Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

30/30 Well done Susan

EDSE 842, Spring 09: Final Exam

Answer each of the following questions, providing examples for each case.

2 1. How is reliability of surveys established? How is survey validity established? How is external validity of survey research established?

In survey research, reliability is established when a sufficient number of items is carefully constructed with language that is easily understood, has a precise format, and is crafted to collect information that will help answer the research questions. It is important to provide some questions that are easy to determine the answers and are more general while others are stronger sentiments and more difficult to answer. It is also important to change the wording of the statements to avoid pattern responses. For example, in Dawson, D'Alois, Rockoff, Reid, and Alston's 2006 study targeting high school students, one item began, "My laptop made it easier…" and another, "The laptops are more trouble …" Another consideration is that when all participants are presented with the same stimulus helping to offset the researcher's subjectivity in observations and increasing reliability. For reliability of the construct, use Cronbachs's Alpha. If it is less than .7 it indicates weak internal consistency.

Validity, however, is more difficult to obtain in survey research because actual feelings are difficult to pinpoint with terms available for survey responses. Not only is it difficult to capture feelings through terms such as "agree/disagree" and through scales one through five, or one through 10, such as those used in Cutler and Graham, (2008) but as the human mind ponders survey questions, myriad factors may influence the chosen answer, and it is very difficult to determine the accuracy of the responses. Internal validity is dependent upon the researcher's

questions and how closely the survey items are aligned. One survey had the goal to determine if the local school board's goals had been met by providing school issued laptops to all high school students in the district. One of the goals was to "prepare students to compete technically in school and work." Items such as, "How often do students in your classroom use their laptops to do the following?" (Dawson et al., 2006) followed by a list of 11 types of activities is likely to provide information being sought. Two important ways to establish external validity of a survey are to get a large number of participants to respond and to replicate the survey. A rigorous meta-analysis, such as Scruggs and Mastropieri (1996), identifying all similar surveys can help one determine whether or not external validity has been accomplished depending on the similarity or diversity of the findings. Good answer

2 2. Describe the problems of using multiple statistical tests on related data (e.g., individual survey items), and ways these problems can be addressed

When multiple statistical tests are performed on related data, chances of committing a Type I error are increased, attributing significance to a random finding. If the alpha level is changed from the usual .05 used with one test to a more stringent number such as .01 or .001 level, depending on the number of statistical tests performed, then it changes the probability of finding the random example of significance, but also increases the chances of committing a type II error, of not finding something that is really significant. Using such methods as the sequential Bonferroni method gives more statistical power with a family-wise error rate as opposed to a pair-wise error rate, which adjusts for that possibility. Although Cullinin, Osborne, and Epstein (2004) performed separate ANOVAs on seven variables, they adjusted the alpha level to .01 and .005 to be more conservative.

Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

2 3. Describe the limitations of survey research.

The strength of surveys is in the possibility of gathering large amounts of data, from a random subset of a great variety of participants in a population, in a time efficient manner, such as was found in Cutler and Graham (2008). Some of the limitations of survey research include;

- Because conditions are not manipulated, as in experimental research, cause and effect is more difficult to explain.
- Although a researcher may acquire randomized lists of prospective participants, it is highly unlikely that a large percentage will respond. It is difficult to randomize the participants who actually complete the survey. It is also impossible to assess the bias that may result from the specific participants who decide to respond and the missing data from those who choose not to participate.
- It is difficult to judge the quality, precision, and the honesty of the responses.

2 4. List and describe several special education research questions that are appropriate for survey research.

Research questions that are appropriate for the survey method are ones that would investigate opinions, beliefs, attitudes, and behaviors of teachers, parents, students, and the community. The survey method can also help with program evaluation.

1. “What are the attitudes of elementary principals toward the inclusion of students with severe/profound disabilities in the general education setting?” (Praisner, 2003, p. 135). Determining attitudes is one of the appropriate directions for a survey research question.

2. "What is the relationship between principals' perception of appropriate placements for students with different types of disabilities and their attitudes and experience? (Praisner, 2003, p. 135). Identifying relationships is also appropriate for survey research.
3. What are some classroom management, motivational, and organizational strategies that teachers have found to be helpful with the integration of technology and/or assistive technology (AT) in their classrooms?
4. What are the attitudes of middle school language arts teachers and special-education teachers towards allowing assistive technology to be used as a classroom tool for writing by students with writing deficits?

Questions three and four could help determine the current trends in attitudes of the use of AT with students who have diverse learning needs. Understanding these attitudes is key to developing an approach that will be more likely to address the teachers' concerns and succeed in motivating teachers to investigate the possibilities that AT can be beneficial to their students without "doing" the work for their students. Then they may be more likely to actively integrate AT into their classroom routines. The second question could provide teachers with descriptions of behaviors and strategies that could increase the potential of understanding and success when attempting to integrate technology.

2 5. Compare and contrast experimental and quasi-experimental research, and give examples of each from special education research.

Experimental research is a rigorous design that includes random assignment and compares the results of two or more equal groups when one or more interventions are

implemented. Each dependent variable is measured or observed only one time. Typical controls are pretests, matching participants, stratified randomization, and accounting for covariates. Using this design helps to control for such validity threats as history, maturation, regression, and selection bias.

The study of Sullivan, Mastropieri, and Scruggs (1995) is an example where students were stratified according to the student's grade level, then each was randomly assigned to a condition. Teachers in each of the conditions were to follow a script so that experiences within conditions would be the same to rule out as many other variables as possible. The students experienced the treatment in a one-to-one situation in an area away from noise. Sessions were recorded and each session lasted the same length of time. A post test and follow up test were given to compare results. Inter-rater reliability was assessed.

Because special education classes are usually in place, and because the classes are by nature not representative of the population, true randomization is not often possible. To adjust for the inability to randomize, quasi-experimental research designs may be employed. Although, as in the experimental design, each dependent variable is measured or observed only one time, and controls such as pretests, matching participants, stratified randomization, and accounting for covariates is also employed, without randomization and other controls, potential threats to validity in the areas of history, maturation, regression, selection bias, and mortality increase. Another obstacle with researching in the educational environment is the problem of withholding treatment in cases of dangerous behavior or in knowing that critical learning time may be missed. Rather than having a true control group, often an alternate treatment or combination of treatments can be used for a comparison group.

Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

Sometimes instead of randomized selection of individuals in the educational settings, classroom or teachers may be matched and randomly selected for one treatment or another. Saenz, Fuchs, and Fuchs (2005) is a good example of a quasi-experimental design. Students were chosen by using criteria and teachers were studied for statistical differences in 9 areas. The randomization was according to classes rather than by individual student. Implementation was carefully controlled for as many variables as possible.

Although the experimental design is more rigorous and is considered the best design to protect for internal validity, the quasi-experimental design is closer to a real situation and the most effective treatment can implemented more easily in other educational setting after the study concludes.

2 6. Describe inferential statistics, and provide examples.

With inferential statistics one takes data that has been collected to try to draw conclusions beyond the actual number themselves. For instance, when comparing performance gains scores from a group who has been in an experimental condition to the gain scores of a group that has been in a control or comparison condition, the researcher tries to figure out which condition is more effective. In this situation a t- test could be used to compare posttest results to infer the effectiveness of the condition as was done with pre-test and post-test scores in Glago's (in press) study of Self-Determination with elementary aged students.

An Analysis of Variance (ANOVA) could be used to show how scores vary from the mean. An Analysis of Covariance (ANCOVA) uses the scores but controls for some factor such as pretest score, age, or another factor that could account for the difference found in a post test scores, so that the conclusions drawn have taken those factors into consideration. For example, in

the study done by Saenz, Fuchs, and Fuchs (2005) using the CRAB scores to infer the comparability of students before the treatment was applied. Cullinin, Osborne, and Epstein (2004) also used and ANOVA to infer that students with ED have more difficulty learning than students without ED. Sullivan, Mastropieri, and Scruggs, (1995) used MANCOVA with the covariate of IQ. They also used a Student Newman-Keuls post-hoc ANOVA analysis to infer differences among the conditions.

2 7. Describe internal validity, external validity, and the problem of induction.

Internal validity asks if the observed changes are related to the factors that have been ascribed to them. Is there a causal relationship between the experimental condition and the results, or could some other factor have been influential. Are questions in a survey actually collecting the data the researcher thinks he/she is collecting? Has the treatment really influenced the change detected? Gersten, Baker, Smith-Johnson, Dimino, and Peterson (2006) used matched pairs and random assignment to increase internal validity.

External validity refers to determining if the changes that have happened during this research would likely happen again if the conditions were replicated. One way to enhance the external validity is to have a large sub sample of the population. Another way is to replicate the study numerous times and to compare results. If results are similar for a number of studies, then a case for external validity has been made. However, if the results vary widely, the external validity is suspect. Meta-analyses such as Scruggs, Mastropieri, and Regan (2006) and Scruggs, Mastropieri, and McDuffie (2007) can provide information needed to determine the degree of external validity for the included studies.

Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

Sometimes inductive reasoning can be faulty. Just because we can observe that something often happens at the same time as another occurrence, does not mean that there is a causal relationship. As was mentioned by Dr. Demetrov Dimitrov in EDUC 811, there is a positive correlation between shoe size and amount of general knowledge, however, shoe size cannot be used to predict general knowledge. When a child grows, both the amount of general knowledge and shoe size do increase. They both happen as a child grows until a certain age but there is no causal relationship. This is a problem of causal reasoning, not exclusively inductive

8. What are major threats to internal validity of group-experimental research, and how can these be addressed? How is external validity established?

Major threats to internal validity of group-experimental research are confounding variables, bias in the selection process, history, maturation, repeated testing, regression toward the mean, and attrition. These threats can be controlled by a variety of methods. Randomization of a large number of participants to the treatment condition and the control condition is one control for internal validity. That helps control for selection bias. Having a large group provides a more representative subpopulation. Maturation is addressed if the study is held over a shorter period of time so that it can't be said that the student knowledge has increased merely because they are older. Sullivan, Mastropieri, and Scruggs, (1995) saw every student one time, for a short period, with random assignment of condition. Although this procedure assured independence of observation and addresses many aspects of internal validity, such as history, maturation, and selection bias, external validity may present a problem. Classrooms are not set up in a similar way so this procedure would not be feasible.

Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

With group-experimental research, external validity is established with a large random representative sample, and/or with replication and through a careful analysis of similar studies such as found in the meta-analysis of Scruggs, Mastropieri, Berkeley, and Graetz, (in press).

1.5 9. Describe the assumptions of the analysis of variance, and list the two most significant assumptions. What can be done when data do not meet these assumptions?

The two most significant assumptions of the analysis of variance are homogeneity of variance and normality of the sampling distribution. Special education classes violate the normality of sampling because by definition students in special education usually score at the bottom of the possibilities and can often vary widely which violates the other assumption! Student special education populations are also much smaller than populations of typical students. To deal with these issues we can adjust the unit of analysis to classes or teachers to help limit the extreme variability. Non parametric tests can be used in the analysis. Independence – more important than normality.

2 10. What are ceiling and floor effects, and why may these lead to invalid conclusions from inferential statistical tests? Why may this be a particular problem in special education research? How can such data be analyzed appropriately?

Ceiling and floor effects are when many or all participants get the highest score possible or the lowest score possible. In a pretest many students may get no answers correct and after the treatment they may earn all possible points. Inferential statistical tests assume that the scores should follow the normal bell curve. This assumption is violated when the participants all score the same. Often in special education students will not have the typical knowledge that their peers

have. When an intervention targeting that knowledge is applied, it might help all students grasp the targeted information.

One way to deal with this is in the design stage to ensure that there will be a way for participants' scores to vary. Glago (in press) used non-parametric statistics, Mann Whitney *U* test to deal with the floor effect with pretest scores and both floor and ceiling effect with posttest results.

2 11. Describe "unit of analysis" and the importance of its consideration in special education research.

The unit of analysis is the smallest identified subject unit to be studied. It can vary from the individual student, to a pair of students who are matched or work together, to small groups, to teachers, to classrooms, to schools, etc.

Sullivan, Mastropieri, and Scruggs, (1995) used the indvidual student as the unit of analysis which insures internal validity. In the study of Sáenz, Fuchs, and Fuchs, (2005), the chosen unit of analysis was each of 12 teachers. In the second study the number became smaller but that choice controlled for widely varied scores that can occur when using individual students. In special education there are often students whose scores very tremendously. When those scores are put in with the others to determine the mean, it can skew the results. For instance if four scores were 32, 31, 30, and 1, the mean would be 23.5 which does not fairly represent the typical scores. The larger the number of participants used to find the mean, the less variation, the smaller the standard deviation, and the more representative the means will be.

Susan Kenney
May 12, 2009

EDSE 842
Dr. Thomas Scruggs

2 12. What are ways of dealing with interventions in intact groups such as special education classrooms?

Some of the ways to deal with intact groups is to randomize teachers and/ or classrooms, use several sets of two classes taught by the same teacher, one with the treatment and one with the comparison condition, find matching groups in the school or at other schools, use pretest as a covariate, calculate and analyze the gain score, use classrooms treated as a nested factor such as was done in the study done by Mastropieri, Scruggs, Spencer, and Fontana, (2003).

2 13. What problems are associated with the use of pre-existing groups in quasi-experimental research, and how can these problems be addressed?

Some of the problems of using pre-existing groups such as boys verses versus girls, older students verses younger students, students with disabilities verses typical students, are that randomization, which is key in a strong experimental design, is not possible, and the groups may vary considerably. A quasi-experimental design may be used to study the groups mentioned by comparing students who are pre-matched by a similar category. It is important to analyze within subjects to get a good idea of how a treatment affects the students as compared to themselves. With these groups, perhaps a different research design such as single subject might be considered.

2 14. Describe and provide an example of an interaction effect in a two-way factorial design, and how it might be interpreted.

A two-way factorial design is when several variations are compared to determine which combination is more effective. Simpkins, Mastropieri, and Scruggs, (in press) entered gain scores into a two by two ANOVA with repeated measures. The groups were normally achieving

students and students with learning disabilities or students at risk. The conditions were the experimental condition with and the control condition. The normally achieving students were exposed to the experimental treatment and the control condition. The students with learning disabilities and students at risk were also exposed to the experimental treatment and the control condition. The results of each of those four situations were compared. The findings were that all students did better when they used the differentiated materials but there was no interaction between the group and condition.

	Normally achieving	Students at risk and LD
Experimental	Results E N 1	Results LDR 1
Control	Results E N 2	Results LDR 2

Both EN1 and LDR 1 did better than EN2 and LDR 2. Both groups responded in a similar way to the treatment condition.

McDuffie (1974??) also did a factorial design with 2 conditions- experimental and traditional by 2 settings- co-taught and non co-taught by 2 groups- general ed and special ed. Eight classes participated in the study with two classes that experienced each condition. Results revealed no statistically significant interaction (although there might have been).

2 15. List special education research questions appropriate to group-experimental and quasi-experimental research.

The following questions are appropriate to be used for group-experimental and quasi-experimental research.

1. Can "students with learning disabilities… benefit from coaching to facilitate immediate and delayed recall of a target and related explanation information"? (Sullivan, Mastropieri, & Scruggs, 1995, p. 312)
2. Can "coaching students to construct explanations … be more effective than providing the same explanations"? (Sullivan, Mastropieri, & Scruggs, 1995, p. 312)
3. "what are the effects of PALs on the reading performance of ELL with LD?" (Sáenz, Fuchs, & Fuchs, 2005 p. 234)
4. "What are the incidental benefits of PALs to ELL with low, average, and high achievement in reading who participate in PALs?" (Sáenz, Fuchs, & Fuchs, 2005 p. 234)

Those questions try to determine the degree of the relative effectiveness of treatments applied to students.

Resources

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100,* 907-919.

Dawson, M. E., D'Alois, L. A., Rockoff, D., Reid, B., & Alston, R. (2006). Formative evaluation of the technology integration in 2004-05. (Available from Alexandria City Public Schools, Monitoring and Evaluation Services. 2000 N. Beauregard, Alexandria, VA 22311).

Gersten, R., Baker, S.K., Smith-Johnson, J., Dimino, J., & Peterson, A. (2006). Eyes on the prize: Teaching complex historical content to middle school students with learning disabilities. *Exceptional Children, 72,* 264-280.

Glago, K., Mastropieri, M.A., & Scruggs, T.E. (in press). Improving problem solving of elementary students with mild disabilities. *Remedial and Special Education.*

Mastropieri, M.A, Scruggs, T.E, Spencer, V., & Fontana, J. (2003). Promoting success in high school world history: Peer tutoring versus guided notes. *Learning Disabilities Research & Practice, 18,* 52-65.

McDuffie, K. (1974)Promoting success in content area classes: Is value added through co-teaching? (Doctoral dissertation, George Mason University, 1974). 2006,mcdu.

Praisner, C.L. (2003). Attitudes of elementary principals toward the inclusion of students with disabilities. *Exceptional Children, 69,* 135-145.

Sáenz, L.M., Fuchs, L.S., & Fuchs, D. (2005). Peer-assisted learning strategies for English language learners with learning disabilities. *Exceptional Children, 71,* 231-247.

Scruggs, T.E., & Mastropieri, M.A. (1996). Teacher perceptions of mainstreaming/ inclusion, 1958-1995: A research synthesis. *Exceptional Children, 63*, 59-74.

Scruggs, T.E., Mastropieri, M.A., Berkeley, S., & Graetz, J. E., (in press). Do special education interventions improve learning of secondary content? *Remedial and Special Education*.

Scruggs, T.E., & Mastropieri, M.A. (1996). Teacher perceptions of mainstreaming/ inclusion, 1958-1995: A research synthesis. *Exceptional Children, 63*, 59-74.

Scruggs, T.E., Mastropieri, M.A., & McDuffie, K.A. (2007). Co-teaching in inclusive classrooms: A meta-synthesis of qualitative research. *Exceptional Children, 73*, 392-416.

Scruggs, T.E., Mastropieri, M.A., & Regan, K. (2006). Statistical analysis for single subject research designs. In T.E. Scruggs & M.A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities: Vol. 19. Applications of research methodology* (pp. 33-54). Oxford, UK: Elsevier.

Simpkins, P.M., Mastropieri, M.A., & Scruggs, T.E. (in press). Differentiated curriculum enhancements in inclusive 5th grade science classes. *Remedial and Special Education.* Reasoning and remembering: Coaching thinking with students with learning disabilities. *Journal of Special Education, 29*, 310-322.

Sullivan, G.S., Mastropieri, M.A., & Scruggs, T.E. (1995). Reasoning and remembering: Coaching thinking with students with learning disabilities. *Journal of Special Education, 29*, 310-322.