnitude of group difference and precision of estimation. However, it yields inadequate coverage probabilities under data non-normality.

- **The $RMSSE_R$ and Noncentral F-distribution based CI for $RMSSE_R$ I developed perform more satisfactorily under various conditions**. Reporting it will be more accurate than the commonly used RMSSE.

24/98

EAST CAROLINA UNIVERSITY

Conclusions on *t*-test, *F*-test, Effect Size, and Robust Effect Size

- ***t*-test *and F*-test depend on sample size, and fail to provide the magnitude of effects and precision of estimation.**

- CI for RMSSE in ANOVA provides the magnitude of group difference and precision of estimation. However, it yields inadequate coverage probabilities under data non-normality.

- **The $RMSSE_R$ and Noncentral F-distribution based CI for $RMSSE_R$ I developed perform more satisfactorily under various conditions**. Reporting it will be more accurate than the commonly used RMSSE.

*******This string of research is published in:***

- *Journal of Modern Applied Statistical Methods,*
- *Middle Grades Research Journal,*
- *Research Supporting Middle Grades Research* (book).

Performance of Kenward-Roger *F*-test

EAST CAROLINA UNIVERSITY

Performance of Kenward-Roger *F*-test

- We investigated the Type I error rate of the Kenward-Roger (KR) *F*-test, using SAS PROC MIXED, through a simulation study for a between- by within-subjects factor split-plot design with non-normal ignorable missing data MCAR (missing completely at random) and MAR (missing at random).

25/98

EAST CAROLINA UNIVERSITY

Performance of Kenward-Roger *F*-test

- We investigated the Type I error rate of the Kenward-Roger (KR) *F*-test, using SAS PROC MIXED, through a simulation study for a between- by within-subjects factor split-plot design with non-normal ignorable missing data MCAR (missing completely at random) and MAR (missing at random).

- The KR *F*-test for the between-subjects factor is robust, in terms of controlling the Type I error, to non-normal ignorable missing data.

EAST CAROLINA UNIVERSITY

Performance of Kenward-Roger *F*-test

- We investigated the Type I error rate of the Kenward-Roger (KR) *F*-test, using SAS PROC MIXED, through a simulation study for a between- by within-subjects factor split-plot design with non-normal ignorable missing data MCAR (missing completely at random) and MAR (missing at random).

- The KR *F*-test for the between-subjects factor is robust, in terms of controlling the Type I error, to non-normal ignorable missing data.

- The KR *F*-test for the within-subjects factor is robust to non-normal data as long as the missing data mechanism is MCAR.

EAST CAROLINA UNIVERSITY

Performance of Kenward-Roger *F*-test

- We investigated the Type I error rate of the Kenward-Roger (KR) *F*-test, using SAS PROC MIXED, through a simulation study for a between- by within-subjects factor split-plot design with non-normal ignorable missing data MCAR (missing completely at random) and MAR (missing at random).
- The KR *F*-test for the between-subjects factor is robust, in terms of controlling the Type I error, to non-normal ignorable missing data.
- The KR *F*-test for the within-subjects factor is robust to non-normal data as long as the missing data mechanism is MCAR.
- The KR *F*-test for the within- by between-subjects interaction performed fairly well under MCAR. Unfortunately, it is not robust under MAR.

Performance of Kenward-Roger *F*-test

- We investigated the Type I error rate of the Kenward-Roger (KR) *F*-test, using SAS PROC MIXED, through a simulation study for a between- by within-subjects factor split-plot design with non-normal ignorable missing data MCAR (missing completely at random) and MAR (missing at random).

- The KR *F*-test for the between-subjects factor is robust, in terms of controlling the Type I error, to non-normal ignorable missing data.

- The KR *F*-test for the within-subjects factor is robust to non-normal data as long as the missing data mechanism is MCAR.

- The KR *F*-test for the within- by between-subjects interaction performed fairly well under MCAR. Unfortunately, it is not robust under MAR.

****This research is supported in part by NIH grant: R01AR052658-03S1*

In collaboration with Padilla & Min.

Published in Journal of Modern Applied Statistical Methods

EAST CAROLINA UNIVERSITY

Cronbach's α

- A typical way of measuring a latent construct is through a scale or questionnaire containing items indirectly measuring the construct.

- It's important that the items be consistent or reliable so that the questionnaire itself is consistent or reliable.

- **Cronbach's α,** aka **coefficient alpha**, was introduced by **Lee Cronbach** in 1951. It's the most commonly used method of measuring reliability of a scale. It will generally increase as the intercorrelations among test items increase, and is thus known as an internal consistency estimate of reliability of test scores.

$$\hat{\alpha}_c = \frac{p}{p-1}\left(1 - \frac{\sum_{i=1}^{p} \hat{\sigma}_{ii}}{\sum_{i=1}^{p}\sum_{j=1}^{p} \hat{\sigma}_{ij}}\right) \text{ estimates } \alpha_c = \frac{p}{p-1}\left(1 - \frac{\sum_{i=1}^{p} \sigma_{ii}}{\sum_{i=1}^{p}\sum_{j=1}^{p} \sigma_{ij}}\right)$$

Cronbach's α for Internal Consistency

Levels of Internal Consistency

Cronbach's alpha	Internal consistency
$\alpha \geq .9$	Excellent
$.9 > \alpha \geq .8$	Good
$.8 > \alpha \geq .7$	Acceptable
$.7 > \alpha \geq .6$	Questionable
$.6 > \alpha \geq .5$	Poor
$.5 > \alpha$	Unacceptable

Cronbach's α for Reliability: can we do better?

28/98

EAST CAROLINA UNIVERSITY

Cronbach's α for Reliability: can we do better?

- Reporting only a point estimate of coefficient alpha is no longer sufficient. Confidence Intervals for coefficient alpha were first introduced in 1960s and further developed.

28/98

EAST CAROLINA UNIVERSITY

Cronbach's α for Reliability:
can we do better?

- Reporting only a point estimate of coefficient alpha is no longer sufficient. Confidence Intervals for coefficient alpha were first introduced in 1960s and further developed.

However, all current coefficient alpha CIs are frequentist-based and thus have the traditional, less desirable CI interpretation: we are 95% confident that the interval captures the true population parameter.

28/98

EAST CAROLINA UNIVERSITY

Cronbach's α for Reliability: can we do better?

- Reporting only a point estimate of coefficient alpha is no longer sufficient. Confidence Intervals for coefficient alpha were first introduced in 1960s and further developed.

- However, all current coefficient alpha CIs are frequentist-based and thus have the traditional, less desirable CI interpretation: we are 95% confident that the interval captures the true population parameter.

- Current coefficient alpha CIs cannot use prior information to stabilize inferences or update information.

Can we do Better?
Bayesian Coefficient Alpha

- Traditional frequentist analyses are only composed of data, **Bayesian** analysis is composed of data and prior knowledge and/or beliefs. Through the combination of data and prior knowledge, more can be learned about the phenomenon under study and knowledge can be updated accordingly.

- We developed a **Bayesian coefficient alpha** $\alpha_b = E(\alpha_c \mid \mathbf{y})$

where

$$\alpha_c^{(t)} = \frac{p}{p-1}\left(1 - \frac{\sum_{i=1}^{p} \sigma_{ii}^{(t)}}{\sum_{i=1}^{p}\sum_{j=1}^{p} \sigma_{ij}^{(t)}}\right)$$

- Bayesian credible intervals can then be obtained by the lower $\alpha/2$ and upper $1-\alpha/2$ percentiles of the sample.

29/98

EAST CAROLINA UNIVERSITY

Investigating Performance of Bayesian Coefficient Alpha

- **A 4 × 3 × 6 Monte Carlo simulation design was utilized to investigate the properties of Bayesian coefficient alpha.**
  - First, the number of items was investigated: 5, 10, 15 and 20.
  - Second, the mean item correlation was investigated: 0.173, 0.223, and 0.314. These mean items correlations were investigated because they generate coefficient alphas that range from 0.50 to 0.90.
  - Third, sample size was also explored: 50, 100, 150, 200, 250 and 300.
  - Multivariate normal data were generated with mean vector zero and correlation matrix R of dimensions defined by the number of items in the simulation.
  - For each condition of the simulation study 1,000 replications were obtained. In each replication, Bayesian Coefficient Alpha was computed along with the SE and 95% BCIs.
- **The results from our Monte Carlo investigations indicate that the Bayesian Coefficient Alpha was relatively unbiased under all investigated conditions.**

EAST CAROLINA UNIVERSITY

Advantages of Bayesian Coefficient Alpha

- Bayesian Coefficient Alpha has the advantage of **having the credible intervals** (BCIs), which have the interpretation researchers really want to make with CIs: we are 95% confident that the true population parameter lies between the bands of the credible interval, a simpler and more powerful statement.

31/98

EAST CAROLINA UNIVERSITY

Advantages of Bayesian Coefficient Alpha

- Bayesian Coefficient Alpha has the advantage of **having the credible intervals** (BCIs), which have the interpretation researchers really want to make with CIs: we are 95% confident that the true population parameter lies between the bands of the credible interval, a simpler and more powerful statement.

- Bayesian Coefficient Alpha can **make use of prior information to stabilize inferences.** Through the combination of data and prior knowledge, more can be learned about the phenomenon under study and knowledge about the phenomenon can be updated accordingly.

EAST CAROLINA UNIVERSITY

Advantages of Bayesian Coefficient Alpha

- Bayesian Coefficient Alpha has the advantage of **having the credible intervals** (BCIs), which have the interpretation researchers really want to make with CIs: we are 95% confident that the true population parameter lies between the bands of the credible interval, a simpler and more powerful statement.

- Bayesian Coefficient Alpha can **make use of prior information to stabilize inferences.** Through the combination of data and prior knowledge, more can be learned about the phenomenon under study and knowledge about the phenomenon can be updated accordingly.

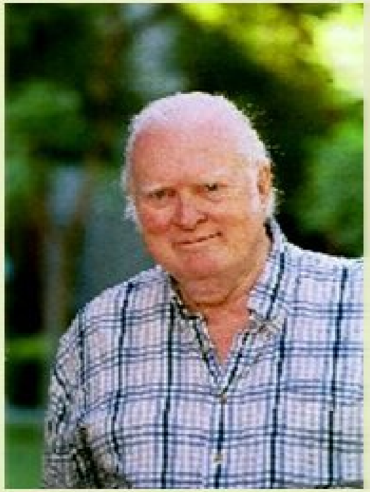
****Presented at the American Educational Research Association annual meeting*

Published in Journal of Modern Applied Statistical Methods

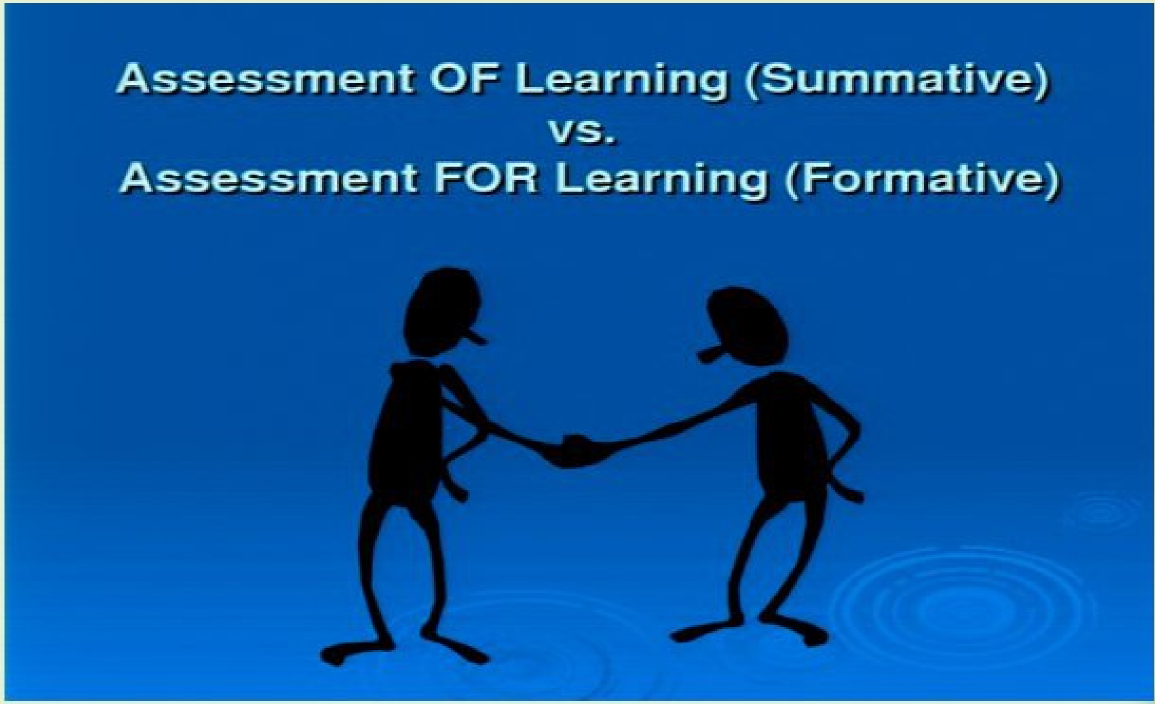
EAST CAROLINA UNIVERSITY

Formative & Summative Evaluation:
Is it systematic enough?

- Michael Scriven coined the terms in 1967.



- Formative evaluation provides information needed for planning a treatment, while summative evaluation provides summary type of judgment regarding the effectiveness of the treatment.

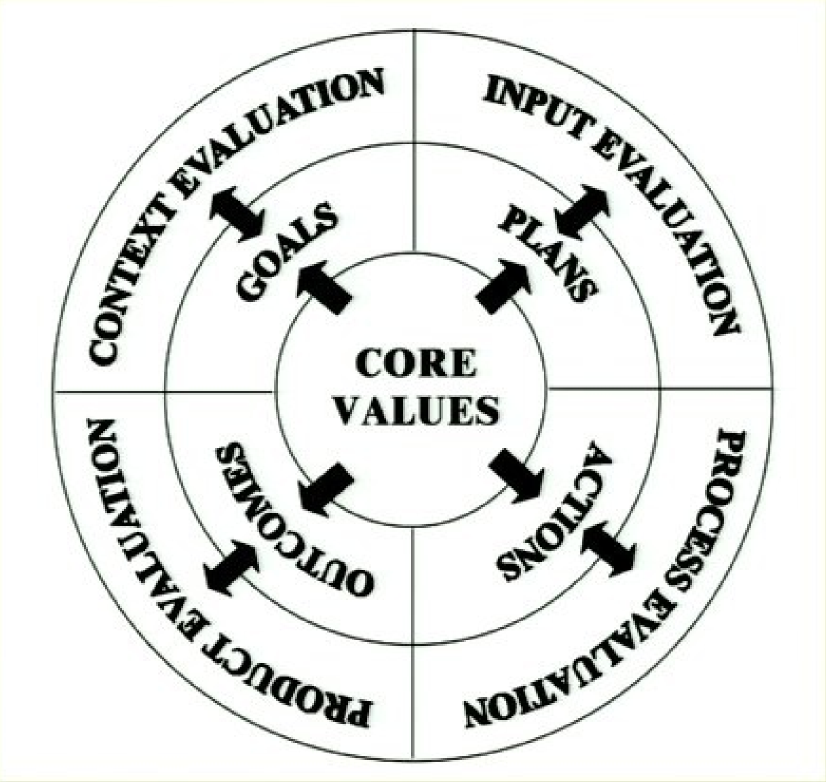


EAST CAROLINA UNIVERSITY

The CIPP Model

- Created in the 1960s by Daniel Stufflebeam, commonly known as the founding father of the evaluation profession.

- The CIPP model includes **C**ontext, **I**nput, **P**rocess, and **P**roduct evaluation.



33/98

EAST CAROLINA UNIVERSITY

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

- **Context Evaluation** is often referred to as needs assessment. It asks the question "What needs to be done?"

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

- **Context Evaluation** is often referred to as needs assessment. It asks the question "What needs to be done?"

- **Input Evaluation** helps prescribe a project to address the identified needs. It asks the question "How should it be done?"

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

- **Context Evaluation** is often referred to as needs assessment. It asks the question "What needs to be done?"

- **Input Evaluation** helps prescribe a project to address the identified needs. It asks the question "How should it be done?"

- **Process Evaluation** monitors the project implementation process. It asks the question "Is it being done?"

EAST CAROLINA UNIVERSITY

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

- **Context Evaluation** is often referred to as needs assessment. It asks the question "What needs to be done?”
- **Input Evaluation** helps prescribe a project to address the identified needs. It asks the question "How should it be done?”
- **Process Evaluation** monitors the project implementation process. It asks the question "Is it being done?"
- **Product Evaluation** identifies and assesses project outcomes. It asks the question "Did the project succeed?", and is similar to outcome evaluation.

34/98

EAST CAROLINA UNIVERSITY

The CIPP Model

as a Comprehensive Guiding Framework for Program Planning, Implementation and Evaluation

- **Context Evaluation** is often referred to as needs assessment. It asks the question "What needs to be done?"

- **Input Evaluation** helps prescribe a project to address the identified needs. It asks the question "How should it be done?"

- **Process Evaluation** monitors the project implementation process. It asks the question "Is it being done?"

- **Product Evaluation** identifies and assesses project outcomes. It asks the question "Did the project succeed?", and is similar to outcome evaluation.

****I am currently completing a book that provides a 21st century update of the CIPP Model with Dr. Stufflebeam.*

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1

Identifying Impact Factors

Applied Research Study Example #1

Identifying Impact Factors

- **Identifying factors influencing engineering student graduation and retention: A longitudinal and cross-institutional study.**

  Published in *Journal of Engineering Education.*

  Supported by a 10-year NSF grant (Grant No. EEC-9727411). NSF developed the Engineering Education Coalition (EEC) program to stimulate bold, innovative, and comprehensive models for systemic reform of undergraduate engineering education.

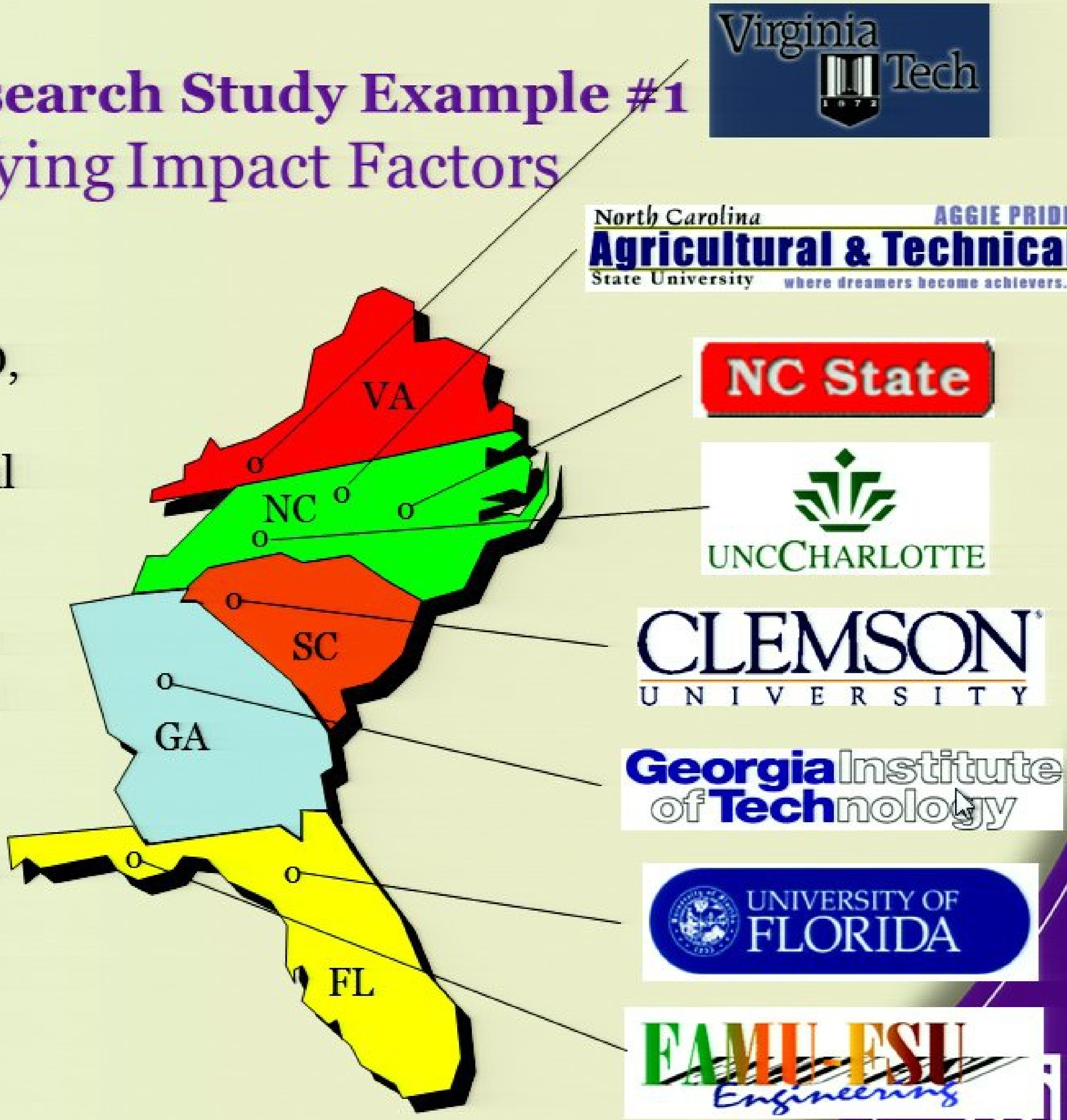
- A total of six EECs were established and funded for 10 years, involving 45 engineering institutions nationwide. The Southeastern University and College Coalition for Engineering Education (SUCCEED) is one of the six engineering Coalitions.

35/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

As the statistical analyst & program evaluator of SUCCEED, I created the SUCCEED longitudinal database (LDB), the largest engineering database of its kind in the U.S. It contains all undergraduate students from 9 universities: Clemson, Florida A&M, FSU, Georgia Tech, NCAT, NCSU, UFL, UNCC, and Virginia Tech.



EAST CAROLINA UNIVERSITY

Applied Research Study Example #1

Identifying Impact Factors

Applied Research Study Example #1
Identifying Impact Factors

Characteristics of SUCCEED Schools

- Enroll over 28,000 engineering undergraduates

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Characteristics of SUCCEED Schools

- Enroll over 28,000 engineering undergraduates
- Award 1/12 of all U.S. engineering degrees

Applied Research Study Example #1
Identifying Impact Factors

Characteristics of SUCCEED Schools

- Enroll over 28,000 engineering undergraduates
- Award 1/12 of all U.S. engineering degrees
- Award 1/4 of all U.S. engineering degrees awarded to Black students

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Characteristics of SUCCEED Schools

- Enroll over 28,000 engineering undergraduates
- Award 1/12 of all U.S. engineering degrees
- Award 1/4 of all U.S. engineering degrees awarded to Black students
- Award 1/10 of all U.S. engineering degrees awarded to women

37/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

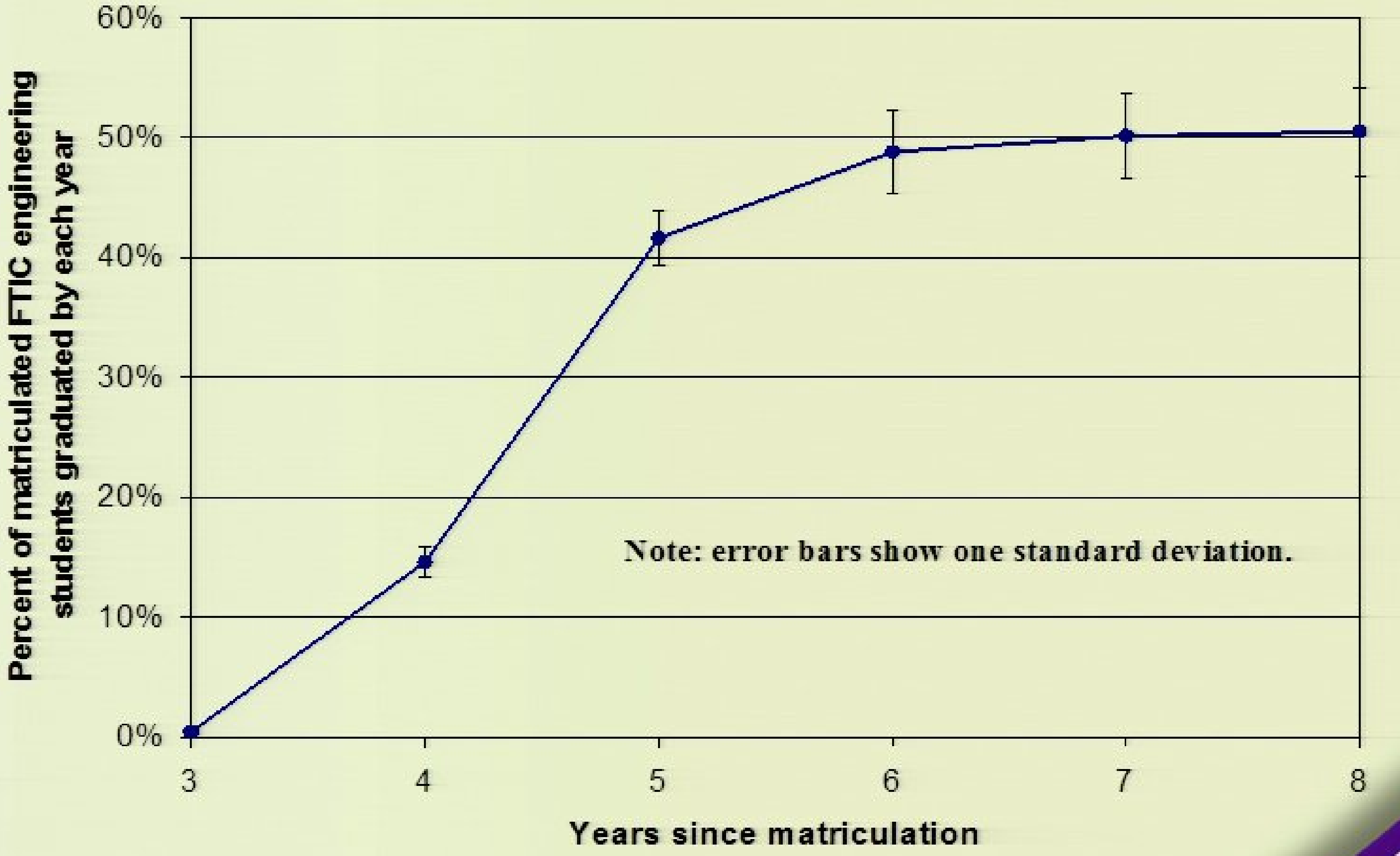
- **The study evaluated pre-existing factors' impacts** on engineering student success.
- **The data** were all engineering students at 9 universities spanning 15 years (87,167 engineering students) extracted from the LDB.
- **A multiple logistic regression model with step-wise selection procedure** is fitted to each institution's data to explore the relationship between graduation and the pre-existing factors. Multiple logistical regression techniques allow us to examine the effect of each predictor while controlling for the other variables.

Test Statistics used:
- Chi-square statistic
- Wald test and Wald confidence interval
- Odds ratio.
- Cochran-Mantel-Haenszel test for homogeneity of odds ratio.
- Coefficient of Determination

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Percentage of First Time in College (FTIC) Engineering Students Graduated by Years since Matriculation



EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Type III analysis of effects: Wald χ^2 and p-value.

	GENDER χ^2 (p-value)	HSGPA χ^2 (p-value)	SATQ χ^2 (p-value)	SATV χ^2 (p-value)	ETHNIC χ^2 (p-value)	CITIZEN χ^2 (p-value)
A	4.287 (0.038)	Not tested*	Not tested*	Not tested*	13.884 (0.031)	---
B	---	66.445 (<0.0001)	22.517 (<0.0001)	---	---	---
C	---	41.097 (<0.0001)	15.705 (<0.0001)	4.041 (0.044)	16.659 (0.005)	---
D	---	600.93 (<0.0001)	234.50 (<0.0001)	39.664 (<0.0001)	24.525 (<0.0001)	35.277 (<0.0001)
E	16.591 (<0.0001)	88.924 (<0.0001)	17.052 (<0.0001)	---	18.262 (0.0001)	---
F	11.365 (0.0007)	168.806 (<0.0001)	74.329 (<0.0001)	39.985 (<0.0001)	17.278 (0.002)	5.594 (0.018)
G	26.090 (<0.0001)	195.343 (<0.0001)	68.543 (<0.0001)	12.036 (0.0005)	11.944 (0.036)	8.578 (0.0137)
H	---	Not tested*	19.635 (<0.0001)	6.183 (0.0129)	---	---
I	0.0057 (0.940)	Not tested*	240.997 (<0.0001)	2.207 (0.0073)	43.947 (<0.0001)	---

40/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

- Odds Ratio Estimates and 95% Wald Confidence Intervals: Universities A to E

	A	*B*	*C*	*D*	*E*
GENDER (Female vs. Male)	0.872 [0.766, 0.993]	—	—	—	1.507 [1.236, 1.837]
HSGPA	Not tested*	4.469 [3.202, 6.238]	2.719 [2.011, 3.677]	3.657 [3.306, 4.045]	3.423 [2.645, 4.430]
SATQ	Not tested*	1.006 [1.004, 1.009]	1.005 [1.004, 1.005]	1.005 [1.004, 1.005]	1.003 [1.002, 1.005]
SATV	Not tested*	0.998 [0.995, 1.000]	0.998 [0.996, 1.000]	0.998 [0.998, 0.999]	—
ETHNIC (Overall)	Not available**	—	Not available**	Not available**	Not available**
ETHNIC (Hispanic vs. Other)	—	—	—	2.744 [1.311, 5.743]	—
CITIZEN (Citizen vs. NRAlien)	—	—	—	—	—
CITIZEN (NRAlien vs. ResAlien)	—	—	—	1.827 [1.328, 2.513]	—

41/98

Not tested because of missing field in database.
** Significant variable overall, but odds ratio can only be computed between specific variable levels.
--- Indicates variable was not found to be significant.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

- Odds Ratio Estimates and 95% Wald Confidence Intervals: Universities F to I

	F	*G*	*H*	*I*
GENDER (Female vs. Male)	0.835 [0.752, 0.927]	0.554 [0.442, 0.695]	---	---
HSGPA	2.118 [1.892, 2.372]	4.103 [3.366, 5.001]	Not tested*	Not tested*
SATQ	1.003 [1.003, 1.004]	1.006 [1.005, 1.007]	1.008 [1.004, 1.011]	1.006 [1.006, 1.007]
SATV	0.998 [0.997, 0.998]	0.998 [0.996, 0.999]	0.997 [0.994, 0.999]	0.999 [0.998, 1.000]
ETHNIC (Overall)	Not available**	Not available**	---	Not available**
ETHNIC (Black vs. Hispanic)	0.542 [0.344, 0.853]	---	---	---
ETHNIC (Black vs. International)	---	---	---	0.355 [0.228, 0.553]
CITIZEN (Citizen vs. NRAlien)	1.975 [1.124, 3.472]	---	---	---
CITIZEN (NRAlien vs. ResAlien)	---	---	---	---

* Not tested because of missing field in database.
** Significant variable overall, but odds ratio can only be computed between specific variable levels.
--- Indicates variable was not found to be significant.

Applied Research Study Example #1
Identifying Impact Factors

Findings

- **High school GPA** and **SAT Math** scores were positively correlated with graduation rates for all universities.

Applied Research Study Example #1
Identifying Impact Factors

Findings

- **High school GPA** and **SAT Math** scores were positively correlated with graduation rates for all universities.

- **SAT Verbal** scores correlated negatively with odds of graduation—for 7 out of 8 universities, making this observation even more notable.

43/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Findings

- **High school GPA** and **SAT Math** scores were positively correlated with graduation rates for all universities.
- **SAT Verbal** scores correlated negatively with odds of graduation—for 7 out of 8 universities, making this observation even more notable.
- While gender, ethnicity and citizenship also showed significant effects, these were not consistently positive or negative. In 3 universities, the graduation rate for males was higher than females, while in 1, it was higher for females.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1
Identifying Impact Factors

Findings

- **High school GPA** and **SAT Math** scores were positively correlated with graduation rates for all universities.
- **SAT Verbal** scores correlated negatively with odds of graduation—for 7 out of 8 universities, making this observation even more notable.
- While gender, ethnicity and citizenship also showed significant effects, these were not consistently positive or negative. In 3 universities, the graduation rate for males was higher than females, while in 1, it was higher for females.
- Ethnicity was significant in 7 universities.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #1

Identifying Impact Factors

Findings

- **High school GPA** and **SAT Math** scores were positively correlated with graduation rates for all universities.
- **SAT Verbal** scores correlated negatively with odds of graduation—for 7 out of 8 universities, making this observation even more notable.
- While gender, ethnicity and citizenship also showed significant effects, these were not consistently positive or negative. In 3 universities, the graduation rate for males was higher than females, while in 1, it was higher for females.
- Ethnicity was significant in 7 universities.
- In 2 universities, citizenship significantly affected graduation.

43/98

EAST CAROLINA UNIVERSITY

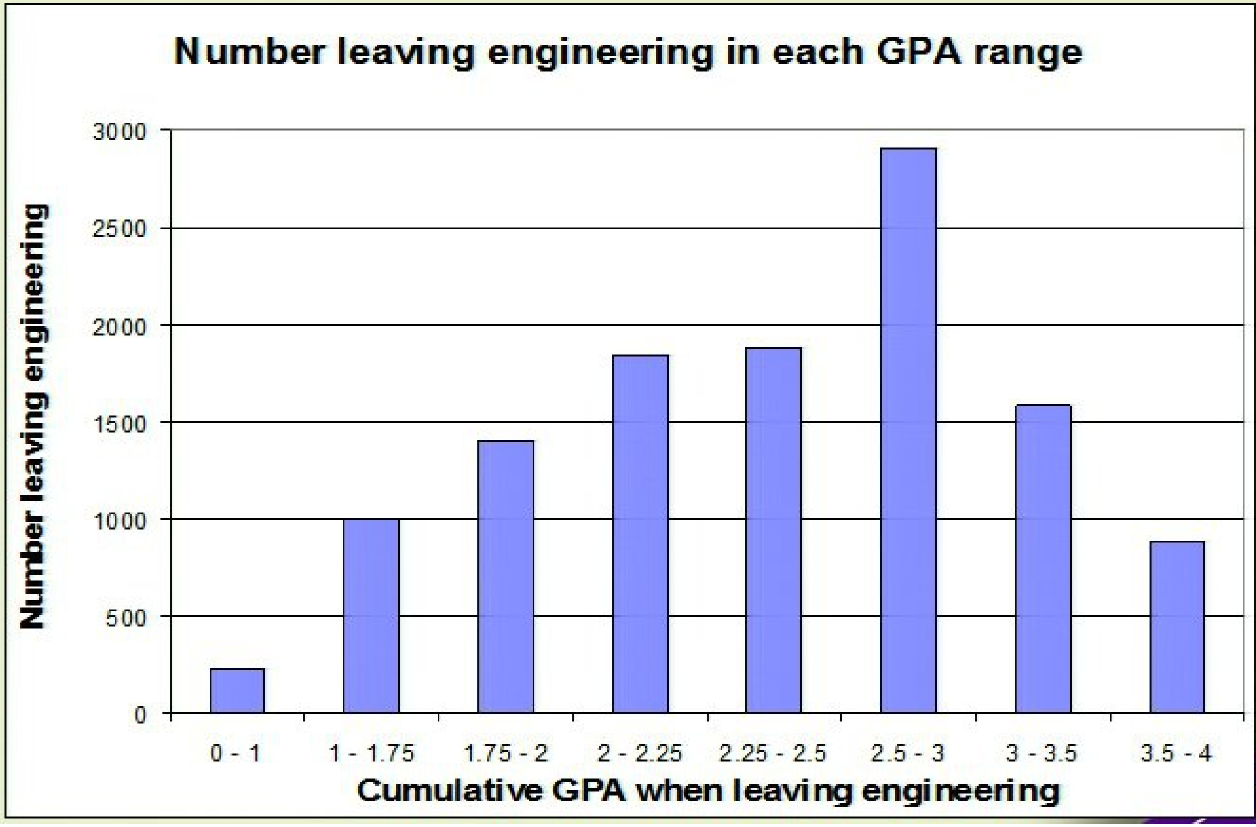
Applied Research Study Example #2

Students Leaving Engineering

Grade-Point Average, changes of major, and majors selected by students leaving engineering.

Received *Frontiers in Education (FIE)* ***Benjamin J. Dasher Best Paper Award.***

A significant majority of students leave engineering with passing GPAs, and that the distribution is actually weighted toward higher GPAs

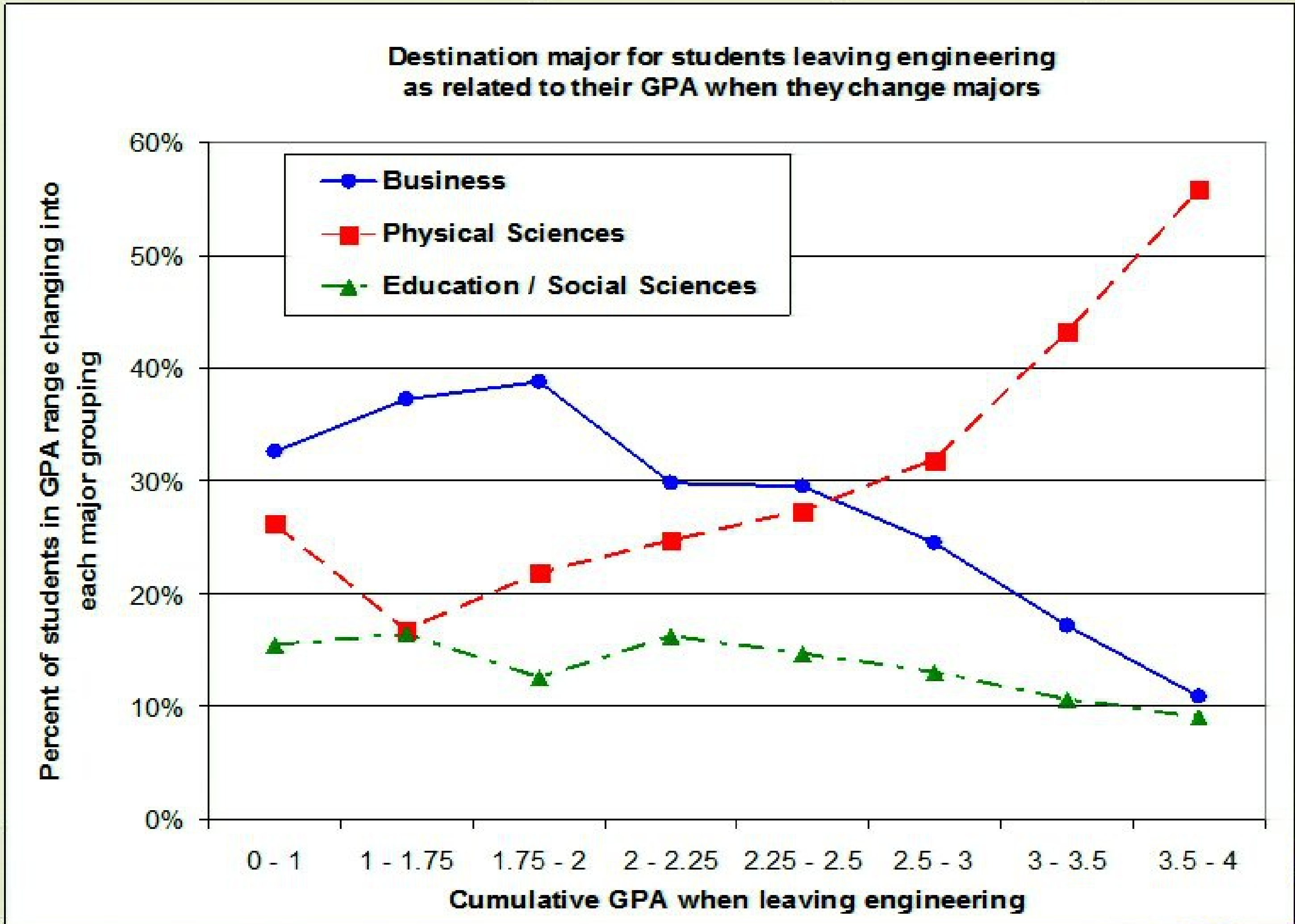


EAST CAROLINA UNIVERSITY

Applied Research Study Example #2

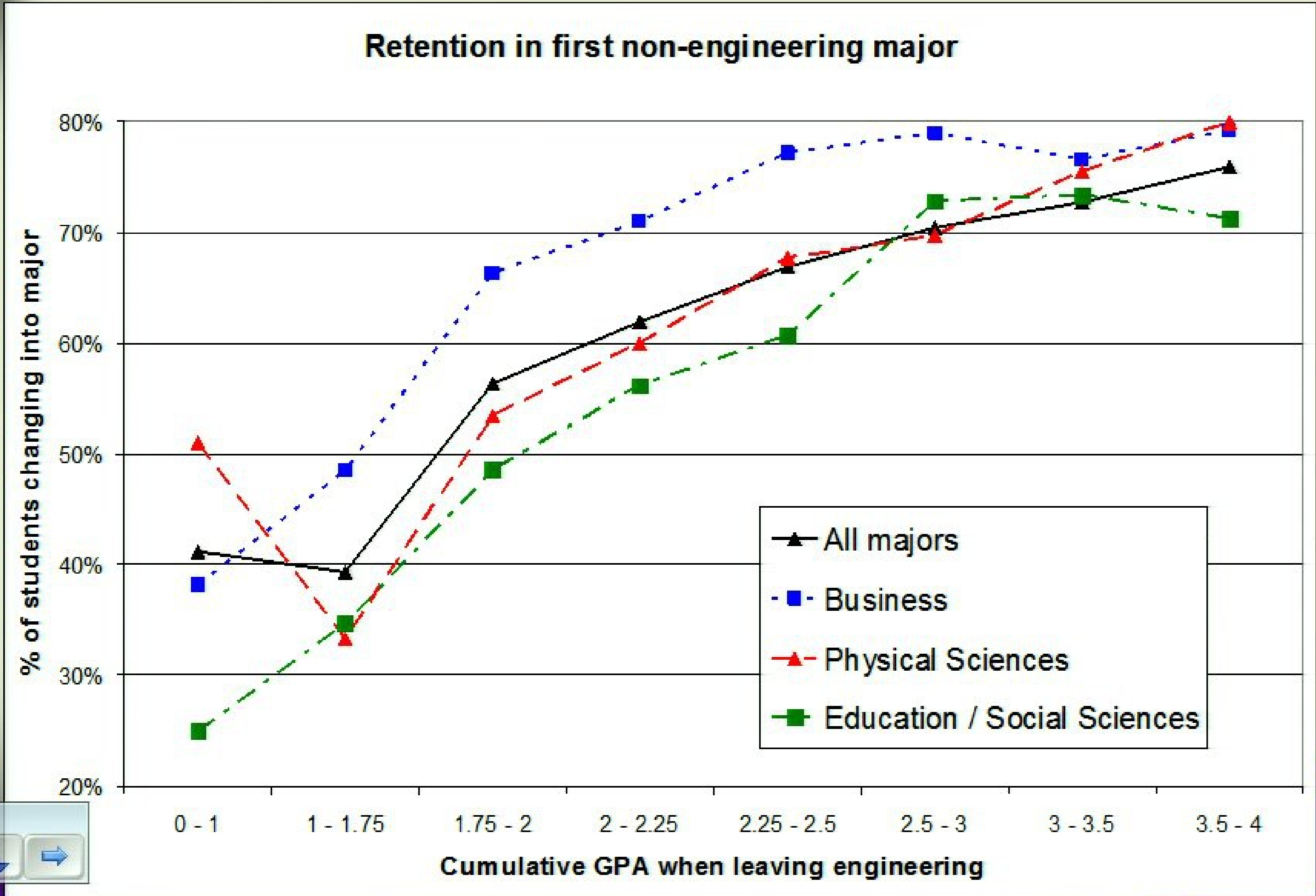
Students Leaving Engineering

Where do students go when they leave engineering?



Applied Research Study Example #2

Students Leaving Engineering

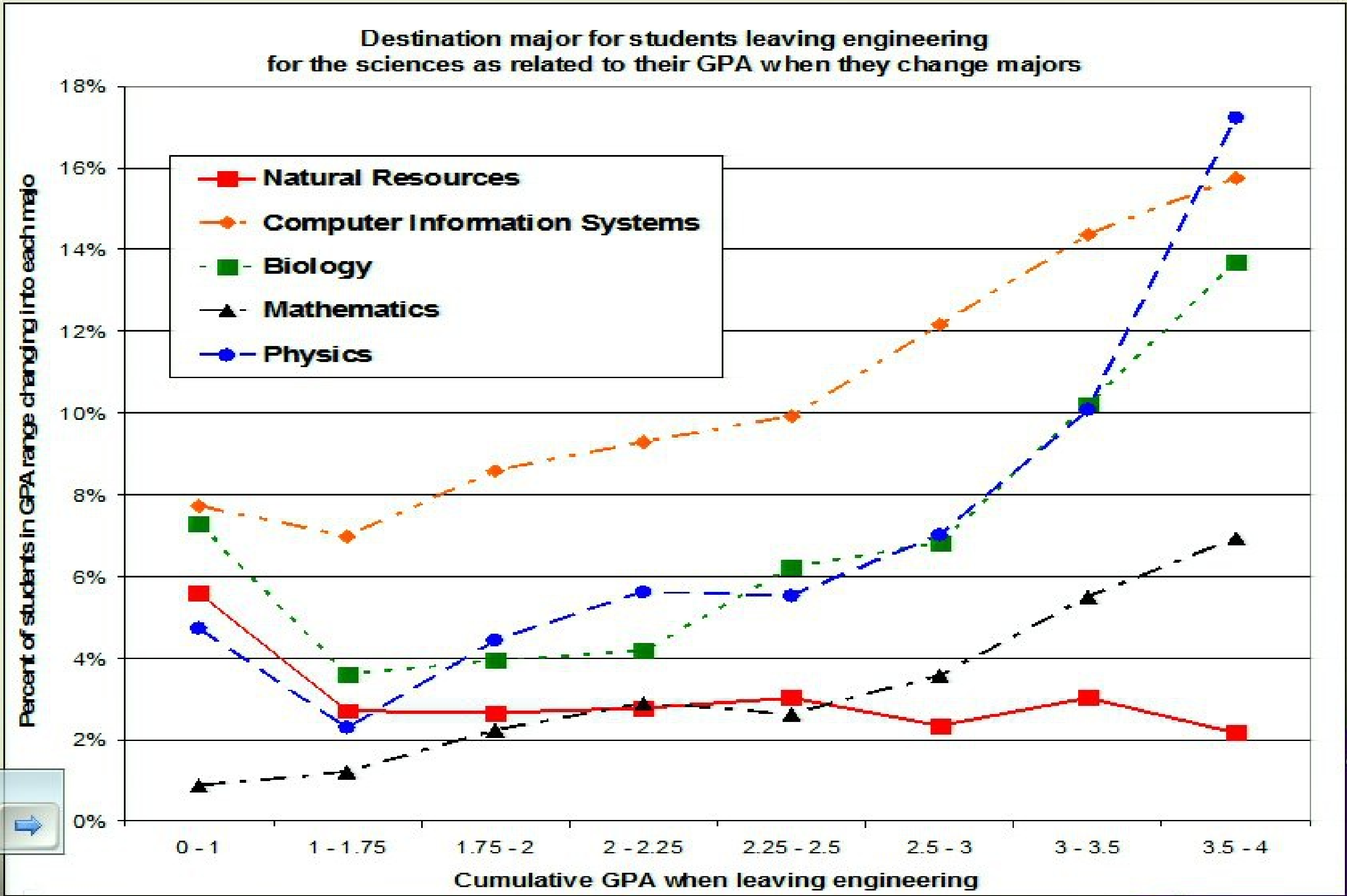


EAST CAROLINA UNIVERSITY

Applied Research Study Example #2

Students Leaving Engineering

Destination Major for Students leaving engineering for the Science



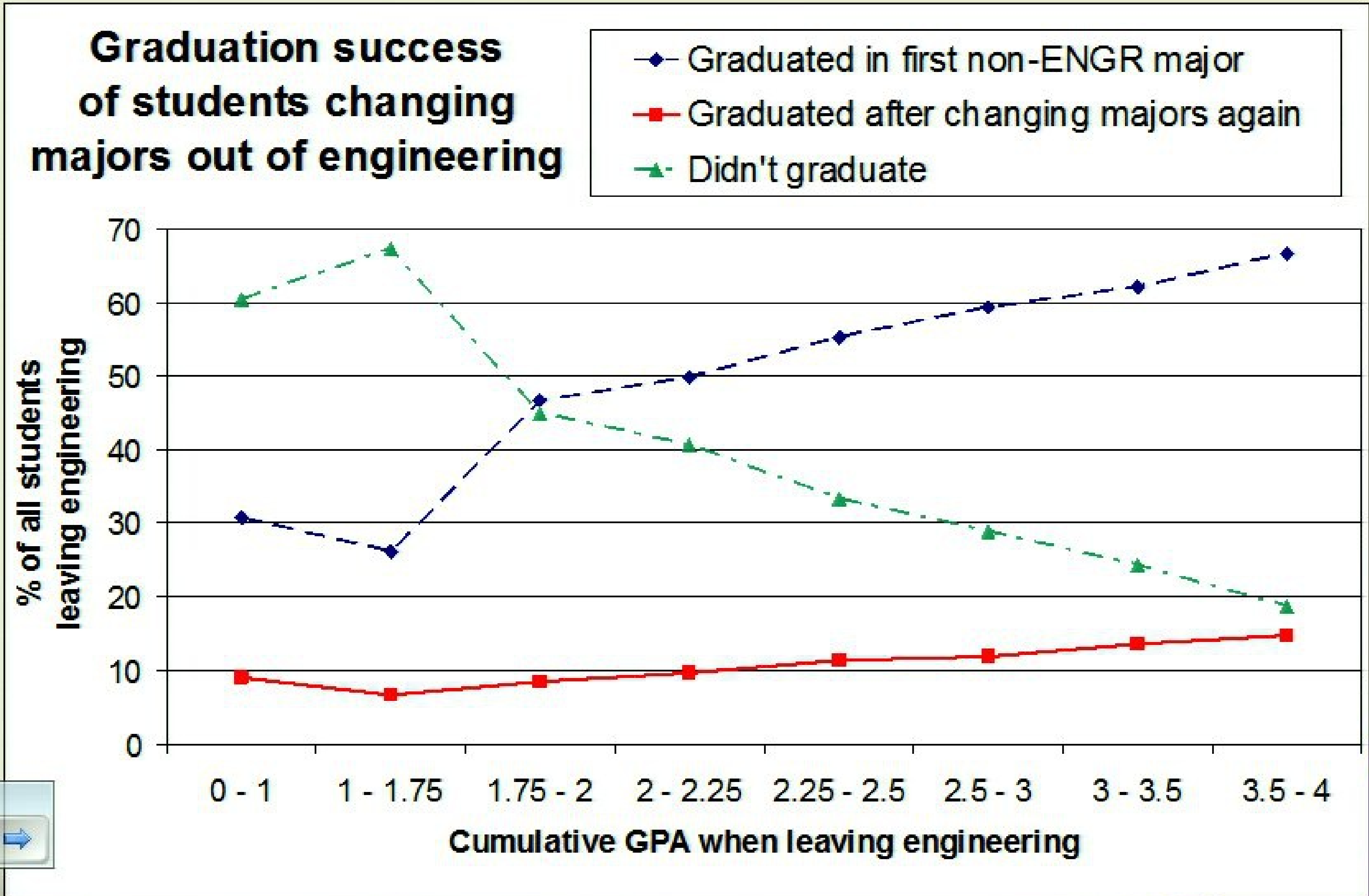
47/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #2

Students Leaving Engineering

What happens to students who leave engineering?



Applied Research Study Example #2

Students Leaving Engineering

49/98

Applied Research Study Example #2

Students Leaving Engineering

Findings and Discussion

- Students do not just leave engineering because they "flunk out" – in fact, many are doing quite well academically.
  - Why do they leave?
  - What can we do to help students choose the right major?

Applied Research Study Example #2

Students Leaving Engineering

Findings and Discussion

- Students do not just leave engineering because they "flunk out" – in fact, many are doing quite well academically.
  - Why do they leave?
  - What can we do to help students choose the right major?
- Students with lower GPAs were more likely to change into business majors; those with higher GPAs were more likely to change into science majors.
  - Does this indicate some stereotyped perceptions about different majors?
  - What can we do to make sure students get the facts, not the stereotypes?

49/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #2

Students Leaving Engineering

Findings and Discussion

- Students do not just leave engineering because they "flunk out" – in fact, many are doing quite well academically.
  - Why do they leave?
  - What can we do to help students choose the right major?
- Students with lower GPAs were more likely to change into business majors; those with higher GPAs were more likely to change into science majors.
  - Does this indicate some stereotyped perceptions about different majors?
  - What can we do to make sure students get the facts, not the stereotypes?
- The fraction of students changing into education and social science majors was 10-15% across all levels of GPA.
  - This may indicate that students change into these majors because of a "calling."
  - Can we make sure students see the social relevance of engineering?

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

How do chemical engineering students differ from others?
Received ***American Society for Engineering Education Best Paper Award.***

Using the SUCCEED longitudinal database (LDB), undergraduate chemical engineering students were compared with other engineering and non-engineering students on demographics and academic performance measures.

Applied Research Study Example #3

How do chemical engineering students differ

Difference in gender among Chemical Engineering, Other Engineering, Science, and Non-science undergraduates

GENDER \ GROUP	*CHE (%)*	*OENG (%)*	*SCI (%)*	*NSCI (%)*
Female	36.25	20.49	44.47	56.97
Male	63.75	79.51	55.53	43.03

Applied Research Study Example #3

How do chemical engineering students differ

Percentage of Student Flow between Engineering Subfields (n=22,610)

BEGIN / *GRAD*	*Chemical*	*Civil*	*Computer*	*Electrical*	*Industrial*	*Mechanical/ Aerospace*	*Other Engineering*
Chemical	59.8	0.9	0.2	0.9	0.5	1.0	1.5
Civil	2.8	72.2	1.1	2.6	2.5	4.1	3.7
Computer	0.3	0.2	70.1	1.9	0.2	0.5	0.6
Electrical	1.5	1.3	10.3	66.1	1.1	2.7	1.7
Industrial	4.8	2.5	2.9	4.4	77.7	4.3	2.9
Mechanical/ Aerospace	2.6	2.9	1.1	3.7	2.6	63.5	4.5
Other Engineering	3.6	1.9	0.3	2.1	1.1	3.3	63.3
Total	75.4	82.0	86.1	81.7	85.6	79.4	78.2

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Percentage of Engineering Student Flow to Non-engineering fields (n=22,610)

GRAD \ *BEGIN*	*Chemical*	*Civil*	*Computer*	*Electrical*	*Industrial*	*Mechanical/ Aerospace*	*Other Engineering*
Biology	3.8	0.8	0.7	1.1	0.6	1.1	1.6
Business	3.9	4.9	3.0	4.5	7.7	5.5	4.7
CIS	0.8	0.4	5.5	2.6	0.6	1.3	1.3
Physical Science	7.4	0.9	0.6	0.6	0.5	1.3	2.0
Social Science	1.6	1.9	1.5	1.8	1.1	2.7	1.7
Other Non-Engineering	7.1	9.1	2.6	7.6	4.0	8.7	10.5
Total	24.6	18.0	13.9	18.3	14.4	20.6	21.8

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

54/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Analyses on the nine variables

Applied Research Study Example #3

How do chemical engineering students differ

Analyses on the nine variables

- **Multivariate Omnibus Test** was done first to determine whether there is an overall group effect. The null hypothesis is that the groups do not differ on any of the 9 variables. The test was done using SAS PROC GLM through a MANOVA with "group" as the between-subjects factor and the 9 academic characteristics as dependent variables. The test showed that the groups differ on at least one variable.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Analyses on the nine variables

- **Multivariate Omnibus Test** was done first to determine whether there is an overall group effect. The null hypothesis is that the groups do not differ on any of the 9 variables. The test was done using SAS PROC GLM through a MANOVA with "group" as the between-subjects factor and the 9 academic characteristics as dependent variables. The test showed that the groups differ on at least one variable.

- **ANOVA on Individual Variables**. We then wanted to find out which of the variables is significant. This was done with an ANOVA on the variable. For a given variable, the null hypothesis is that the groups do not differ on that particular variable.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Analyses on the nine variables

• **Multivariate Omnibus Test** was done first to determine whether there is an overall group effect. The null hypothesis is that the groups do not differ on any of the 9 variables. The test was done using SAS PROC GLM through a MANOVA with "group" as the between-subjects factor and the 9 academic characteristics as dependent variables. The test showed that the groups differ on at least one variable.

• **ANOVA on Individual Variables**. We then wanted to find out which of the variables is significant. This was done with an ANOVA on the variable. For a given variable, the null hypothesis is that the groups do not differ on that particular variable.

• **Pair-wise Comparison among Groups.** To determine the extent of the variation among the groups, each variable was subsequently tested using pair-wise comparison. The Shaffer-Holm procedure was used to control the family-wise error rate.

54/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Comparisons among Chemical Engineering, Other Engineering, Science and Non-science Majors

Variable	Statistic	CHE	OENG	SCI	NSCI
*SAT Math Score	M	645.92 [a]	635.88 [b]	609.11 [c]	553.16 [d]
	SD	82.29	83.83	96.20	92.20
*SAT Verbal Score	M	533.96 [a]	517.34 [b]	519.66 [b]	490.06 [c]
	SD	86.73	87.62	95.66	87.45
**High School GPA	M	3.72 [a]	3.56 [b]	3.57 [b]	3.31 [c]
	SD	0.36	0.41	0.47	0.52
Time to Graduation	M	53.88 [a]	55.90 [b]	51.77 [c]	51.82 [c]
	SD	8.88	10.24	11.22	10.88
Cumulative GPA	M	3.17 [a]	2.98 [b]	3.04 [c]	2.97 [d]
	SD	.49	.55	.68	.54
Number of Major Changes	M	.63 [a]	.95 [b]	.89 [c]	1.29 [d]
	SD	.76	.83	.92	1.03
Semesters to Graduation	M	13.04 [a]	13.87 [b]	11.69 [c]	10.94 [d]
	SD	4.07	4.35	3.62	2.62
Cumulative Semester Hours	M	163.35 [a]	168.87 [b]	146.05 [c]	136.11 [d]
	SD	38.42	42.12	36.29	26.10
Average Semester Hours	M	13.01 [a]	12.60 [b]	12.84 [c]	12.68 [d]
	SD	2.34	2.22	1.99	1.75

55/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

  CHE majors had much **better SAT verbal** scores than NSCI majors, as well as OENG and SCI majors, although the scores were closer.

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

- CHE majors had much **better SAT verbal** scores than NSCI majors, as well as OENG and SCI majors, although the scores were closer.

- CHE majors came in with **better high school GPAs** than all other groups.

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

- CHE majors had much **better SAT verbal** scores than NSCI majors, as well as OENG and SCI majors, although the scores were closer.

- CHE majors came in with **better high school GPAs** than all other groups.

- CHE students took significantly more **time to graduate** than both SCI and NSCI groups, but required less time to graduate than OENG students.

Applied Research Study Example #3

How do chemical engineering students differ

Findings

- Chemical engineering students (CHE) majors had significantly **better SAT math** scores than other engineering majors (OENG), science majors (SCI) and non-science majors (NSCI).

- CHE majors had much **better SAT verbal** scores than NSCI majors, as well as OENG and SCI majors, although the scores were closer.

- CHE majors came in with **better high school GPAs** than all other groups.

- CHE students took significantly more **time to graduate** than both SCI and NSCI groups, but required less time to graduate than OENG students.

- CHE majors have significantly **higher cumulative GPAs** than others.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

- CHE majors **changed major significantly fewer times** than other majors.

Applied Research Study Example #3

How do chemical engineering students differ

- CHE majors **changed major significantly fewer times** than other majors.
- CHE students took an average of **two more semesters to graduate** than their SCI and NSCI counterparts, but required about a semester less time to graduate than OENG students.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

- CHE majors **changed major significantly fewer times** than other majors.
- CHE students took an average of **two more semesters to graduate** than their SCI and NSCI counterparts, but required about a semester less time to graduate than OENG students.
- CHE students took significantly **more semester hours to graduate** than both SCI and NSCI students, but required less semester hours to graduate than OENG students.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #3

How do chemical engineering students differ

- CHE majors **changed major significantly fewer times** than other majors.
- CHE students took an average of **two more semesters to graduate** than their SCI and NSCI counterparts, but required about a semester less time to graduate than OENG students.
- CHE students took significantly **more semester hours to graduate** than both SCI and NSCI students, but required less semester hours to graduate than OENG students.
- CHE students took slightly **fewer hours each semester** than OSCI, SCI and NSCI students.

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

58/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Nonparametric Survival Analysis of the loss rate of undergraduate engineering students.

Supported by NSF's MIDFIELD grant (Multiple-Institution Database For Investigating Engineering Longitudinal Development).

Published in *Journal of Engineering Education*.

58/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Nonparametric Survival Analysis of the loss rate of undergraduate engineering students.

Supported by NSF's MIDFIELD grant (Multiple-Institution Database For Investigating Engineering Longitudinal Development).

Published in *Journal of Engineering Education*.

Background:

Student dropout from colleges and universities has long been a concern and focus for educators.

58/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Nonparametric Survival Analysis of the loss rate of undergraduate engineering students.

Supported by NSF's MIDFIELD grant (Multiple-Institution Database For Investigating Engineering Longitudinal Development).

Published in *Journal of Engineering Education.*

Background:

Student dropout from colleges and universities has long been a concern and focus for educators.

- The snapshot approach taken by the existing studies does not offer the means to fully understand the time-dependent dropout issue.

58/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Nonparametric Survival Analysis of the loss rate of undergraduate engineering students.

Supported by NSF's MIDFIELD grant (Multiple-Institution Database For Investigating Engineering Longitudinal Development).

Published in *Journal of Engineering Education*.

Background:

Student dropout from colleges and universities has long been a concern and focus for educators.

- The snapshot approach taken by the existing studies does not offer the means to fully understand the time-dependent dropout issue.
- In this study, using a longitudinal database of 100,179 engineering students from 9 universities and spanning 19 years, undergraduate student dropout by cohort group, gender, ethnicity, and SAT verbal and math scores were investigated by using nonparametric survival analysis.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Key Research Questions

- Does the risk of leaving engineering differ among groups with different background (i.e., cohort, gender, ethnicity, SAT math score, and SAT verbal score)?
- When are students most likely to leave engineering?
- Is SAT score a good predictor of the risk of dropping out?

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

- We introduced survival analysis to engineering student dropout research. Survival analysis allows:
  - ✓ modeling patterns of occurrence,
  - ✓ comparing these patterns among groups, and
  - ✓ building statistical models of the risk of occurrence over time.
- SAS software's PROC LIFETEST function was used to perform survival analysis.
- The data for this study were extracted from the MIDFIELD database that we developed.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Table 1. The frequency of gender by cohort groups

MIDFIELD	Cohort Groups					
Gender	1987-1990	1991-1994	1995-1998	1999-2002	2003-2004	Total
Female	4,096	4,697	4,813	5,309	1,867	20,782
Male	16,756	17,541	16,937	20,088	8,075	79,397
Total	20,852	22,238	21,750	25,397	9,942	100,179

Table 2. The frequency of gender by ethnicity

MIDFIELD	Ethnicity				
Gender	White	Asian	Minority	Other	Total
Female	14,144	1,478	4,796	364	20,782
Male	61,435	5,404	10,445	2,113	79,397
Total	75,579	6,882	15,241	2,477	100,179

EAST CAROLINA UNIVERSITY

Applied Research Study Example #4

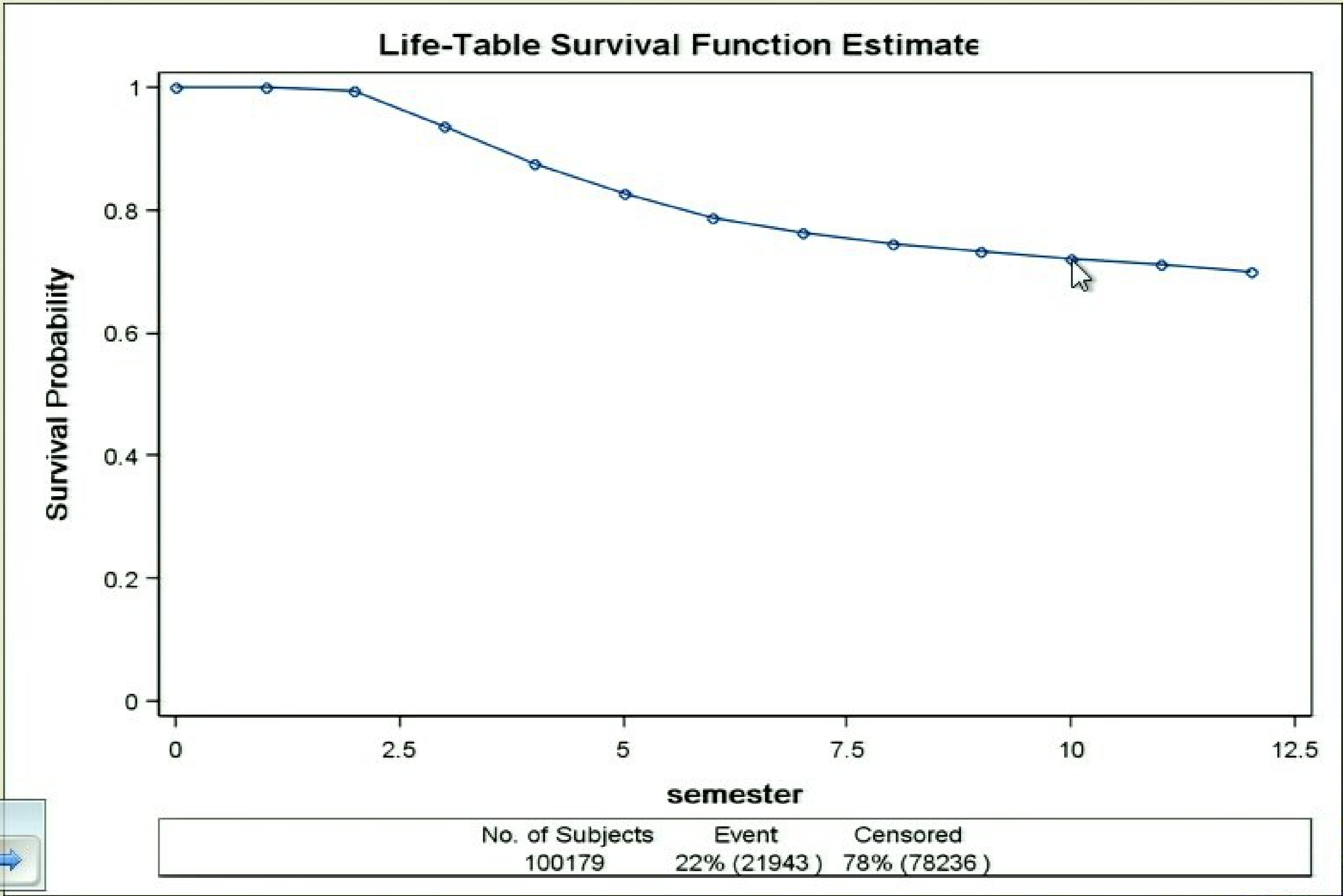
Survival Analysis of Engineering Attrition

Table 3. The frequency of SAT math score groups

SAT math score groups	Frequency	Percent	Cumulative Frequency	Cumulative Percent
SATM <500	5,430	5.42	5,430	5.42
$500 \leq$ SATM < 550	8,854	8.84	14,284	14.26
$550 \leq$ SATM < 600	15,826	15.80	30,110	30.06
$600 \leq$ SATM < 650	22,987	22.95	53,097	53.00
$650 \leq$ SATM < 700	24,036	23.99	77,133	77.00
$700 \leq$ SATM < 750	15,510	15.48	92,643	92.48
$750 \leq$ SATM ≤ 800	7,536	7.52	100,179	100.00

Results

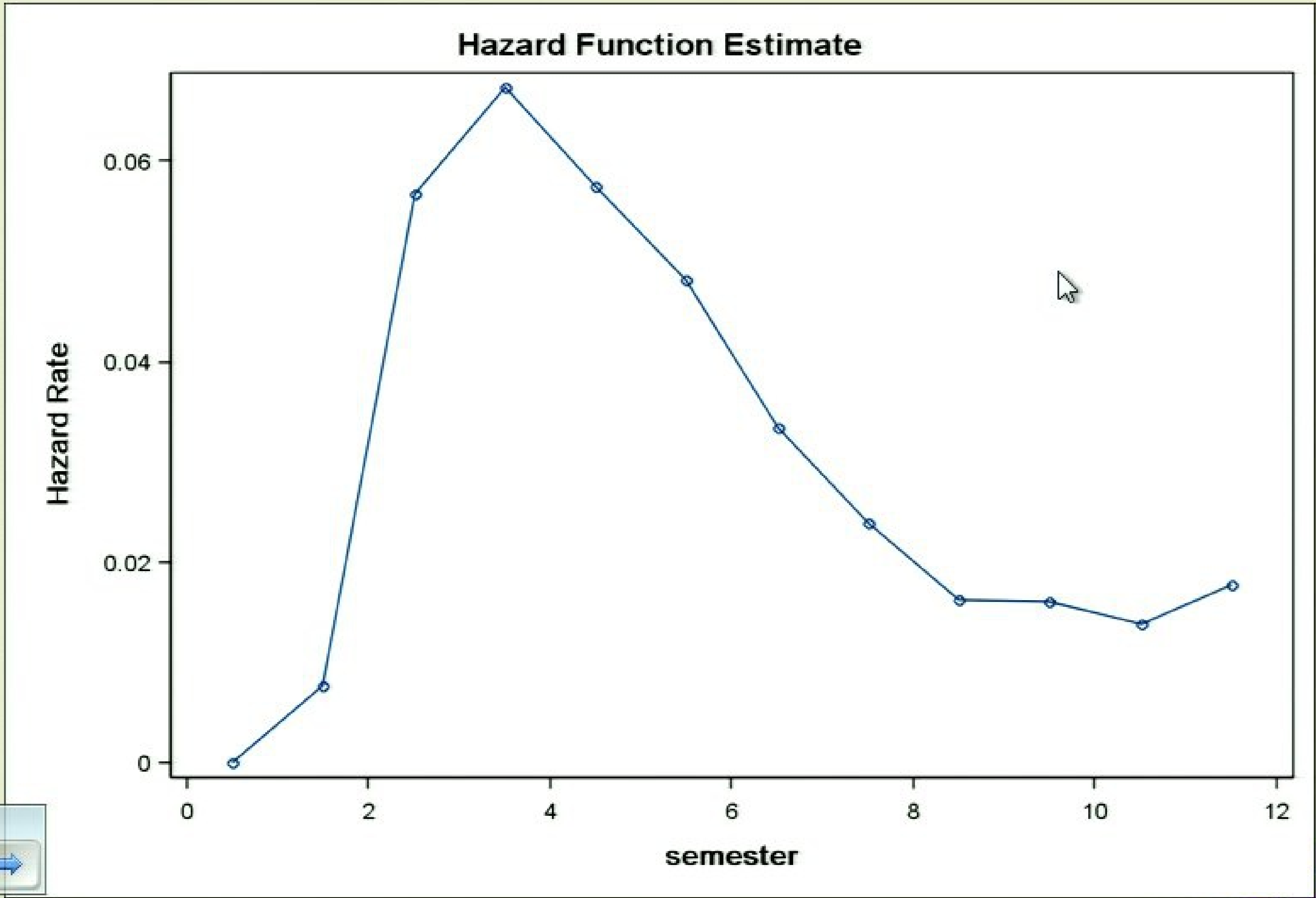
Figure 1. Survival function for first-time-in-college students matriculating in engineering.



53/98

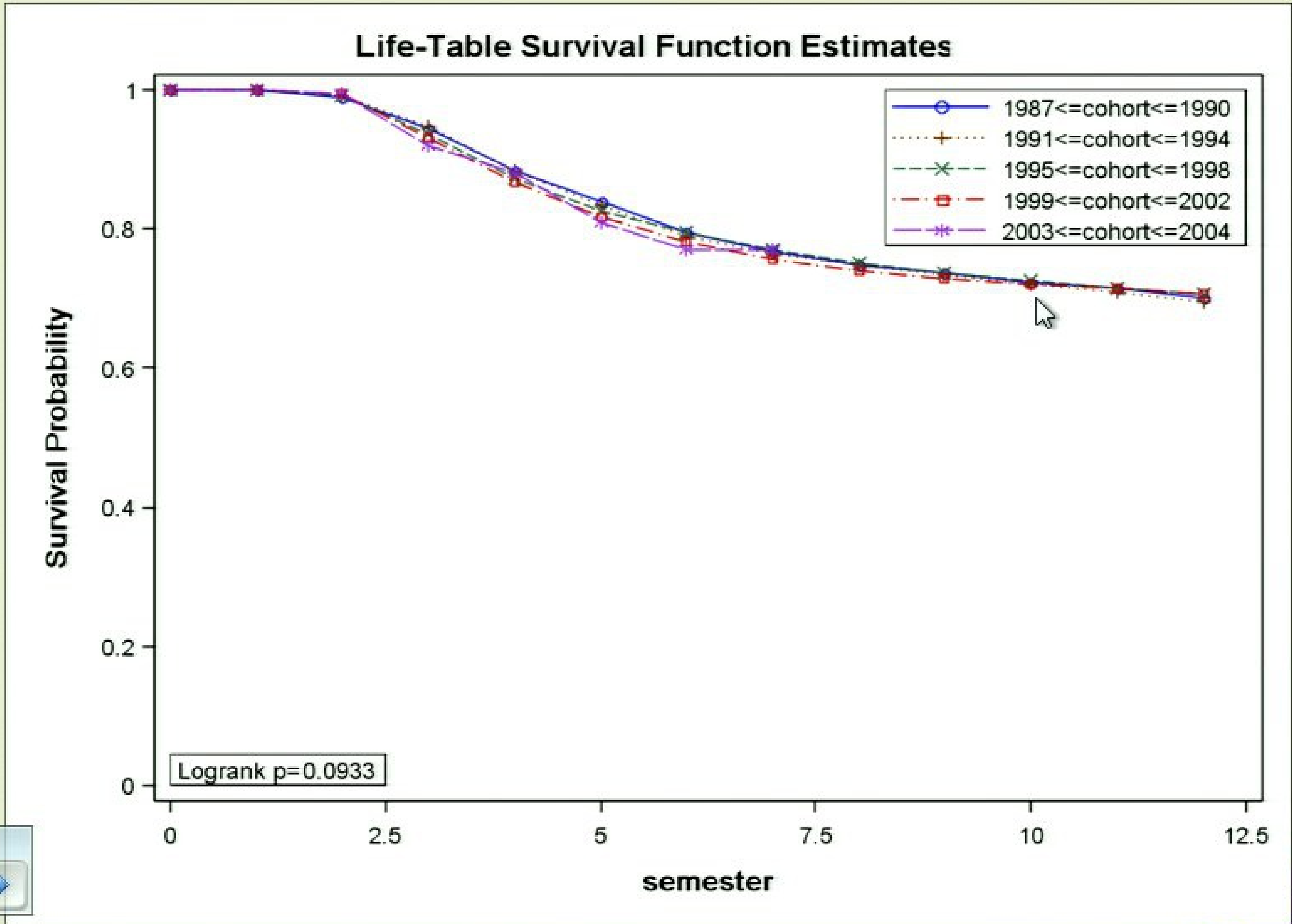
Results

Figure 2. Hazard function for first-time-in-college students matriculating in engineering.



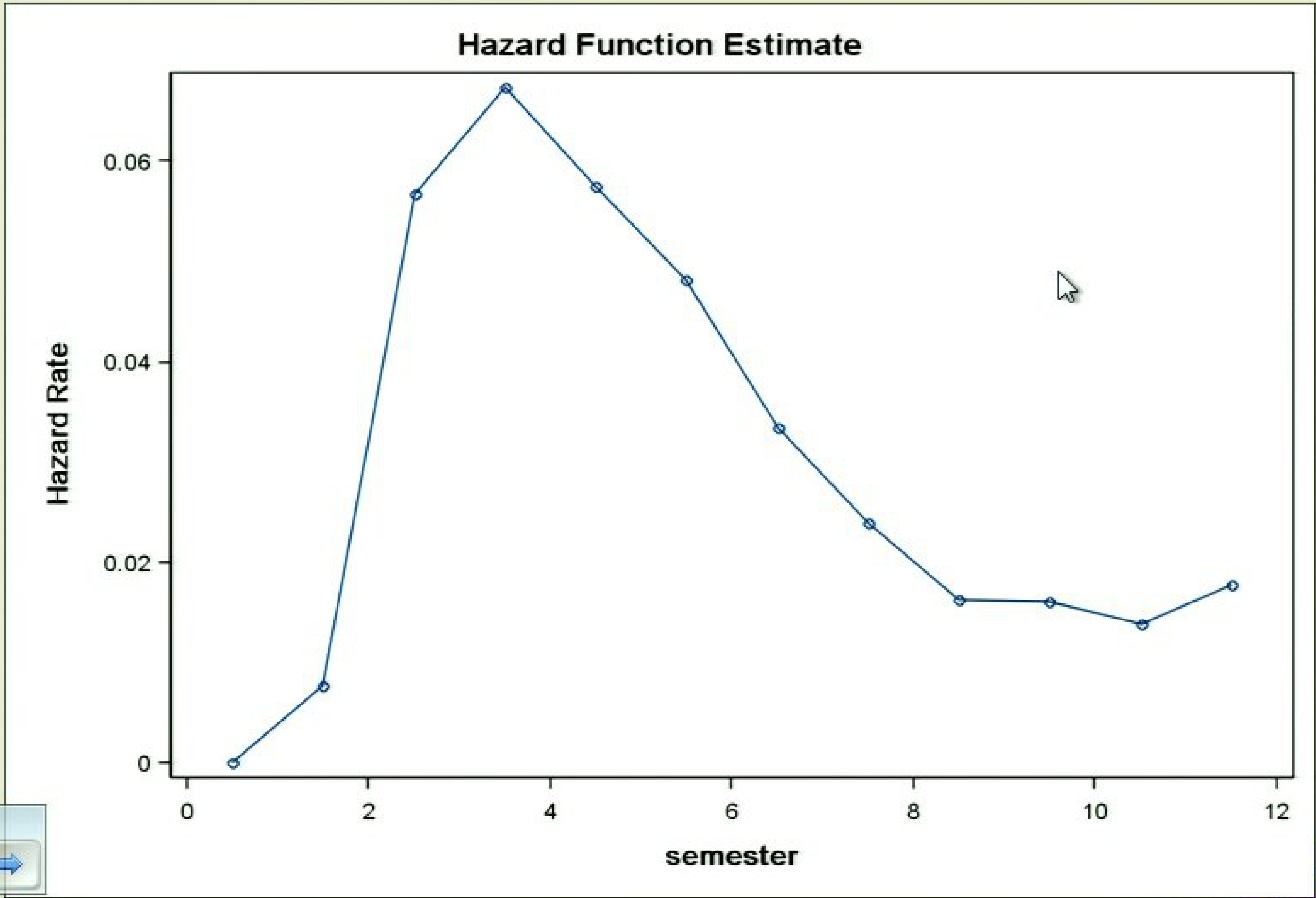
Results

Figure 3. Survival functions by cohort groups



Results

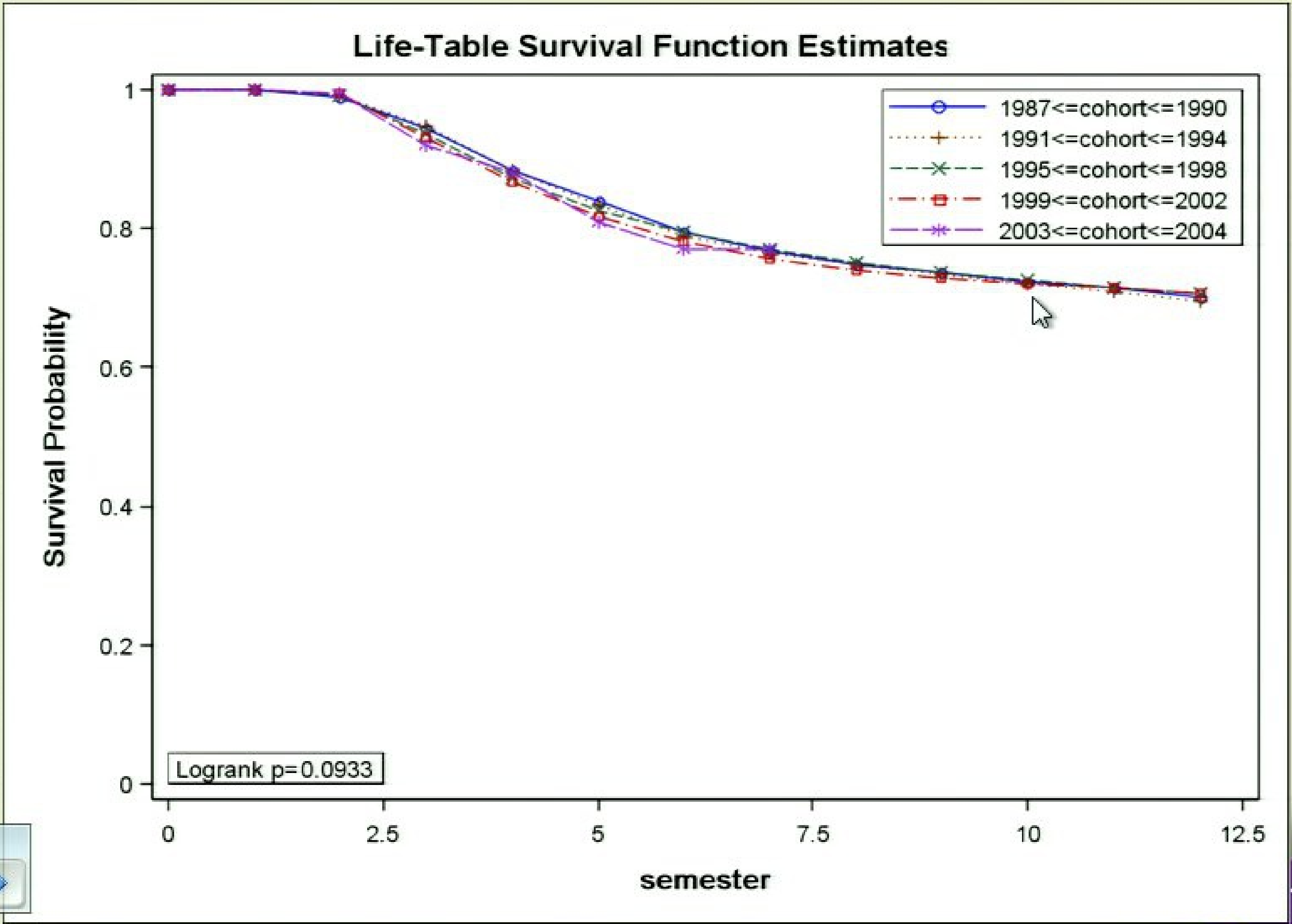
Figure 2. Hazard function for first-time-in-college students matriculating in engineering.



54/98

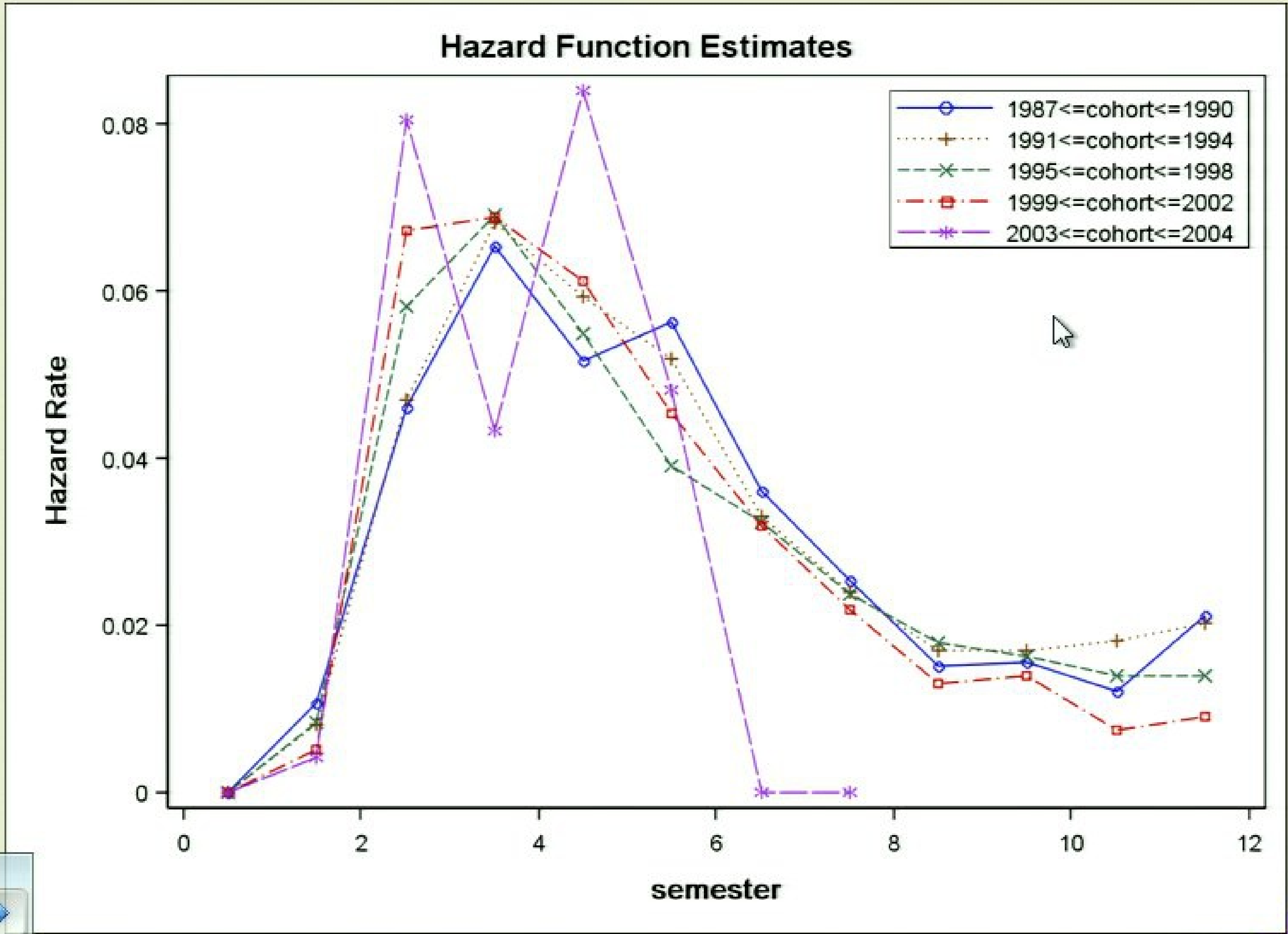
Results

Figure 3. Survival functions by cohort groups



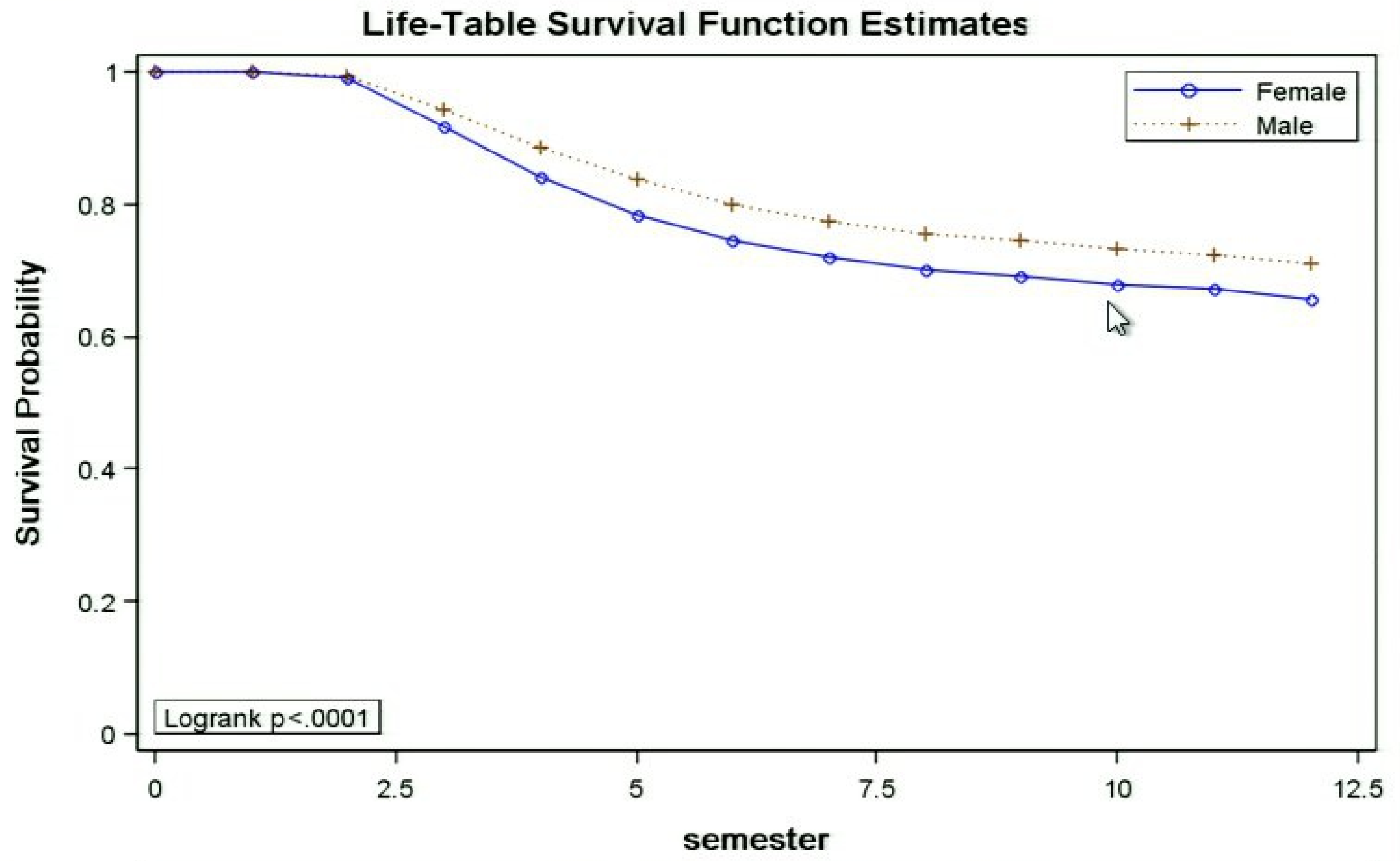
Results

Figure 4. Hazard functions by cohort groups



Results

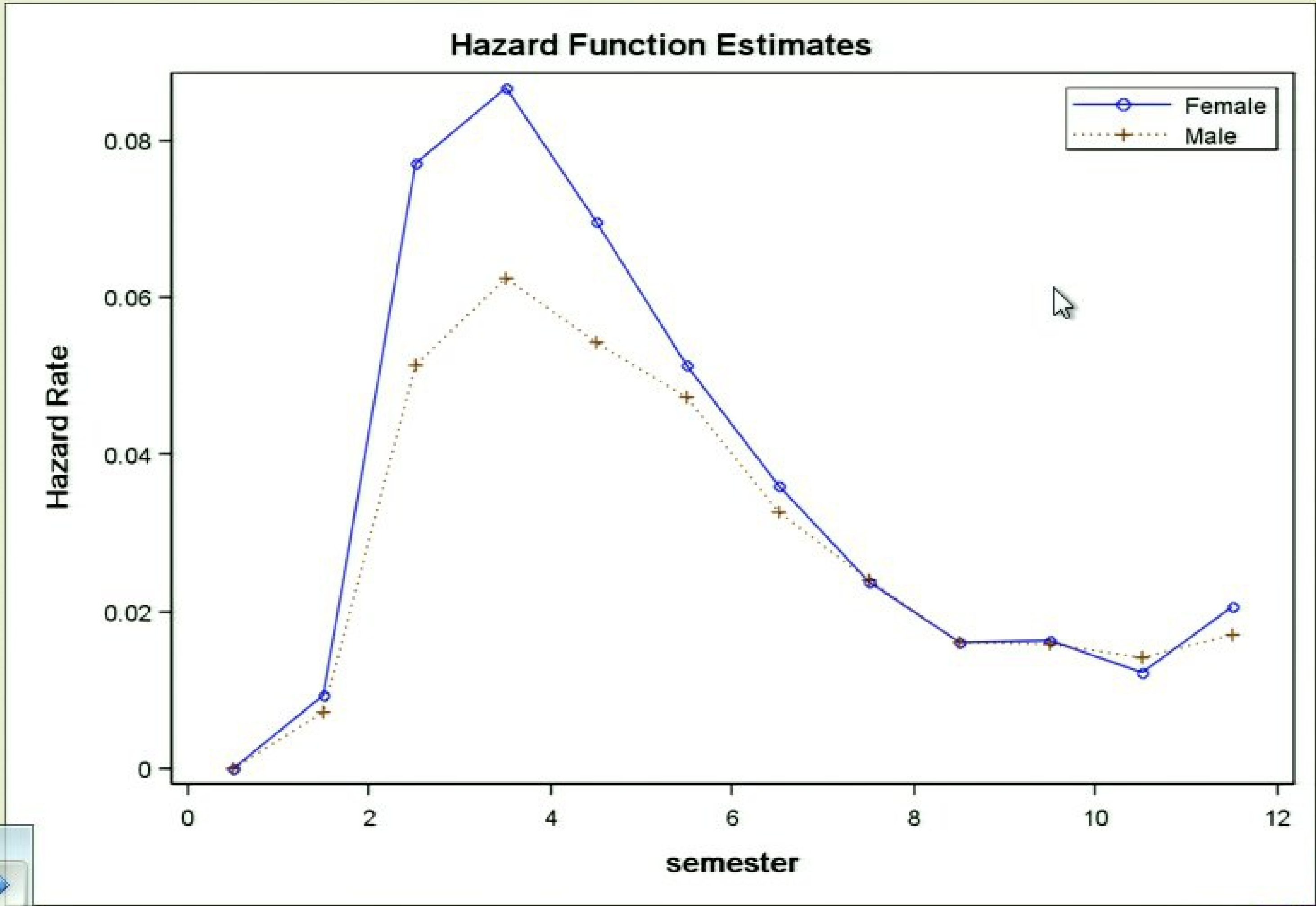
Figure 5. Survival functions by gender



	No. of Subjects	Event	Censored
Female	20782	26% (5407)	74% (15375)
Male	79397	21% (16536)	79% (62861)

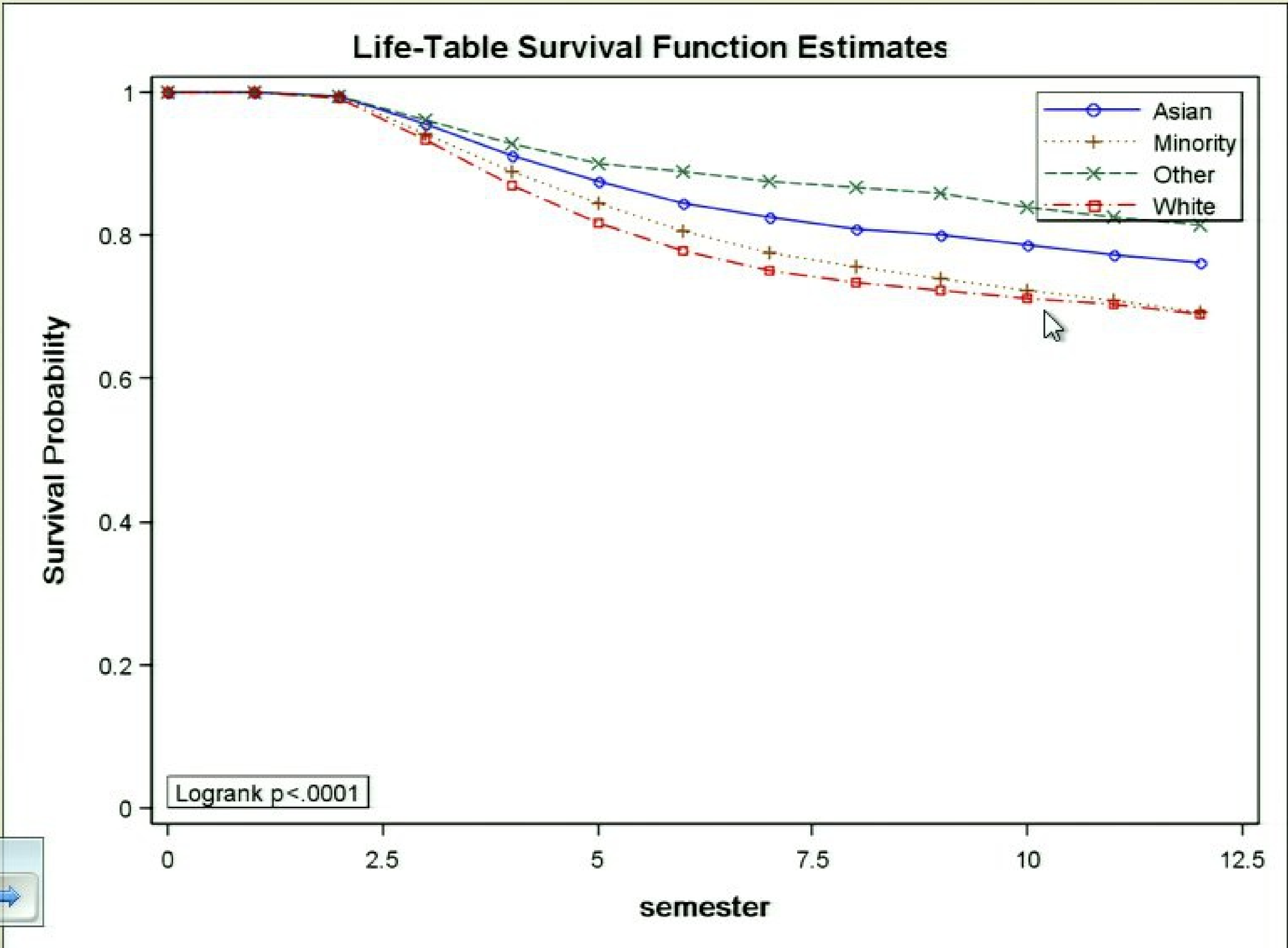
Results

Figure 6. Hazard functions by gender



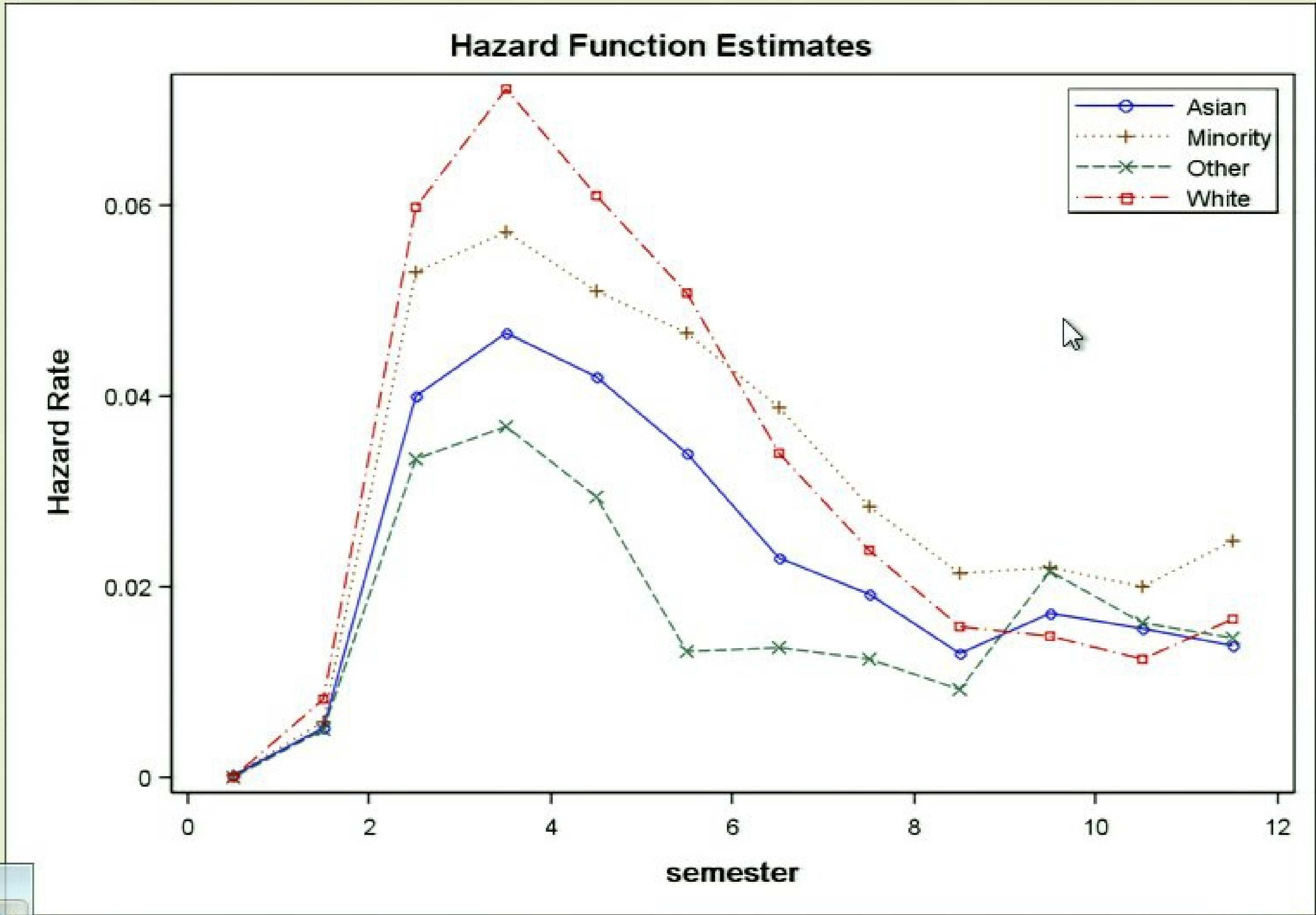
Results

Figure 7. Survival functions by ethnicity



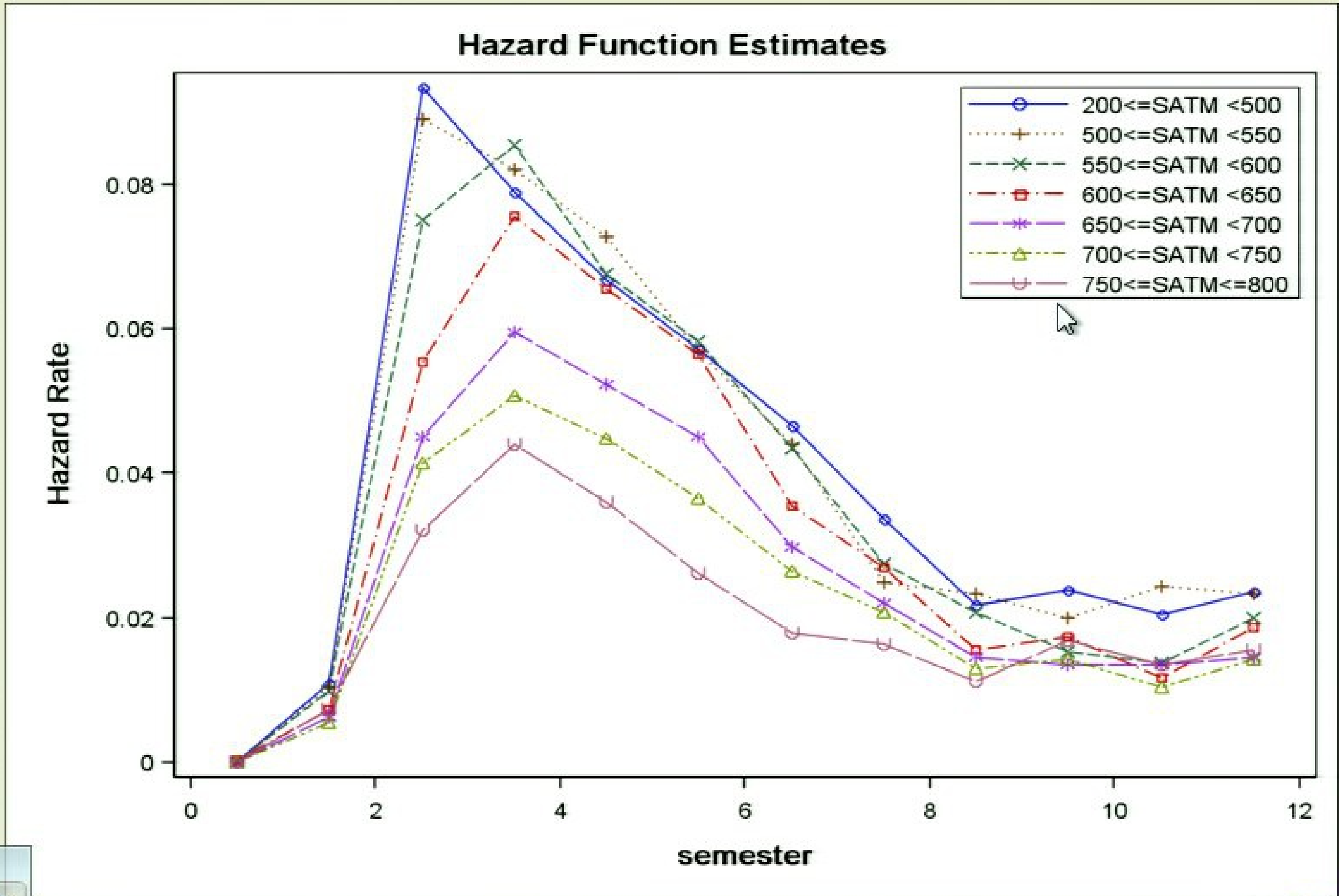
Results

Figure 8. Hazard functions by ethnicity



Results

Figure 9. Hazard functions by SAT math score groups

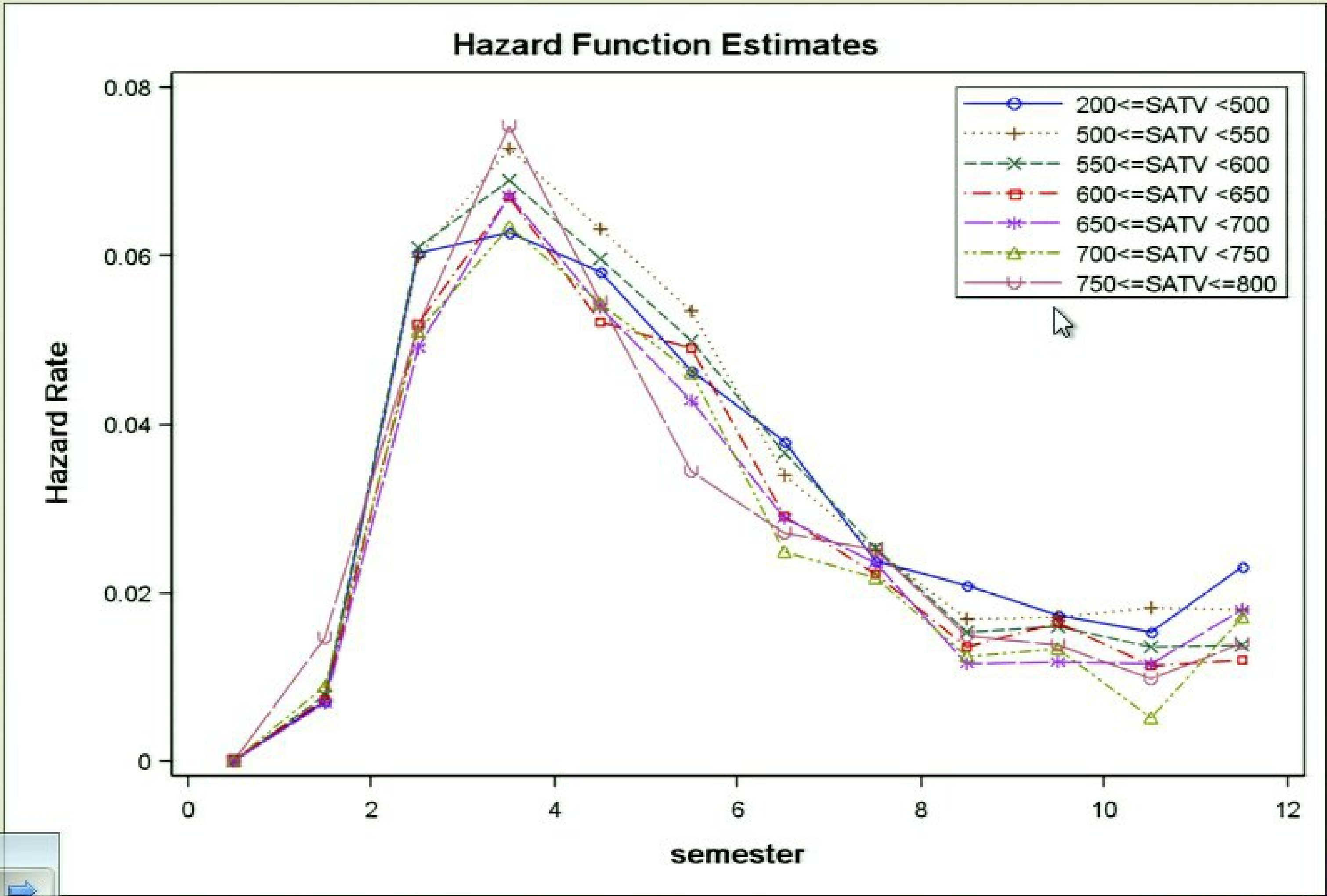


71/98

EAST CAROLINA UNIVERSITY

Results

Figure 9. Hazard functions by SAT Verbal score groups

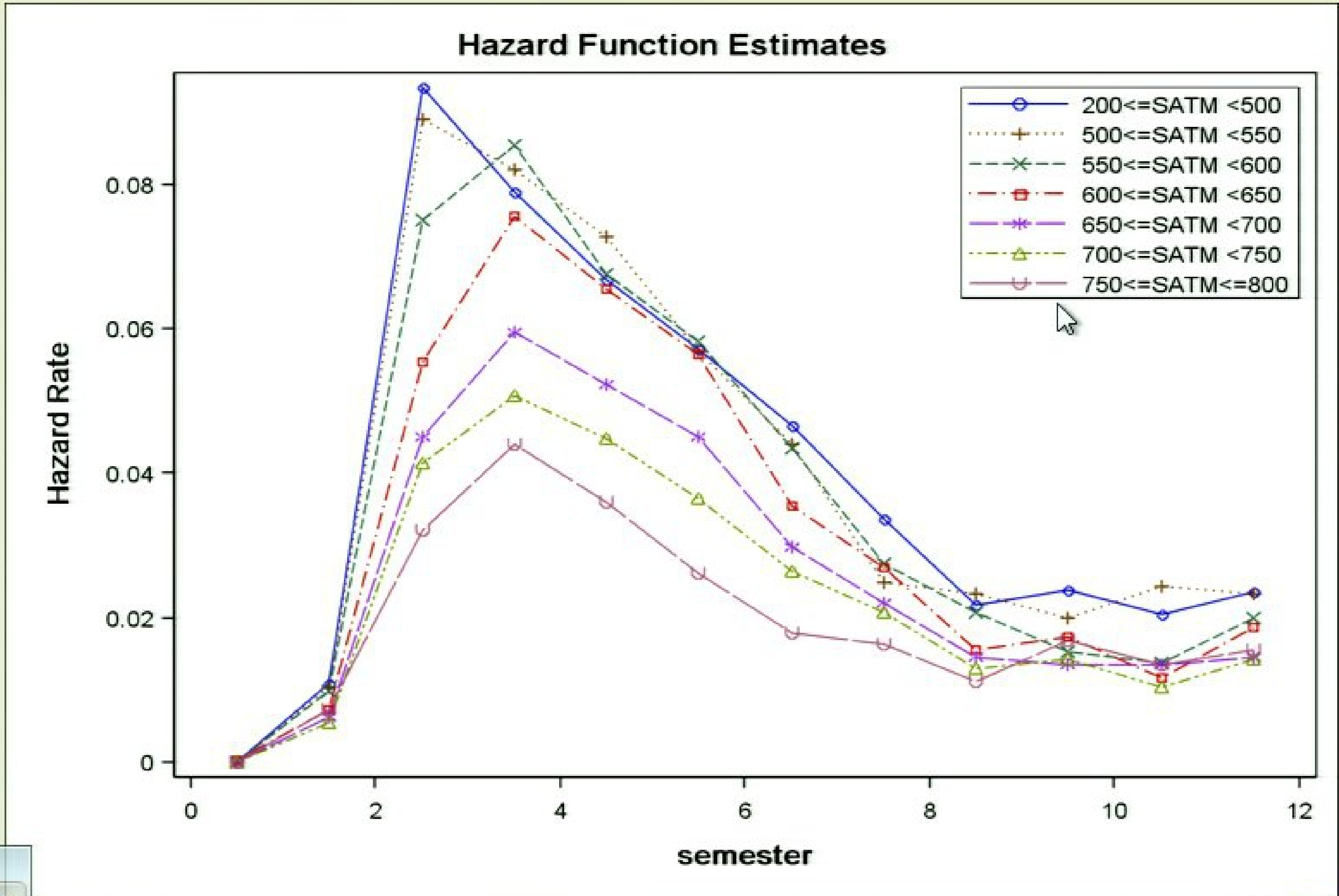


72/98

EAST CAROLINA UNIVERSITY

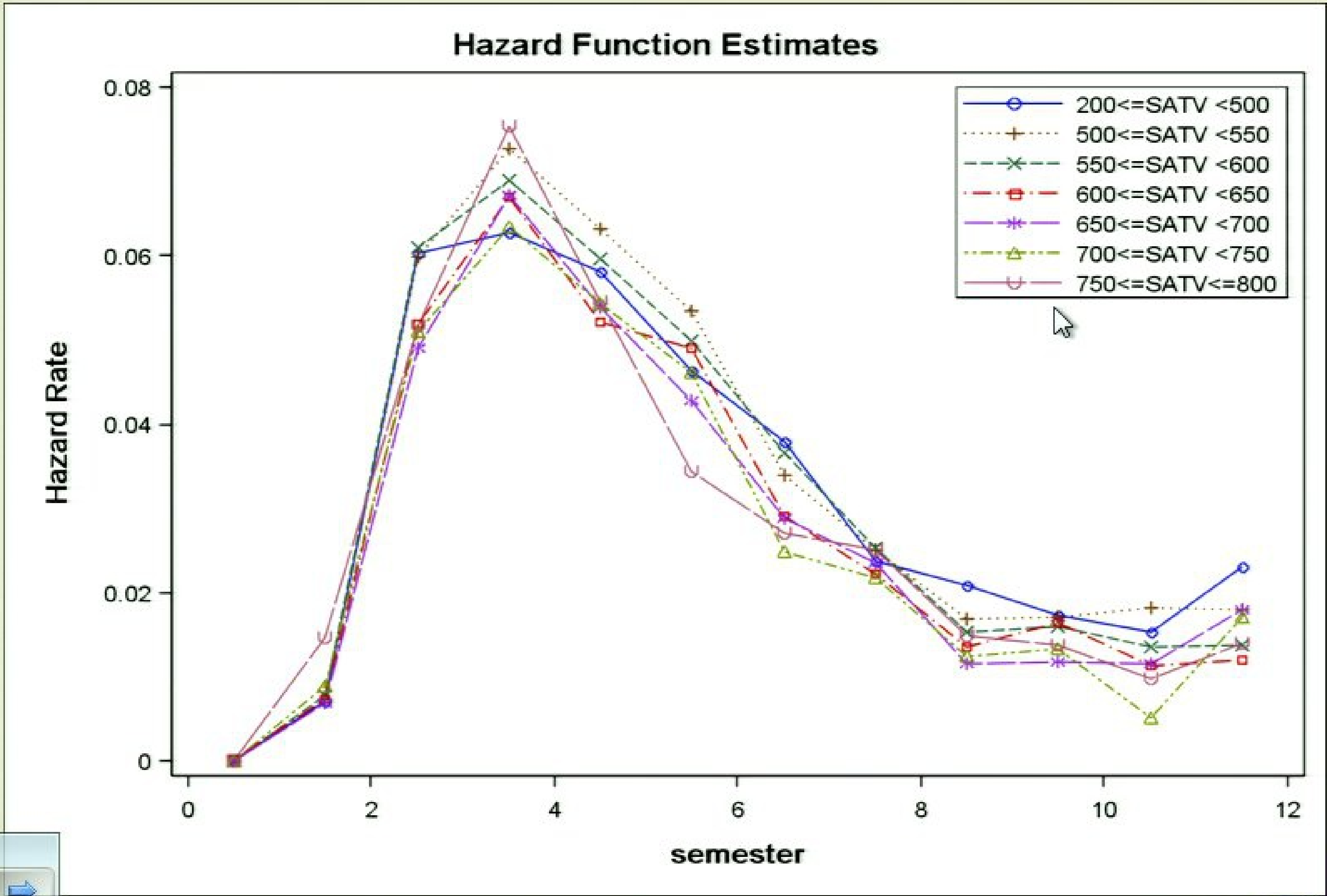
Results

Figure 9. Hazard functions by SAT math score groups



Results

Figure 9. Hazard functions by SAT Verbal score groups



72/98

E A S T
CAROLINA
UNIVERSITY

Summary of Key Results

73/98

EAST CAROLINA UNIVERSITY

Summary of Key Results

•There are **no significant differences** among cohort subgroups for long survival times, but there are significant differences between cohort subgroups **for early survival times**, and for gender, ethnicity group, and SAT math and verbal scores subgroups.

Summary of Key Results

- There are **no significant differences** among cohort subgroups for long survival times, but there are significant differences between cohort subgroups **for early survival times**, and for gender, ethnicity group, and SAT math and verbal scores subgroups.

- **Females are significantly more likely to be at risk of leaving engineering in semesters 3 to 5 than males**, while the risks are similar during other semesters.

Summary of Key Results

•There are **no significant differences** among cohort subgroups for long survival times, but there are significant differences between cohort subgroups **for early survival times**, and for gender, ethnicity group, and SAT math and verbal scores subgroups.

•**Females are significantly more likely to be at risk of leaving engineering in semesters 3 to 5 than males**, while the risks are similar during other semesters.

•**White students tend to leave engineering slightly more than minority** students, who leave engineering more than Asians, and Asian students leave engineering more than other students.

Summary of Key Results

- Generally, engineering college students **leave engineering** during the **3rd semester** rather than other semesters, which may be due to probationary periods offered in earlier periods. Students with an SAT math score < 550, however, tend to leave engineering during the 2nd semester.

CollegeBoard SAT
connect to college success™

EAST CAROLINA UNIVERSITY

Summary of Key Results

•Generally, engineering college students **leave engineering** during the **3rd semester** rather than other semesters, which may be due to probationary periods offered in earlier periods. Students with an SAT math score < 550, however, tend to leave engineering during the 2nd semester.

•**SAT math score better predicts the risk of failure than SAT verbal score,** and, not surprisingly, the lower a student's SAT math score, the more likely that student is to leave engineering.

CollegeBoard SAT
connect to college success™

EAST CAROLINA UNIVERSITY

74/98

Summary of Key Results

•Generally, engineering college students **leave engineering** during the **3rd semester** rather than other semesters, which may be due to probationary periods offered in earlier periods. Students with an SAT math score < 550, however, tend to leave engineering during the 2nd semester.

•**SAT math score better predicts the risk of failure than SAT verbal score,** and, not surprisingly, the lower a student's SAT math score, the more likely that student is to leave engineering.

•Engineering college students with SAT verbal score **< 500 tend to survive more than** the students whose SAT verbal is **between 500 and 600.**

CollegeBoard SAT
connect to college success™

EAST CAROLINA UNIVERSITY

74/98

Applied Research Study Example #4

Survival Analysis of Engineering Attrition

Significance of this Research

- **Survival analysis is superior in analyzing time-to-event** or longitudinal data because it exceptionally affords important information on variation in dropout risk at various times, and permits the effects of predictors to fluctuate over time.
- This research can **stimulate more advanced research on undergraduate student dropout** and **provide important knowledge** as basis for better dropout prevention.

Applied Research Study Example #5

Using CIPP Model to Guide Service Learning Project

- **Using the Context, Input, Process, and Product Evaluation Model (CIPP) as a Comprehensive Framework to Guide the Planning, Implementation, and Assessment of Service-learning Programs.**

  (in collaboration with Zeller, Griffith, Metcalf, Williams, Shea, & Misulis)

  Published in *Journal of Higher Education Outreach and Engagement*.

- **What is Service Learning**?

  Service-learning involves the integration of community service into the academic curriculum. Students learn while providing services that meet a community's needs.

EAST CAROLINA UNIVERSITY

Applied Research Study Example #5

Using CIPP Model to Guide Service Learning Project

Successful service-learning requires researchers to:

1. **Identify** the learning needs of service providers and the needs of community partners,
2. **Design** a program that can effectively address both needs, and
3. Successfully **implement** the programs to **generate** desired outcomes to meet both needs.

EAST CAROLINA UNIVERSITY

Using the CIPP Model to Guide Service-learning Program

•In spring **2008, a service-learning tutoring program** was initiated to address the identified learning needs of the preservice teachers in the Elementary Education program at ECU and those of at-risk readers in the local school system.

•**26 preservice teachers** taking a course in Diagnostic/Prescriptive Teaching of Reading completed a service-learning component by **tutoring 26 Response-to-Intervention students** (RTI at-risk readers) in kindergarten, first, and second grades at a local elementary school.

•**The CIPP model was used to systematically guide the conception, design, implementation, and assessment** of this tutoring program and led to its successful completion and attainment of desired outcomes. The four CIPP components were methodically conducted using a variety of assessment techniques.

EAST CAROLINA UNIVERSITY

Use the CIPP Model to Guide Service-learning Project

Use the CIPP Model to Guide Service-learning Project

•First, **Context Evaluation identified the needs** of pre-service teachers to gain firsthand experience of working with students from diverse background and the elementary school RTI at-risk readers' needs for individualized assistance on reading.

EAST CAROLINA UNIVERSITY

Use the CIPP Model to Guide Service-learning Project

- First, **Context Evaluation identified the needs** of pre-service teachers to gain firsthand experience of working with students from diverse background and the elementary school RTI at-risk readers' needs for individualized assistance on reading.
- Next, **Input Evaluation incorporated input** from experts, practitioners, and various stakeholders, and prescribed an effective tutoring project.

EAST CAROLINA UNIVERSITY

Use the CIPP Model to Guide Service-learning Project

- First, **Context Evaluation identified the needs** of pre-service teachers to gain firsthand experience of working with students from diverse background and the elementary school RTI at-risk readers' needs for individualized assistance on reading.
- Next, **Input Evaluation incorporated input** from experts, practitioners, and various stakeholders, and prescribed an effective tutoring project.
- Then **Process Evaluation helped monitor** the project implementation process and provided ongoing feedbacks for needed adjustments to the project.

Use the CIPP Model to Guide Service-learning Project

- First, **Context Evaluation identified the needs** of pre-service teachers to gain firsthand experience of working with students from diverse background and the elementary school RTI at-risk readers' needs for individualized assistance on reading.

- Next, **Input Evaluation incorporated input** from experts, practitioners, and various stakeholders, and prescribed an effective tutoring project.

- Then **Process Evaluation helped monitor** the project implementation process and provided ongoing feedbacks for needed adjustments to the project.

- Finally, **Product Evaluation assessed** the service-learning project's impacts and provided judgment of the project's effectiveness.

Use CIPP Model to Guide Service-learning Project

Context, Input, Process, and Product Model Components	Methods Used in the Service-learning Tutoring Program
Component I: Context evaluation Identify the needs, assets and opportunities for addressing the needs	• Assessed the setting for the intended service • Interviewed school principal, teachers, and reading specialists • Reviewed school records • Identified at-risk readers and their needs • Administered diagnostic tests to at-risk readers • Conducted initial quantitative assessment of at-risk readers • Conducted pre-service teacher focus group interviews • Conducted initial quantitative assessments of pre-service teachers
Component I: Input evaluation Prescribe a project to meet the identified needs and identify and assess project strategies and procedural designs	• Reviewed relevant literature • Interviewed school principal, teachers, and reading specialists • Consulted university reading faculty and other experts • Viewed exemplary projects • Consulted Learn and Serve America • Formed advocate teams • Service-learning taskforce members met bi-weekly • Conducted pre-service teacher focus group interviews
Component III: Process evaluation Monitor project's process and potential procedural barriers and identify needs for project adjustments	• Identified what activities should be monitored • Receivd bi-weekly update from service-learning taskforce • Observed service-learning activities • Kept a log of the activities • Interviewed at-risk readers • Interviewed preservice teachers • Interviewed school principal, teachers, and reading specialists • Reviewed preservice teachers' self-reflections • Reviewed students' work samples • Conducted debriefing with preservice teachers
Component IV: product evaluation Measure, interpret, and judge project outcomes and interpret their merit, worth, significance and probity	• Conducted post-project quantitative assessments of pre-service teachers • Conducted post-project focus group interview of pre-service teachers • Conducted post-project quantitative assessment of at-risk readers • Administered at-risk readers survey • Interviewed or surveyed other stakeholders including faculty instructor, principal, teacher, reading specialist, and parents of at-risk readers

EAST CAROLINA UNIVERSITY

Applied Research Study Series Example #6

Providing physical activity for students with intellectual disabilities

- **A Close Look at the Fitness Level of Elementary Students with Intellectual Disabilities.**

  Published in *Sport Science Review*.

- **Providing physical activity for students with intellectual disabilities: The motivate, adapt, and play (MAP) program.**

  Published in *Journal of Physical Education, Recreation & Dance*.

- **Promoting Health-Related Fitness for Elementary Students with Intellectual Disabilities through a Specifically Designed Activity Program.**

Published in *Journal of Policy and Practice in Intellectual Disabilities*.

*** *in collaboration with Boswell, Davis, Decker, & Hodson.*

31/98

EAST CAROLINA UNIVERSITY

Applied Research Study Series Example #6

Providing physical activity for students with intellectual disabilities

Background

- In NC, 20% of children are overweight or obese; NC is #5 in the nation for childhood obesity (BeActiveNC, 2006).
- Children with disabilities achieve insufficient physical activity for health purposes (Faison-Hodge & Poretta, 2004).
- Children with disabilities have high risk of Cardiovascular disease, diabetes, and hypertension.
- There has been little help on improving physical activity levels of students with intellectual disabilities.

Participants

- 4 elementary self-contained classes
- 25 students with intellectual disabilities, ages 8-12
- 9 girls, 16 boys
- 11 Caucasian, 11 African American, 3 Hispanic

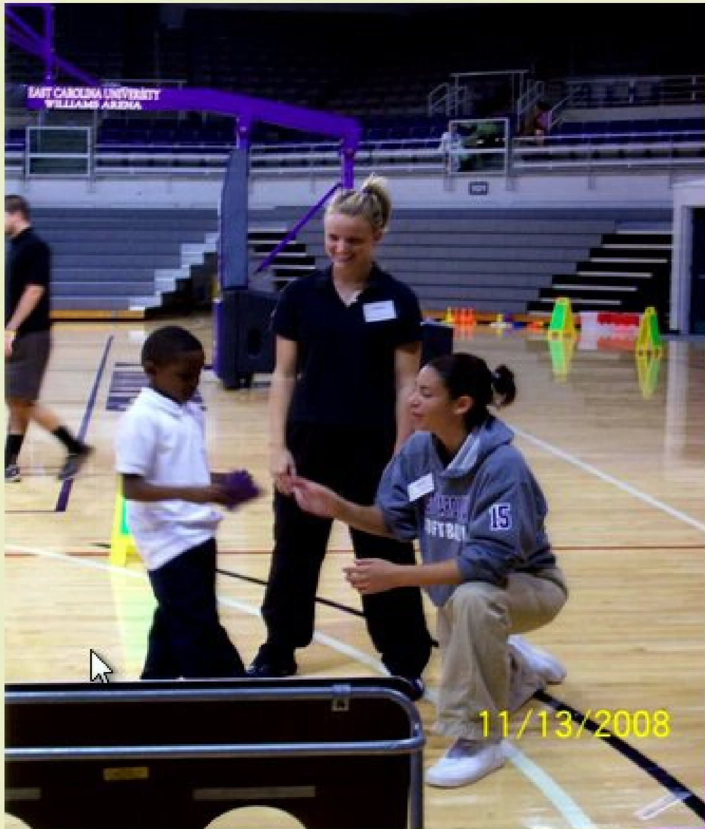
33/98

EAST CAROLINA UNIVERSITY

Pre-test Fitness Data

The pre-test data showed that students with intellectual disabilities had low fitness levels compared to healthy standards set for typical students:

- 31% had an unhealthy weight
- 93% did not meet standard for PACER
- 77% not able to complete the minimum standard for the modified pull-up
- 35% did not meet criteria for sit-and-reach



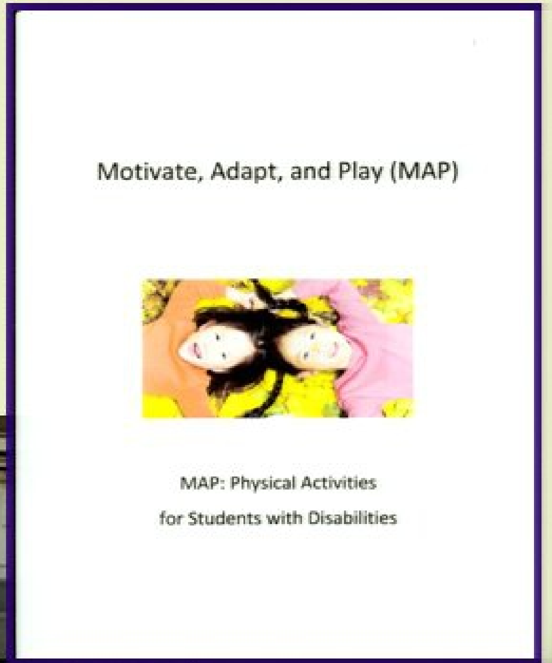
EAST CAROLINA UNIVERSITY

Applied Research Study Series Example #6

Providing physical activity for students with intellectual disabilities

- **M**otivate: The key is to find out what motivates students with intellectual disabilities to participate (work hard) in physical activity.
- **A**dapt: Teachers adapt all MAP activities to fit the ability/development of each student with intellectual disabilities.
- **P**lay: MAP activities were designed to be fun, make students laugh, and continue physical activity. MAP activities included dance, yoga, motor skills, aerobics, and manipulative skills.

****MAP program is designed by: Dr. Boni Boswell, Whitney Woodhall-Smith, Danielle Bogner .*


Motivate, Adapt, and Play (MAP)

MAP: Physical Activities
for Students with Disabilities

Applied Research Study Series Example #6

Student Performance Before and After the MAP Program (n=25)

Measures	Pre-Post	Mean	Sd	*t*	*df*	p-val	ES (*d*)
BMI	Before After	19.96 19.94	5.92 5.71	.056	24	.956	.01
PACER	Before After	10.28 13.24	11.72 10.89	2.69	24	.013*	.55
Modified Curl-up	Before After	7.12 12.64	11.80 10.55	2.089	24	.048*	.42
Medicine Ball Throw	Before After	101.52 129.08	73.92 80.02	2.081	24	.048*	.42
Sit-and-Reach Right	Before After	23.08 27.10	7.49 8.67	4.278	24	.000*	.882
Sit-and-Reach Left	Before After	21.86 26.68	7.84 8.07	5.329	24	.000*	1.07
Pedometer Steps	Before After	3185.20 3220.72	1236.68 1740.29	.111	24	.913	.02

Note: The * indicates statistical significance.

36/98

EAST CAROLINA UNIVERSITY

Applied Research Study Example #7

Podcasting: Engaging preservice teachers in learning about reading education with Web 2.0 technologies.

(with Swaggerty & Atkinson) published in *Teacher Education Quarterly*.

- This study **examines the impact of a podcasting project on preservice teachers' interests and attitudes** in two sections of an undergraduate reading methods course (n=54).

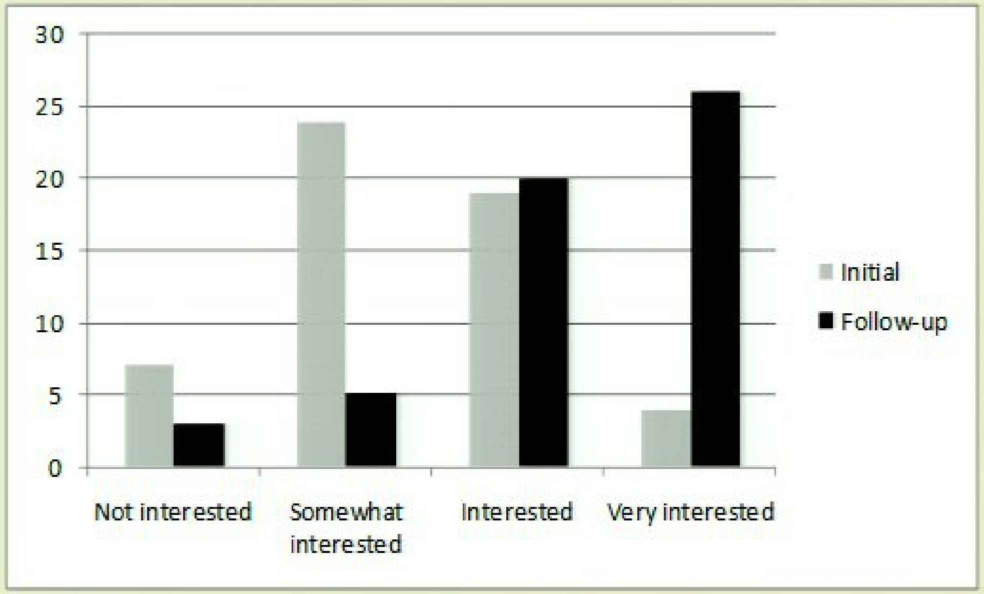
- **Students** worked in pairs or groups of three to **create an episode of the "Literacy Superstar Podcast"** for the course website. Students chose a prominent figure in the field of literacy education, researched the literacy superstar's life and his or her impact on the field, synthesized the information, and created a script for the project.

EAST CAROLINA UNIVERSITY

37/98

Applied Research Study Example #7

Comparison of Preservice Teachers' Interest in Using Podcasts for Educational Purposes before and after the Podcast Project



Survey	n	Mean	*sd*	*t*	*p*	*d*
Before	54	2.37	0.808			
After	54	3.28	0.856	6.829	.000	0.932

38/98

EAST CAROLINA UNIVERSITY

Selected Grants

- **Guili Zhang, PI/PD. IGERT: Spatial Ecology and Evolution: Integrative Quantitative Training in Biology, Statistics, and Mathematics, Project evaluation grant.** Funded by the National Science Foundation (NSF). The Integrative Graduate Education and Research Traineeship (IGERT) is NSF's flagship interdisciplinary training program, educating U.S. Ph.D. scientists and engineers by building on the foundations of their disciplinary knowledge with interdisciplinary training.

- **Guili Zhang, Program Evaluator & Statistical Analyst. SUCCEED (Southeastern University and College Coalition for Engineering Education).** Funded by NSF. It takes a holistic approach to curriculum reform and leads the systemic reform in the nation's engineering education.

- **Guili Zhang, Program Evaluator & Statistical Analyst. "Studies Using the Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD)".** Funded by NSF. The MIDFIELD database has been a unique and an exceptionally useful source for the nation's engineering education research.

39/98

EAST CAROLINA UNIVERSITY

Selected Grants, cont'd

- **Guili Zhang, Project Evaluator, Teacher Quality Partnership Project (TQP).** Funded by the U.S. Department of Education. The TQP Grants Program seeks to improve the quality of new teachers by creating partnerships among IHEs, high-need school districts (local educational agencies (LEAs)) their high-need schools, and/or high-need early childhood education (ECE) program.

Guili Zhang, PI. 2010 American Educational Research Association Statistics Institute for Faculty Grant. Funded by NSF and the American Educational Research Association**.**

- **Guili Zhang, Co-PI** (with Davis, Boswell, Decker, & Hodson). **Motivate, Adapt, and Play (MAP).** Funded by Division of Research and Graduate Studies Research Development Grant. This project involves developing, implementing, and assessing a physical activity program for youth with intellectual disabilities in order to increase their health-related fitness level.

90/98

EAST CAROLINA UNIVERSITY

Selected Grants, cont'd

- **Guili Zhang, PI. Learning to Teach, Learning to Serve Evaluation Grant.** Funded by Learn & Serve America. Students achieve learning goals while providing meaningful service to the community.

- **Guili Zhang, Co-PI** (with Zeller)**. Learning to Teach, Learning to Serve Evaluation Grant.** Funded by Learn & Serve America. . Students achieve learning goals while providing meaningful service to the community.

- **Guili Zhang, Co-PI** (with Shea)**. Learning to Teach, Learning to Serve Program Grant.** Funded by Learn & Serve America**.** It aims to strengthen teacher preparation programs by incorporating service learning in teacher education curriculum.

- **Guili Zhang, PI. 2012 AERA Institute on Statistical Analysis for Education Policy Grant.** Funded by the American Educational Research Association (AERA) Grants Program which is supported by NSF.

EAST CAROLINA UNIVERSITY


National Leadership Services
Chair, Quantitative Methods TIG, American Evaluation Association.
Executive Council Member of American Educational Research Association Division D. (Measurement and Research Methodology).
Chair, Outstanding Quantitative Dissertation Award Committee of American Educational Research Association (AERA), 2009-2011.
Program Co-Chair, Quantitative TIG, American Evaluation Association.
Chair, Phenomenology, Hermeneutics, & Critical Theories in Qualitative Research session of the 2012 AERA Annual Meeting in Vancouver, British Columbia, Canada.
92
EAST CAROLINA UNIVERSITY

National Leadership Services, cont'd

- **Advisory Board Member:** *Journal of Curriculum and Instruction,* 2009-present.

- **Chair,** Research on Evaluation Session, American Educational Research Association Annual Conference.

- **Chair,** Statistics and Effect Sizes in Experimental Designs Session, Educational Statisticians SIG, American Educational Research Association Annual Conference.

- **Co-Chair**, Outstanding Quantitative Dissertation Award Committee of American Educational Research Association (AERA), 2008-2009.

- **Member**, American Educational Research Association (AERA) Division D Measurement & Research Methodology Outstanding Quantitative Dissertation Award Committee, 2007-present

EAST CAROLINA UNIVERSITY

National Leadership Services, cont'd

- **Chair**, Learning Communities Session, American Institute of Higher Education National Conference, Orlando, FL, April 2008
- **Co-Editor**, *Journal of Curriculum and Instruction*, 2006-2010.
- **Invited Presenter:** Beijing Teacher's Training Center for Higher Education, People's Republic of China
- **Advisor:** Shandong Normal University, People's Republic of China

Selected National Awards & Honors

- ***Outstanding Professor and Researcher***

  The Government of the United States (USCIS category E17 – Outstanding Professor and Researcher).

- ***2010 Edward C. Pomeroy Award for Outstanding Contributions to Teacher Education*** (with JoCI Co-editors: Atkinson, O'Connor, Steinweg, Kester, Rodriguez, & Swaggerty)

  The American Association of Colleges for Teacher Education.

- ***Benjamin J. Dasher Best Paper Award***

  Frontiers in (engineering) Education (FIE)

- ***Best Paper Award***

  American Society for Engineering Education (ASEE)

Selected National Awards & Honors

- ***2012 AERA Institute on Statistical Analysis for Education Policy Award.***

  American Educational Research Association & National Science Foundation.

- ***2010 American Educational Research Association Statistics Institute for Faculty Award***

  National Science Foundation & American Educational Research Association

- ***American Educational Research Association Outstanding Quantitative Dissertation Award Finalist***

  American Educational Research Association

- ***Excellent Lesson Award***

  National Teaching Competition, Education Ministry of China

- ***Excellent Lesson Award***

  The 6th Annual Teaching Research Conference of Eastern China (Six Provinces)


Thank You
98/98