

## Quantitative Research Design for Language and Linguistics

English 618

### April Ginther

aginther@purdue.edu

Young 816

Office Hours:

### Messages and appointments

Nina Barron-Burke (OEPP Secretary)

nbarronb@purdue.edu

Young 809

494-3780

### Texts

Leavitt, F. (2001). Evaluating Scientific Research: Separating Fact from Fiction.

Upper Saddle River, NJ: Pearson Education. ISBN 0-13-012845-7

Moore, D. S. (1999). Against All Odds: Inside Statistics. Annenberg CPB Project.

Telecourse Study.

### Additional Readings

Anghoff, W. (1988). Validity: An evolving concept. In H. Wainer & H. Braun (Eds.) Test Validity (pp. 19-32). Hillsdale, NJ: Erlbaum.

Booth, W.C., Colomb, G.G., and Williams, J.M. (2003). The Craft of Research, 2<sup>nd</sup> Ed.

Chicago: University of Chicago Press. Chapters 3 and 4: 35-70.

Eisenhart, M. and DeHaan, R.L. (2005). Doctoral Preparation of Scientifically Based Ed Researchers. Educational Researcher, 34 (4) 3-13.

Ercikan, K. and Roth, W-M. (2006) What good is polarizing research into qualitative and quantitative? Educational Researcher, 35 (5) 14-23.

Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. Language Testing, 19, 133-167.

Kirp, D.L. (2006). After the bell curve. New York Times Magazine. July 23, 2006, 15-16.

Klinger, J.K., Scanlon, D., and Pressley, M. (2005). How to publish in scholarly journals. Educational Researcher, 34 (8) 14-20.

Lazaraton, A. (2000). Current Trends in Research Methodology and Statistics in Applied Linguistics. TESOL Quarterly, 34 (1) 175-181.

Quantitative Research Design for Language and Linguistics (618) is an introduction to quantitative research methodology associated with the social sciences. Members of this class generally have different interests and are enrolled in different programs within the university but share the need and desire to become both effective consumers and efficient producers of quality research. Throughout the course, the idea of evidence, or more accurately *the quality of evidence*, as the basis for inferential reasoning will be emphasized.

### Class Requirements

Class Participation	10%	<p>Each homework assignment carries the same weight.</p> <p>Your final grade will be determined by your average for each class component as associated with the percentage for each component.</p>
Homework Assignments	20%	
Summaries	--	
Problems	--	
Article presentation	--	
Proposals	25%	
Proposal Reviews	20%	
Final Paper	25%	
<b>Total</b>	<b>100%</b>	

### Class Participation/Preparation

Prepare for class by carefully reading each assignment and by writing a question, series of questions, and/or comments to facilitate discussion. E-mail these questions to everyone in the class by using the "respond to all" function. Submit your questions to the class by 9:00 a.m. every Wednesday. Put 618, the week, and your topic on the subject line (e.g. 618, Week 2, falsifiability).

### Homework

Summaries: The majority of your homework assignments (11 of 16) involve watching and then summarizing selections from Against All Odds: Inside Statistics. DVD and VHS copies of all chapters of Against All Odds are available for checkout in from the OEPP library (see Nina Barron Burke, OEPP secretary). The summaries should not exceed 500 words - short is fine. Comment on anything you found interesting. Your goal is to demonstrate to me that you watched the DVD and understood the concept(s) presented.

However, please let me know if there's something that you did not understand or are confused about -- chances are others have encountered the same problem. We will touch on the content of the telecourse in class but only briefly.

Problems: Homework problems (#8, #11, #12, #15) are provided on the pages following the class schedule.

Article Presentations: Each student will present an article. (You have already been assigned a date.) Select anything of interest that includes a quantitative analysis. Use the **Checklist for Evaluating Research** (Leavitt, p. xv-xvi) to prepare your 20 minute presentation, and be prepared to answer questions as well (10 minutes). Select an article that presents a study that you would like to conduct (think about replication). Make an appointment to discuss your choice with me and to address any questions you may have at least one week before your presentation.

## **Proposals**

On Friday, March 9<sup>th</sup>, your preliminary research papers (aka -- 'proposals') are due. These papers can be thought of as proposals and should include the following sections: abstract (250 words), statement of the problem (250-500 words), research questions (250 words), literature review (5-10 pages), methodology (5 pages), and references (at least 10); that is, the preliminary paper will include all of the sections of your final research paper except the results and discussion. (The word and page specifications are approximations.) Your methodology section should include a description of the subjects, data collection procedures, and the methodology that you plan to use to collect your data and analyze your results. Your methodology section might also include appropriate references for the methodology, i.e., cite references, texts that explain and studies that have used comparable methods. These proposals should be double-spaced with one-inch margins on all sides and should follow the style specifications outlined by the American Psychological Association (APA) Style Manual of Style. Your preliminary papers are identical in format to what is minimally required for most research proposals that involve funding and can easily be expanded in order to meet your program's prospectus requirements.

You are required to submit three copies of your preliminary paper to me in class on Wednesday, March 7<sup>th</sup>. Two copies of your paper will be distributed to two of your classmates who will review your submission over Spring Break.

### Reviews of Preliminary Papers

As a peer reviewer, you should function as both an editor and reviewer. Mark any grammatical or stylistic problems that you notice. Your job is to help the writer improve the quality of the final draft. In addition to editing the paper, you should write a formal review that comments on the proposal's strengths and weaknesses. Your review should not exceed 3 double-spaced pages in length. These reviews will be returned to the authors and to me on Wednesday, March 21.

### Final Paper

On Wednesday, May 2nd (the Wednesday of finals week), your final papers are due. Please turn them in to Nina Barron-Burke in the OEPP office —Young 809. These papers should be a revision of your preliminary paper with the results and discussion added. In a separate section, all comments that were made by the anonymous reviewers about your prelim should be directly addressed. The paper, with all required sections, appendices, tables, figures, and references should not exceed 30 pages in length.

<u>Against All Odds: Inside Statistics</u>			
1	What is Statistics?	14	Samples and Surveys
2	Picturing Distributions	15	What is Probability?
3	Describing Distributions	16	Random Variables
4	Normal Distributions	17	Binomial Distributions
5	Normal Calculations	18	The Sample Mean and Control Charts
6	Time Series	19	Confidence Intervals
7	Models for Growth	20	Significance Tests
8	Describing Relationships	21	Inference for One Mean
9	Correlation	22	Comparing Two Means
10	Multidimensional Data Analysis	23	Inference for Proportions
11	The Question of Causation	24	Inference for Two-Way Tables
12	Experimental Design	25	Inference for Relationships
13	Blocking and Sampling	26	Case Study: Developing AZT

## Class Schedule and Homework Assignments\*

1	01/10	Introduction
2	01/17	Leavitt, Chapter 1: What is Science? Leavitt, Chapter 2: Reading and Reviewing Scientific Literature. Eisenhart & DeHaan: Doctoral Preparation of Scientifically Based Ed Researchers. Lazarton: Current Trends in Research Methodology and Statistics in Applied Linguistics. <b>HW1: <i>Against All Odds</i>, 1: What is Statistics? Summary Due.</b>
3	01/24	Leavitt, Chapter 3: Conflicts of Interest and Bias. Leavitt, Chapter 4: Finding Interesting Problems and Studying Them Creatively. Booth, Colomb, & Williams, Chapter 3: From Topics to Questions. Booth, Colomb, & Williams, Chapter 4: From Questions to Problems. <b>HW2: <i>Against All Odds</i>, 2: Picturing Distributions. Summary Due.</b>
4	01/31	<u>Lecture</u> : Measurement Scales and Measures of Central Tendency. Leavitt, Chapter 5: Selecting and Measuring Variables. Leavitt, Chapter 6: Choosing the Best Research Design. <b>HW3: <i>Against All Odds</i>, 3: Describing Distributions. Summary Due.</b> <b>HW4: <i>Against All Odds</i>, 4: Normal Distributions. Summary Due.</b> Article Presentation 1: Lixia Cheng
5	02/07	Leavitt, Chapter 7: Experimenting: Two Groups. <b>HW5: <i>Against All Odds</i>, 21: Inference for One Mean. Summary Due.</b> <b>HW6: <i>Against All Odds</i>, 22: Comparing Two Means. Summary Due.</b> Article Presentation 2: Fatima Esseili
6	02/14	<u>Lecture</u> : Interactions. Leavitt, Chapter 8: Variations on the Simple Experiment. Ginther, Context and Content Visuals and Performance on Listening Comprehension Stimuli. <b>HW7: <i>Against All Odds</i>, 13: Blocking and Sampling. Summary Due.</b> Article Presentation 3: Christopher La Cross
7	02/21	Kirp: After the Bell Curve. <b>HW8: Interactions. Due.</b> Article Presentation 4: Xinqiang Li
8	02/28	<u>Lecture</u> : Correlation. Leavitt: Chapter 10: Correlational Strategies to Predict and Assess Relationships. <b>HW9: <i>Against All Odds</i>, 8: Describing Relationships. Summary Due.</b> <b>HW10: <i>Against All Odds</i>, 9: Correlation. Summary Due.</b> <b>HW11: Volleyball Variables. DUE Monday March 5<sup>th</sup>.</b> Article Presentation 5: Karen Morgan

**Class Schedule and Homework Assignments, cont.\***

9	03/07	<p><u>Lecture:</u> Correlation, continued.  Volleyball Variables. Discussion.  <b>HW12:</b> Reviewers Comments of Ginther, Context and Content Visuals and Performance on List  Comprehension Stimuli. Discussion.  <b>Proposals Due. Three copies.</b>  Article Presentation 6: Sunny Park</p>
10	03/14	Spring Break
11	03/21	<p>Leavitt: Chapter 11: Case Studies.  Leavitt: Chapter 12: Observing.</p>
12	03/04	<p>Leavitt: Chapter 13: Surveys.  <b>HW13:</b> <i>Against All Odds</i>, 14: Samples and Surveys. Summary Due.  <b>HW14:</b> <i>Against All Odds</i>, 19: Confidence Intervals. Summary Due.</p>
13	04/11	<p><u>Lecture:</u> Validity.  Anghoff: Validity: An Evolving Concept.  Leavitt: Chapter 17: Data Analysis.  Leavitt: Chapter 18: Philosophical Challenges.</p>
14	04/18	<p>Klinger, Scanlon, and Pressley: How to Publish in Scholarly Journals.  <b>HW 15:</b> Abstract Reviews.</p>
15	04/25	Paper Presentations. Class Evaluations.
16	05/02	<b>Final Papers Due // Wednesday May 2nd by 4:00 p.m. // OEPP OFFICE</b>

Note: \* The schedule may change depending on our ability to work through the material.

## HW8: Interactions. DUE February 21.

### Preparation

- 1) Read Ginther, Context and Content Visuals and Performance on Listening Comprehension Stimuli, carefully long before the homework assignment is due. Ask any questions in class or via e-mail that need to be addressed.
- 2) Use the **Checklist for Evaluating Research** (Leavitt, p. xv-xvi) to make sure you understand what's going on - you should be able to answer all of the questions on the guide.
- 3) Make sure you can identify the DV and the IVS. How many independent variables are there and how many levels do each of the independent variables have?
- 4) Leavitt (p.111) explains that there are six possible outcomes for a 2x2 factorial study. This means that there are 2 IVs with 2 levels each; for example, with reference to the article, if the design were actually a 2x2, it could have been = PROFICIENCY (high or low) x VISUALS (present or absent).

### Assignment

- 5) **Imagine that the study was actually a 2X2 with Proficiency by Visual condition as the IVs. Graph (both ways, i.e., 12 graphs total) the six possible outcomes and write a paragraph discussing each. Indicate which representation you would choose for the publication. Include believable means and standard deviations. Include, at least, a one sentence interpretation: Does the interaction make sense? Why might it have occurred?**
- 6) **Repeat above with variables of interest that you have chosen.**

**HW11: Volleyball Variables. DUE Monday March 5<sup>th</sup>.**

**HW12: Reviewers Comments of Ginther, Context and Content Visuals and Performance on Listening Comprehension Stimuli.**

Read the reviewers comments. Ask questions in class or via e-mail about what I changed or did not change.

**Reviewer #1**

I would recommend this paper for publication in *Language Testing*. It is detailed and well-written. It needs no major revisions. The design was complex, but adequately explained. I have a few clarification issues that the author(s) may want to address, but these are optional (unless the editors think otherwise).

1. At the end of the first paragraph of the introduction, the author states the three stimulus types. Although explained later in detail, a very brief description of these here might help, since they are talked about quite a bit at the end of the introduction but are not fully described until after the lit review. Again, the description need only be very brief.
2. At the beginning of the results section, in order to make the results of the complex analysis more concrete, the author may want to provide the means for the main effects (proficiency) and nested effects (status), perhaps main effects for form (although there were 16 forms!), and main effects of stimulus type, time, and presence or absence of visuals. I can understand, however, that these might distract the reader from the author’s main argument, but some type of simple descriptive statistics on performance towards the beginning of the results section would be helpful.
3. Right before “Insert Figure 1”, the author writes “If the level of proficiency of the high group were extended....” This is a bit confusing. Perhaps the author could just point out that these two lines are not parallel, thus an interaction.
4. I think something is missing in this sentence from the discussion section: “Given the very generic nature of both the visual and the verbal information presented in TOEFL Dialogues/Short Conversations, it is hardly surprising that the presence of visuals in this condition was virtually nonexistent.”

This was an excellent paper. I enjoyed it thoroughly.

-----

**Reviewer #2**

Accept with amendments.

This is a very professional piece of research and the report is clearly written. The descriptions of all the stages of the research are clear, and the tables and graphs are lucid. There is, though, evidence of last minute rush with the occasional typo, and references to wrong tables etc.

The findings of the research are very interesting.

I have various questions though, two of which can be easily answered:

1. At what stage are the visuals provided for the Dialogues? Do they appear *before* the listening texts, or at exactly the same time? This might make a difference to the subjects' responses. (If it is *before*, then I don't see that there would be any problems with the prompts contradicting the hearer's own imagined context, see Salomon's third quote [Page 3, according to my numbering].)
2. I assume that for the Short Conversations and the Academic Discussions, the context visuals appear just before the listening text. Is this right, or do the photos not appear until the text starts to be played?

Following on from these two questions, there needs to be something in the text to tie in the TOEFL Listening test use of visuals to theories about the effect of when the context of a test item is given.

My other main query relates to the interaction between the supply of visuals and proficiency level. The answer to this research question is supplied under the heading '*Proficiency and visual condition*', but, presumably because of the lack of a significant interaction, there is little discussion of the results. Since this is such an interesting question, shouldn't there be more discussion about this, and possible reasons for the lack of significance in the findings? In any case, I think this section should be given more prominence, as it is providing the answer to one of the research questions.

### **Specific Points**

Just under the three numbered items referring to the content of the CBT Listening test (Page 3, according to my numbering) the first sentence starts with the word 'clearly'. Why do the majority of the visuals accompanying the audio texts 'clearly' depict context-related rather than content-related information?' Is it because there are more items relating to dialogues/short conversations and academic discussions than there are to mini-talks?

Four paragraphs below '*4 Design*': I'm not sure why we have to '*imagine*' that low proficiency subject #1 was administered form 1. Why can't we simply have statements of fact? E.g. Subject 1 (or whoever), who was of low proficiency, took form 1, and answered the 40 items in the visual conditions presented in the first column ...'. The subjects are completely anonymous, so I should have thought this form of reporting would be perfectly acceptable.

That same paragraph ('Finally, the stimulus subsets within each type of stimulus were intended to have comparable content and difficulty'): Had the items been pre-tested so that their difficulty levels were at least approximately known?

Two paragraphs below '*Insert Table 1*': The three 'md's should be spelt out in full, as otherwise some readers will not know whether they refer to the mode or the median.

Three paragraphs below '*Insert Table 1*': The education of the students should be explained more fully since not all readers will be familiar with the USA education system. For example, what is a 'high school degree', and how does this differ from a 'college degree'?

If *Language Testing* wants this article to be slightly shorter, the article could be shortened under 'Insert Table 3' by reducing the description of the TOEFL CBT. This material is slightly repetitive; we already know a lot of the information provided here from the description of the test at the beginning of the paper so I suggest the description of the test be much reduced here. Perhaps there should be a full description of the test at the beginning of the paper, and then simply a reference to it under '*Characteristics of TOEFL CBT Listening Comprehension Stimuli*'.

If *Language Testing* is short of space I would have also thought that Appendix 1, and Table 4 could be reduced so that only the top one or two boxes are provided and the others are explained verbally. Similarly it seems to me that not so many details are needed in Table 3. I suggest that languages with more than two speakers (Arabic, Chinese, German, Hindi, Indonesian, Korean, Malay and Spanish) should be listed and that the other languages should be covered by a sentence saying something like 'the other subjects spoke as their L1s a range of different languages from the Middle East, Europe, Africa and Asia'.

I'm not quite sure what the purpose is of the page at the end which follows Fig. 3. This information, where interactions are in any case ascribed to the wrong figs, seems to me to be repeating information adequately given in the three figs.

The paragraph above '*Insert Figure 3*': Are the differences in means significant? This may be clear to some readers, but it wasn't to me.

The paragraph below '*Insert Figure 3*': What does part of the first sentence mean ('... in order for an effect to be considered small ...')? Does this mean that the reader wants a small rather than a large effect, or does it mean that in order to have any affect at all...?  
-----

### Reviewer #3

Overall comments: Overall this is a very good paper. It is an important study that is well executed and interesting to read. I commend the author(s) on the paper.

Specific comments: The results are interesting to me because they are worth reporting and considering even though they are small statistical effects. These smallish effects are being detected because of the degrees of freedom in the denominator of the F-tests conducted. I am therefore, pleased to see the author(s) report effect sizes. Although interpreting the magnitude of effect sizes is somewhat subjective, it is still true that small effects are sometimes worth reporting and considering (see Abelson, 1985, Prentice & Miller, 1992). My point is that the effects reported are, in my opinion, small but worthy of consideration. This may be an important point to communicate to the readers.

Comments on writing and presentation:

1. On page 2, the citation for the Cronbach quotation should also include the page number from the Cronbach article.
2. In the last paragraph on Page 2, the word "conservative" is unnecessarily distracting when describing the experimental approach. That word seems like the beginning of an anti experimentalism polemic -- and this takes away from the point of the sentence and paragraph.
3. On page 3, four lines from the bottom, the "TD" should be written out in full.
4. On page 4, the paragraph that begin with "The focus of this study involves a fundamental concern ..." seems to miss the point that an artificially inflated performance due to the presence of visuals is also problematic. It is true that in the high stakes world of testing the primary focus would be on the associated debilitation of performance due to the presence of visual stimuli. I suppose that this is a minor subtle point but I do not like the (unintended) implication that as long as we only concern ourselves with downward bias (i.e., debilitation) then everything is ok in high stakes testing. We should also concern ourselves with the artificially upward bias (i.e., the

positive effect on performance) associated with the presence of visual stimuli. In terms of good testing practice, the positive effect of stimuli is also problematic, even in high stakes testing. If I have misinterpreted the author(s) then she/he should make their point much clearer in that paragraph.

References:

Abelson, R.P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, 97, 129-133.

Prentice, D.A., & Miller, D.T. (1992). When small effects are impressive. *Psychological Bulletin*, 112, 160-164.

**HW 15: Abstract Reviews. Due 4/18.**

**Homework 2 (Abstracts 1 – 3):** As a reviewer for conferences and journals, I have read abstracts similar to the three presented below. Your task in this homework assignment is to evaluate these three abstracts and make an explicit recommendation to an imagined applied linguistics/language learning conference committee: accept or reject. Consider audience interest in the topic, theoretical orientation, research design and methodology. Write a short response (500 words max.) ***that explains your decision*** to the presentation committee and, *potentially, to the submitter*. Give specific reasons for your decision, and if you have recommended that the paper be rejected make specific suggestions as to how the abstract might have been improved.

**ABSTRACT 1: Nikapuendan students' attitudes towards Alginan English**

Alginan English, a new variety of English, which has its own lexical and phonological features, serves as the lingua franca, the medium of instruction, and an expression of regional identity in the Alginan region of Nikapuenda. However, little is known about whether Alginan English is “advanced” enough to be accepted as a medium of instruction by students who are of other Nikapuendan ethnicities. This paper presents the results of a survey conducted to examine Nikapuendan students' (who are NOT Alginan) attitudes toward Alginan English.

A questionnaire was administered to 197 pre-university students from other regions in Nikapuenda who did not identify themselves as of Alginan ethnicity. The 22 statements on the questionnaire were designed to survey students' motivations for learning English and their attitudes about the variety of English used by Alginans. A four point Likert scale ranging from “Strongly Agree” (1) to “Strongly Disagree” (4) was used to force a positive or negative choice.

Results showed that students most strongly agreed with the statement “Alginans should learn American or British English” (mean=1.92, sd=0.71). They agreed that “Alginans

can write English accurately" (mean=2.11, sd=0.49) but strongly disagreed with the statement "English spoken by Alginans is easy to understand" (mean=3.16, sd=0.70). One possible explanation is that these Nikapuendan students have been exposed to "classroom English" using American or British standards as the norm. They are not familiar with other accents, nor have they been exposed to "real life" English used in different domains. The students' less favorable view of Alginan pronunciation, however, does not appear to affect their attitudes towards Alginans, as they strongly disagreed with the statement, "I do not make friends with Alginans easily" (mean=3.60, sd=0.58).

This paper concludes that language teachers should increase students familiarity with the varieties of English used in Nikapuenda. Language activities can be designed to help these Nikapuendan students improve their intercultural communication skills.

The presenter will begin by showing examples of written Alginan English and playing audio recordings of spoken Alginan English. Overhead transparencies are used to present the outline of the paper and the survey results. The presenter finishes by leading a discussion about raising students' familiarity with new varieties of English in multicultural contexts.

## **ABSTRACT 2: Multimedia, computer-based tests, and test-taking strategies**

In comparison with a tradition paper-pencil test, which limits the generalizability of inferences of test takers' listening comprehension to real-life situations beyond the test itself, a computer-based test offers a wide range of advantages for test users. One of its most attractive features is to have test takers actively participate in the test utilizing audio, visual, and contextual cues. This feature can be supported by a number of researchers, acknowledging that listening involves linguistic knowledge as well as extra-linguistic information processing such as dealing with contextual or situational cues. However, the effects of multimedia and test takers' listening strategies on the computer-based test performance have not yet been investigated. In order to provide evidence for valid inferences, an investigation of test takers' behavior while taking the computer-based listening test is of great importance, as this will provide some insight into the constructs to be measured in the computer-based listening test.

The purpose of the present study is to investigate test takers listening strategies and examine whether these contribute construct - relevant or - irrelevant variance. In addition, the relationship between those strategies and test performance is explored. The instrument being used is a Chinese computer-based placement test developed at the University of Northern Montana. The subjects are from beginning, intermediate, and advanced level of Chinese courses at UNM. The preliminary results have indicated that test takers' strategies were diverse, depending on task types and difficulty levels. The test taker's proficiency level has also been found be a critical factor in strategy use. For instance, learners at higher levels tended to focus on specific information that was directly related to the answer of the question ignoring the rest text, while learners at lower level tended to listen to the whole text repeatedly. In addition, computer-based specific strategies were observed, raising such issues as the effect of visual images and the text on the video clip, and the number of listening repetitions on the test performance. Although this study was conducted with a small sample size in a

qualitative way, it is hoped that the study will provide some valuable guidelines for developing and validating the computer-based listening tests.

### **ABSTRACT 3: Metaphor and communicative effectiveness in teaching simulations**

The trend toward more authentic communicative teaching has led to a broadening of the factors that are considered to be justifiably presented in classroom situations to include sociolinguistic, strategic, and pragmatic competencies. One recurrent pragmatic feature in the presentation of scientific or technological information is the use of metaphors or analogies (Kuhn 1993, Gentner & Jeziorski 1993, Dunbar 1995). However, recent investigations into the use of metaphors and analogies in the presentation of scientific and technological information in the classroom, have suggested that they can lead to miscommunications and misunderstandings if they are not sufficiently explained and extended (Engle 2000). The present study investigates the way in which prospective International Teaching Assistants' (ITAs) use of metaphors and analogies in the context of a teaching simulation affects the ratings they are given for sociolinguistic competence. From a sample of 89 performance tests, we analyzed the transcripts of 30 prospective ITAs whose presentations involved the use of metaphor to describe technological topics related to civil, electrical, industrial or mechanical engineering. These thirty metaphors were analyzed for their effectiveness. Effectiveness was evaluated by two independent raters in terms of correspondence to the explicated concept by counting the frequency of explicit and implicit links between the two (rater agreement 73%). A cut-off for effectiveness was established which classified half of the thirty presentations as containing metaphors that were evaluated as "sufficiently explicated"; the other half did not. Another two independent raters then assessed the simulations in four categories: linguistic competence (0-20 points), aural competence (0-10 points), sociolinguistic/strategic competence (0-10 points), and discourse/functional competence (0-10 points) and made a general pass/fail recommendation [inter-rater reliability .82]. We then examined the relative contribution of linguistic, aural, sociolinguistic and discourse competence to overall simulation performance for the two groups of ITAs. For students with combined linguistic and aural competence scores between 23 and 28, we found the greatest contributions in pass/fail recommendations between the two groups, with those using well-explicated metaphors receiving higher discourse and sociolinguistic competence ratings and correspondingly higher overall scores than those with insufficiently explicated metaphors. This finding suggests that the effective use of metaphors in teaching simulations may be especially helpful for intermediate students perhaps allowing them to compensate for other areas of weakness in their simulations. There was no apparent effect of metaphor use for students who had lower or higher scores. An outline of a metaphor-training curriculum component will be discussed.