The Value of Research

Spring Research Conference
University of Louisville
University of Kentucky
University of Cincinnati
March 31, 2012

Christopher Rakes, Ph.D.
Mathematics Education
Associate Research Scientist
Institute of Education Sciences
U.S. Department of Education

Email: christopher.rakes@gmail.com
Web: http://csrakes.yolasite.com

Introduction:
My background is in mathematics education. I was a high school mathematics teacher for 8 years and taught an additional 2 years at the college level. I am currently serving as an associate research scientist at the Institute of Education Sciences within the U.S. Department of Education.

My position involves two primary tasks. The first is editing and reviewing reports released by IES. The second is coordinating and managing grant review panels. In both of these tasks, the value of research plays a critical role.

The Challenge

I believe that only scientists can understand the universe. It is not so much that I have confidence in scientists being right, but that I have so much in nonscientists being wrong.

I believe in evidence. I believe in observation, measurement, and reasoning, confirmed by independent observers. I'll believe anything, no matter how wild and ridiculous, if there is evidence for it. The wilder and more ridiculous something is, however, the firmer and more solid the evidence will have to be. ~Isaac Asimov (1920-1992)

"The prevailing view is that findings from education research studies are of low quality and are endlessly contested—the result of which is that no consensus emerges about anything" (Shavelson & Towne, 2002, p. 28)

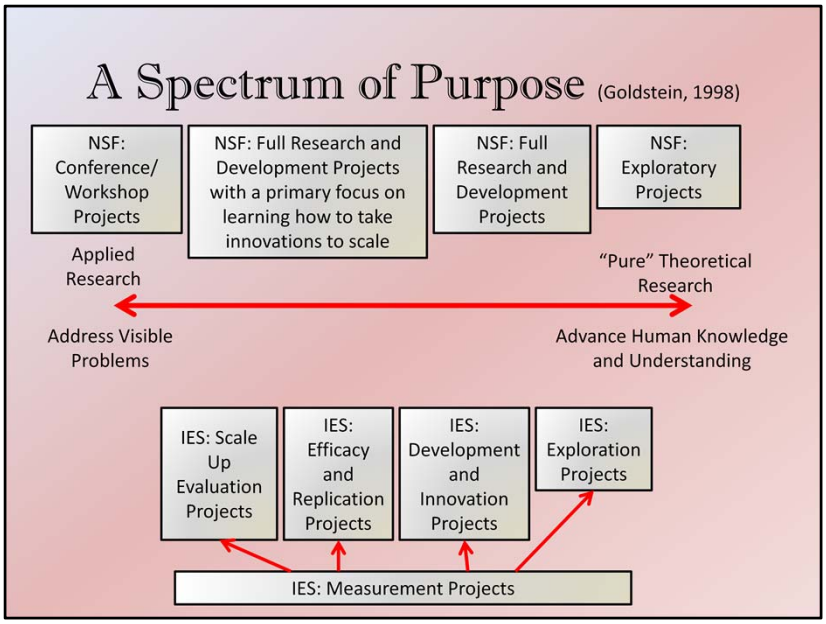
Why did Isaac Asimov have such a radically different consideration of research than the education community? Consider the example of calculator use in mathematics education.

Forty years of calculator research have reached consensus: the worst case scenario resulting from calculator usage for both instruction and assessment is that students are no better or worse at computation, but they are better problem solvers in unfamiliar situations (Ronau et al., 2011).

In contrast, more than forty years of tradition have produced a general attitude that calculators will harm students' computational skill development. Recent research has shown how particular strategies for using calculators as instructional tools can actually improve students' computational skills.

When such a discrepancy arises, research is often presumed to be biased ("well OF COURSE a traditionalist found that calculators harm;" "well OF COURSE a progressive found that calculators did no harm. What else would they find?"). Practitioners rely on other bodies of literature that has few (if any) connections to research. The result is that rather than having a seamless transition from research to practice, more time and money is spent spinning our wheels: arguing positions based on intuition, experience (and perhaps some research support), developing policies (with little or no support), and entrenching our profession with practices that have no evidentiary support.

RNR: We can and must do better. But to do better, we must, as researchers, attend to the specific tools and structures of research and connect that research to practice; and as practitioners, become better consumers and users of research to support our decisions about our practice.



The tension between research and practice implies a dichotomy; the situation is more like a continuum. It is often TREATED like a dichotomy; but that is due to an inadequacy of our understanding of the nature of research and practice, not due to the research process itself.

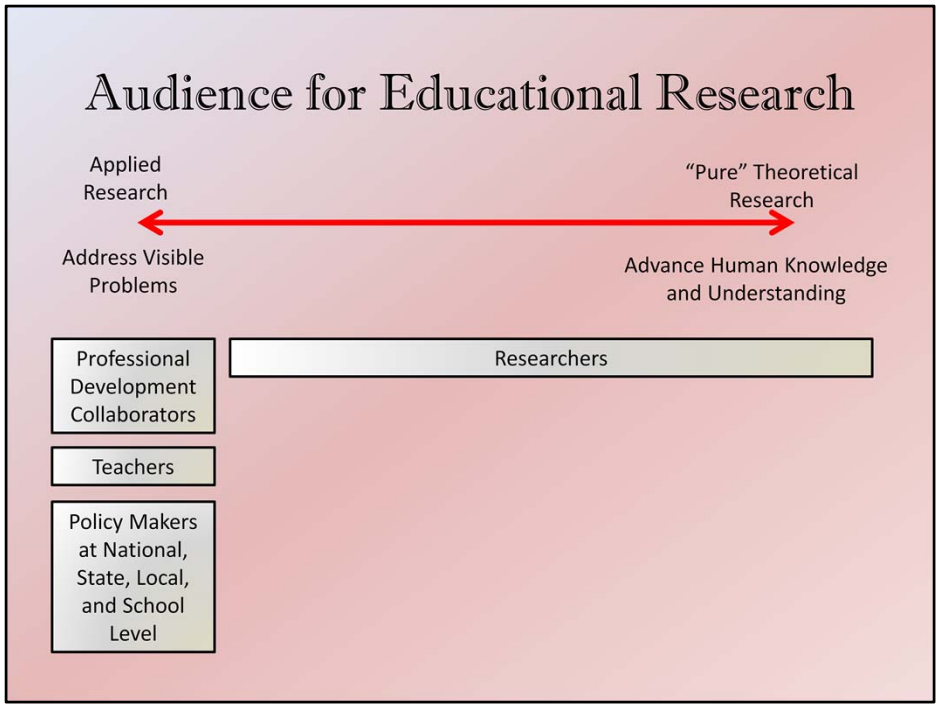
There is a general tension in education between applied research and "pure" research.

Applied research tends to focus on addressing visible problems while "pure" or theoretical research is more concerned with advancing human knowledge and understanding. The practical benefits of theoretical research may be apparent either immediately or over time, but often the benefit is not easily applied to a practical scenario.

The National Science Foundation and the Institute of Education Sciences are the two major federal government agencies that fund educational research. Both agencies value research across the spectrum of purposes. Notice that the NSF goal structure currently supports projects at the more extreme ends of the spectrum while the IES goal structure concentrates more toward the middle chunk of the spectrum.

RNR: Note-A particular line of research would move from right to left in these spectrums

IES and NSF are currently working on the development of a common goal structure that would apply to both agencies.



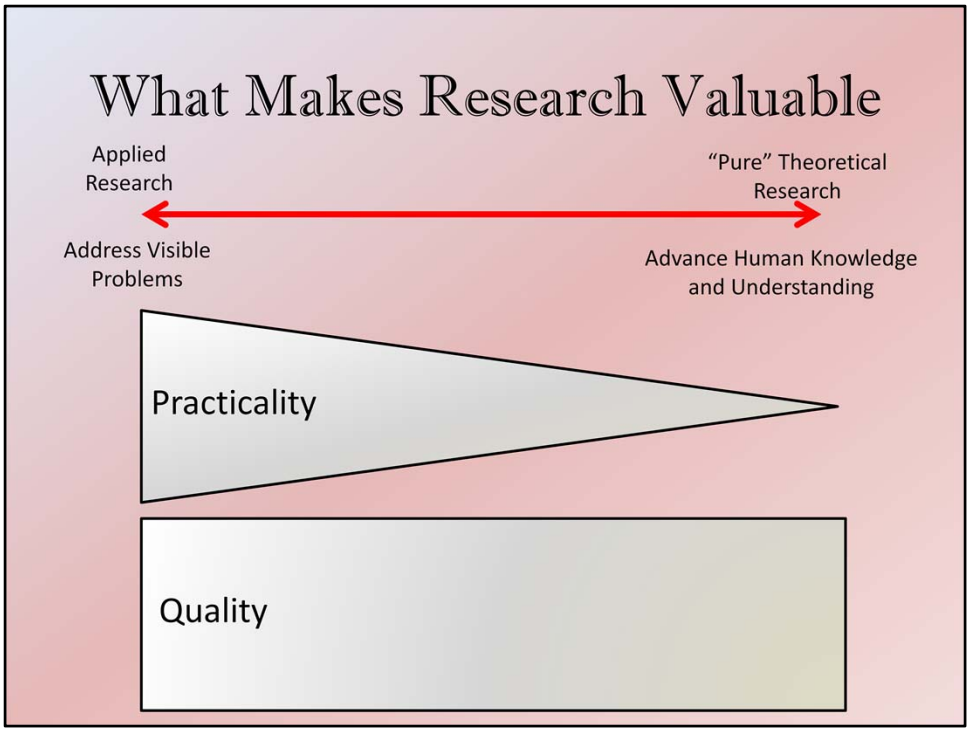
The audience for research across the spectrum differs.

Instructional Coaches, Academic Deans, Principals, Content Specialists, Mentor Teachers, and other professional development providers may focus on how research applies to their specific area of professional development.

Teachers may focus on how strategies that have been shown to have promise can be applied to their classrooms.

Policy makers may rely on research and data to guide systemic reforms and the development of appropriate support structures.

Researchers may also be concerned with the applicability of research, but their interests should not be limited to applied research. The advancement of the field occurs only when research is well-coordinated, founded in scientific principles, and applied appropriately.



What makes research valuable changes depending on where the purpose falls on this spectrum.

As purposes fall more to the theoretical end of the spectrum, the importance of practicality decreases.

On the other hand, regardless of where the purpose falls on the spectrum, quality remains critical to the value of the research.

Six Scientific Principles (Shavelson & Towne, 2002)

1. Pose Significant Questions That Can Be Investigated Empirically
2. Link Research to Relevant Theory
3. Use Methods That Permit Direct Investigation of the Question
4. Provide a Coherent and Explicit Chain of Reasoning
5. Replicate and Generalize Across Studies
6. Disclose Research to Encourage Professional Scrutiny and Critique

The National Research Council proposed 6 principles that need to be followed for research to be considered of scientific quality.

These six principles (posing significant questions, linking research to theory, using appropriate methods, providing a logical chain of reasoning, replicating findings, and submitting for peer review) provide a useful framework for considering the components needed to produce high quality research.

For each principle, Bob Ronau and I have developed a set of guiding questions to support self-assessment for researchers based on three systematic reviews of research quality on a variety of educational topics (e.g., educational technology in mathematics, language arts, social studies, science, CTE, teacher knowledge).

Principle 1: Pose Significant Questions That Can Be Investigated Empirically

1. Does the introductory narrative clearly and succinctly identify the **purpose** of the study?
2. Can readers **distinguish the research questions** from the introductory narrative?

Ronau & Rakes (2011)

For research to be clear and easily understandable, authors should explicitly tell readers what the study intended to accomplish generally and what research questions guided the study.

In some cases, research questions are considered the “end point” rather than the beginning point. In such research, hypotheses should be stated, and the research questions should still be explicitly given in the discussion.

Principle 2:
Link Research to Relevant Theory

1. Does the manuscript identify a guiding **theoretical framework** for the study?
2. Does the manuscript explicitly state how the **guiding theoretical framework** informed the methodology, analysis, and interpretation of the study?
3. Does the literature review **include explicit connections** between the research questions and purpose of the study with the chosen methodology?
4. Does the literature review **provide an argument for the study**; that is, it explicitly makes a case for present study, clearly justifies the conceptual framework used to guide the study, demonstrates how the study builds from previous research, and shares how the study contributes to a need in the current research foundation?
5. Are the connections to prior research **re-visited in the discussion** of results?

Ronau & Rakes (2011)

Linking a study to a framework adds a stronger foundation for situating the study to the relevant body of research. The framework and associated literature should be addressed in both the introduction and discussion to fully connect the study to prior research.

Principle 3: Use Methods That Permit Direct Investigation of the Question

1. Is the **research design** stated explicitly?
2. Does the manuscript provide a **rationale for the type of research** being conducted (e.g., quantitative, qualitative, mixed methods)?
3. Are the **population** specifics, **sampling** techniques, characteristics of the **sample**, and **grouping** assignment techniques stated explicitly?
4. Are **threats to validity** from the sample addressed explicitly?
5. Does the manuscript explicitly discuss how the research methodology **balances relevant threats to validity**?

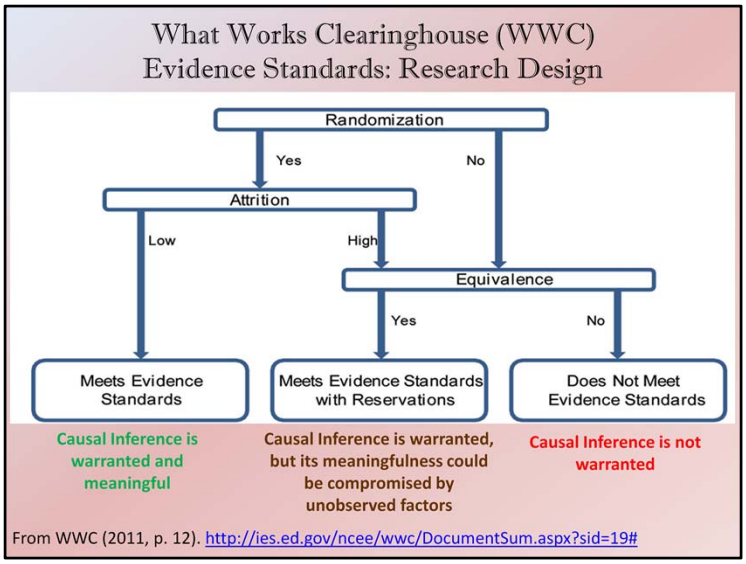
For more information:

Creswell, J. W. (2007). *Qualitative inquiry and research design* (2nd ed.).
Patton, M. W. (2002). *Qualitative research and evaluation methods* (3rd ed.).
Shadish, Cook, & Campbell (2002). *Experimental and quasi-experimental design*.
Teddlie, & Tashakkori (2009). *Foundations of mixed methods research*.

Ronau & Rakes (2011)

The choice of research design should be linked closely to the purpose:
Do the study questions seek to describe a phenomena, establish an association, or establish impact of an intervention?

- Descriptive studies may use qualitative methods.
- Studies seeking to establish an association (i.e., correlation) may be
  - Multiple group comparisons ("Some research may show that students participating in a given program tend to have **better outcomes**, such as higher math achievement scores or lower dropout rates," WWC website),
  - but they may also compare two or more variables within a **single group** (e.g., students in a given program improved during the period of intervention; students' scores in math and reading were highly correlated).
- Studies seeking to establish impact of an intervention (i.e., causation; "to be evidence of **effectiveness**, the research has to show that the program *caused* the improvement," WWC website) need to consider randomization, attrition, and equivalence issues before declaring causal impact.



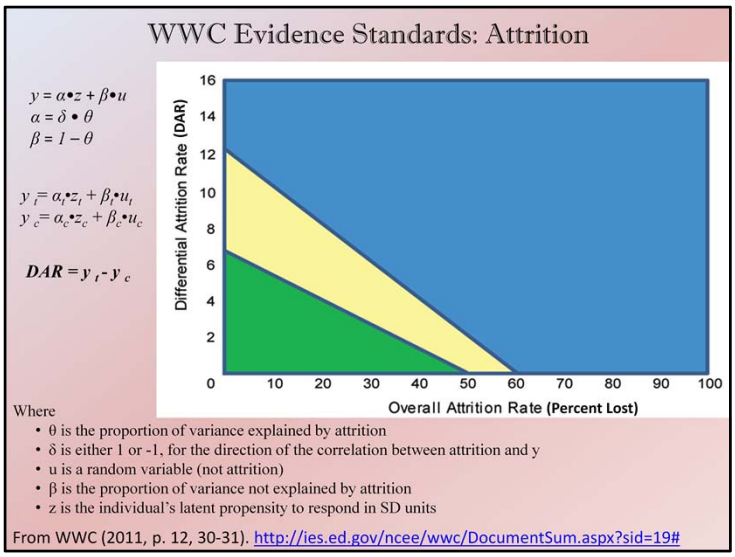
Research Design

"In an RCT, researchers use random assignment to form two groups of study participants…Any movement or nonrandom placement of students, teachers, classrooms, or schools after random assignment jeopardizes the random assignment design of the study. "

If a study randomizes the process of assigning subjects to the treatment groups, and overall attrition is low, equivalence of groups at the outset of the study is assumed and causal inferences can be made.

If a study randomizes the treatment groups, but overall attrition is high, the study must demonstrate the treatment and control groups are alike (addresses the attrition threat to internal validity). If a study does not randomize the treatment groups, it must demonstrate that the treatment and control groups are alike, regardless of attrition (addresses the selection threat to internal validity). Studies in these categories that demonstrate group equivalence are considered to have set up conditions in which causal inference appears to be warranted, but the possibility remains that unobserved or unmeasured characteristics could influence the outcome.

If a study does not demonstrate equivalent groups in one of these ways, causal inference is not warranted.



Attrition

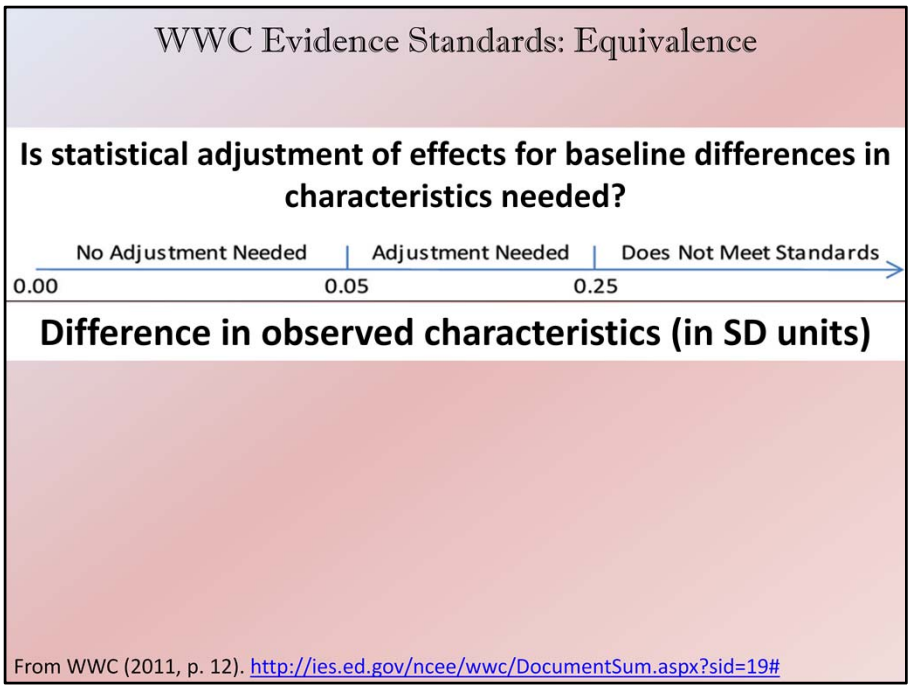
"Randomization, in principle, should result in similar groups, but attrition from these groups may create dissimilarities. **Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups**. The WWC is concerned about overall attrition as well as differences in the rates of attrition for the intervention and comparison groups. **If there are high levels of attrition, the initial equivalence of the intervention and comparison groups may be compromised and the effect size estimates may be biased**."

In this graph, the overall attrition rate is the percentage of respondents lost overall.

The differential attrition rate, however, is a more complicated scale that takes into account the relationship between an outcome variable and the treatment/control group attrition rate.

- **Start by computing an individual's propensity to respond (z).**
- Consider an outcome variable (y) that is composed of attrition effects (α•z) and non-attrition effects (β•u)
- **Compute the outcome for the treatment group (y_t) and control group (y_c).**
- The difference is the DAR.

"Within the yellow/middle region of the figure, the potential bias depends on the assumptions of the model (WWC, 2011, p. 13)"



Equivalence

"The WWC requires that RCTs with high levels of attrition and all QEDs present evidence that the intervention and comparison groups are alike. Demonstrating equivalence minimizes potential bias from attrition (RCTs) or selection (QEDs) that can alter effect size estimates.

Baseline equivalence of the analytical sample must be demonstrated on observed characteristics defined in the topic area protocol, using these criteria:

- The reported difference of the characteristics must be less than 0.25 of a standard deviation (based on the variation of that characteristic in the pooled sample).
- In addition, the effects must be statistically adjusted for baseline differences in the characteristics if the difference is greater than 0.05 of a standard deviation" (WWC, 2011, p. 14)

Principle 3: Use Methods That Permit Direct Investigation of the Question (cont'd.)

1. Do the chosen data analysis techniques **align** with the research questions and purpose of the study?
2. Does the **discussion** of the **results** clearly **address the research questions**?
3. Does the data analysis **directly support all conclusions** made in the discussion?

For more information:

Standards for reporting on empirical social science research in AERA Publications. (2006). *Educational Researcher, 35*, 33-40. Retrieved from http://www.sagepub.com/upm-data/13127_Standards_from_AERA.pdf

Ronau & Rakes (2011)

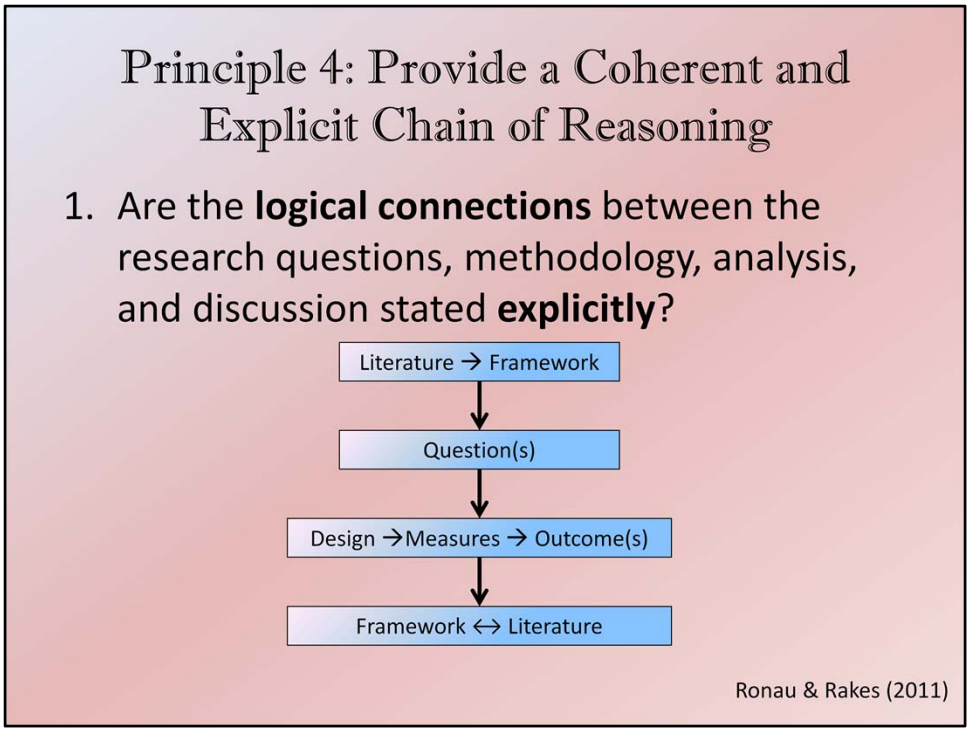
"Need a clear statement of the purpose and scope of the study.

Make clear how the study is a ***contribution to knowledge.***

Review relevant scholarship that bears directly on the topic.

The rationale for the ***conceptual, methodological, or theoretical orientation*** of the study should be described and explained with relevant citations to what others have written about it.

A rationale should be provided for the ***problem formulation as it relates to the groups studied*** (especially with respect to relevant features of the historical, linguistic, social, and cultural backgrounds of the group) where questions about appropriateness of the connections may arise" (AERA, 2006, p. 34).



"The design and logic of a study flows directly from the problem formulation. It is shaped by the intellectual tradition(s) in which the authors are working and the ways in which they view the phenomenon under study. This in turn influences the identification of questions, the choice of methods of data collection, the approach to analysis and interpretation, and the format of reporting. These decisions constitute the logic of inquiry that researchers report." (AERA, 2006, p. 34)

Principle 5: Replicate and Generalize Across Studies

1. Are the **necessary data reported** to compare future replication studies with current results (e.g., means, standard deviations, sample sizes, effect sizes)?
2. Does the manuscript describe the **measures** used, the reported **validity** and **reliability** from previous studies (if applicable), and **validity** and **reliability** statistics from the current sample?

For more information:

Urbina, S. (2004). *Essentials of psychological testing*.

Ronau & Rakes (2011)

Replication with variety of settings or conditions and/or different types of students is important for determining how broadly inferences can appropriately be interpreted.

Principle 6: Disclose Research to Encourage Professional Scrutiny and Critique

1. Has the study been presented at one or more peer-reviewed **conferences**?
2. Has the study been submitted to a **peer-reviewed** book, journal, or other publishing agency?

Ronau & Rakes (2011)

I have found conference presentations to be quite helpful for (a) publicizing my research, (b) identifying other scholars interested in my areas of interest, and (c) getting feedback on the study prior to submission to a journal for publication.

I have found that the revisions based on peer review have always helped my own conceptualizations of the study; help point out blind spots; catch errors that I overlooked; point out where something is unclear to readers.

BUT, peer reviewed does not equal high quality.
Editor may not have had high quality articles to choose from
Well-known issue of publication bias (Rosenthal, 1979). "Careful sampling and identification of the fugitive literature are the best protection against a publication bias" (Lipsey & Wilson, 2001, p. 166).

A Measure of Quality (Rakes et al., 2012)

¶	Mixed Methods (up to 16 pts)	
Non-Research (up to 5 pts)	**Quantitative (up to 15 pts)**	**Qualitative (up to 11 pts)**
Theoretical Connections (up to 4 pts) • Literature Support (up to 2 pts) ➢ Well Grounded (2 pts) ➢ Partially Grounded (1 pt) ➢ Not Grounded (0 pts) • Framework/Theory Connections (up to 2 pts) ➢ Well Connected (2 pts) ➢ Partially Connected (1 pt) ➢ Not Connected (0 pts) **Design Clarity and Validity (up to 1 pt)** • Purpose Statement (1 pt)	**Theoretical Connections (up to 4 pts)** • Literature Support (up to 2 pts) ➢ Well Grounded (2 pts) ➢ Partially Grounded (1 pt) ➢ Not Grounded (0 pts) • Framework/Theory Connections (up to 2 pts) ➢ Well Connected (2 pts) ➢ Partially Connected (1 pt) ➢ Not Connected (0 pts) **Design Clarity and Validity (up to 9 pts)** • Purpose Statement (1 pt) • Research Questions/Hypotheses (1 pt) • Design Robustness (up to 3 pts) ➢ Randomized Experiment (2 pts) ➢ Regression Discontinuity Design (2 pts) ➢ Quasi-Experimental Design with: • Sampling Strategies Unclear (1 pt) • Convenience Sample (1 pt) • Other Sampling Strategies (2 pts) ➢ Use of Control Group (1 pt) • Threats to Validity Addressed (up to 4 pts) ➢ Internal (1 pt) ➢ External (1 pt) ➢ Construct (1 pt) ➢ Statistical Conclusion (1 pt) **Measurement Trustworthiness (up to 2 pts)** • Reliability (1 point) ➢ Internal Consistency ➢ Split-Half ➢ Test-Retest ➢ Inter-Rater ➢ Alternate Forms • Validity (1 point) ➢ Content → Construct ➢ Concurrent Criterion → Discriminant ➢ Predictive Criterion → Convergent	**Theoretical Connections (up to 4 pts)** • Literature Support (up to 2 pts) ➢ Well Grounded (2 pts) ➢ Partially Grounded (1 pt) ➢ Not Grounded (0 pts) • Framework/Theory Connections (up to 2 pts) ➢ Well Connected (2 pts) ➢ Partially Connected (1 pt) ➢ Not Connected (0 pts) **Design Clarity and Validity (up to 5 pts)** • Purpose Statement (1 pt) • Research Questions/Hypotheses (1 pt) • Threats to Validity Addressed (up to 3 pts) ➢ Internal (1 pt) ➢ External (1 pt) ➢ Construct (1 pt) **Measurement Trustworthiness (up to 2 pts)** • Reliability (1 point) ➢ Internal Consistency ➢ Inter-Rater • Validity (1 point) ➢ Persistent Observation ➢ Triangulation ➢ Peer Debriefing ➢ Negative Case Analysis ➢ Referential Adequacy ➢ Member Checks ➢ Thick Description ➢ Dependability Audit ➢ Confirmability Audit ➢ Reflective Journal

36-55% of Points Possible Earned in Mathematics Educational Technology Research

In a recent study, my team and I developed a scale for measuring quality, by type of research. You will notice that the criteria here are based on the criteria we've been discussing throughout this presentation.

We applied this scale to a systematic review of literature in mathematics educational technology, and found that connections to frameworks were consistently lacking.

Framework points (average/points possible)

Practitioner-Oriented: 0.24/2 = 12% of points possible
Quantitative: 0.66/2 = 33% of points possible
Mixed Methods: 1.08/2 = 54% of points possible
Qualitative: 1.28/2 = 64% of points possible

Research: The Foundation of the Field

- The quality of research is critical to establishing the credibility of the knowledge held and practices used within a field.
- Frameworks and resources are available to help researchers produce high quality findings.

The importance of research quality ties directly back to you as doctoral students.

As you prepare or complete your dissertations, consider the goal of the research.
What type of questions are you asking?

Have you designed a study that is capable of answering these questions in a valid, reliable manner? Think about both your general approach and your analytic techniques.

Are you attempting to make causal claims? Does your design support such claims? If not, can the design be adjusted. If it can't be adjusted, then the goal of making statements about "impact" should be reconsidered. If your study design is correlational, avoid words such as "impact."

Produce Research that can be Valued

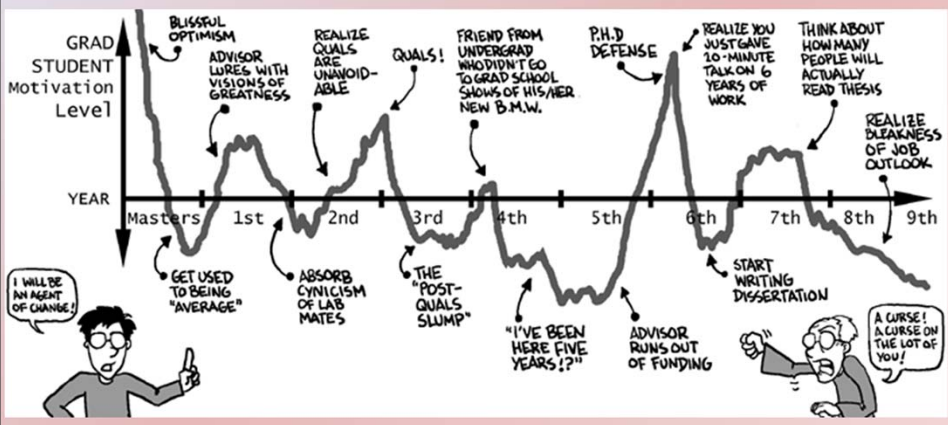
- Identify strengths – What activities can you do to support continued growth in these areas?
- Identify weaknesses – what barriers prevent you from producing high quality research?
- Remove Roadblocks
  - Additional Coursework (Do more than the minimum requirements for research methodology).
  - Independent studies (on or off your program sheet).
  - Volunteer to help with your mentors' research.
  - Consider helping to teach basic methodology classes.
  - Training Workshops

While one of your goals is to get a degree, a more important goal is to get a high-quality job, and even more important is making a positive impact on your field with your career. Shooting for a degree will not necessarily lead you to have a career with an impact; but shooting for the career will lead you to the degree. Pack your tool belt as full as you can to conduct the high quality research needed for an impactful career.

1. Understanding the basic statistical tools available such as Factor Analysis, HLM, SEM, and IRT are a baseline.
2. Understand how research design allows or obstructs valid causal inferences regardless of the statistical analyses being run.
3. Do not graduate limited to "only quantitative" or "only qualitative." Position yourself to be able to let the research questions dictate the design and methodology.
   a) Additional Coursework (Do more than the minimum requirements for research methodology).
   b) Independent studies (on or off your program sheet).
   c) Volunteer to help with your mentors' research.
   d) Consider helping to teach basic methodology classes.
   e) Training Workshops (Conferences, especially AERA; IES conducts Summer Research Training Institutes: This summer, the two workshops are on Single Subject Designs (6/25) and Randomized Control Trials (7/15).

4. Read beyond your classes
5. Look up and read key citations in your readings
6. Develop a database of key articles and manuscripts in one or two key areas of study (the actual articles since so much is electronic)
7. Attend faculty job presentations; read their vitae; understand the market
8. Study, research, and write in teams.

Thank You!



Email: christopher.rakes@gmail.com
Web: http://csrakes.yolasite.com

References

Goldstein, H. (1998). The importance of educational research. *Improving Schools, 1*, 33-34.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Rakes, C. R., Ronau, R. N., Bush, S. B.[1], Driskell, S. O.[1], Pugalee, D.[1] & Niess, M. L. (2012, February). *A structured inquiry of research in mathematics educational technology: Findings and implications*. Symposium presented at the annual meeting of the Association of Mathematics Teacher Educators, Fort Worth, TX.

Ronau, R. N., & Rakes, C. R. (2011). Making the grade: Reporting educational technology and teacher knowledge research. In R. N. Ronau, C. R. Rakes, & M. L. Niess (Eds.), *Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches* (pp. 323-332). Hershey, PA: IGI Global. DOI: 10.4018/978-1-60960-750-0.ch014

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638–641.

Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academies Press.

Standards for reporting on empirical social science research in AERA Publications. (2006). *Educational Researcher, 35*, 33-40. Retrieved from http://www.sagepub.com/upm-data/13127_Standards_from_AERA.pdf

What Works Clearinghouse Procedures and standards handbook. (2011). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. Retrieved from http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19

[1] Bush, Driskell, and Pugalee contributed equally and are presented in alphabetical order.