

**THE PREPARATION OF ASPIRING EDUCATIONAL RESEARCHERS IN THE  
EMPIRICAL QUALITATIVE AND QUANTITATIVE TRADITIONS  
OF SOCIAL SCIENCE**

**Methodological Rigor, Social and Theoretical Relevance, and More**

**Report of a Task Force of the  
Spencer Foundation Educational Research Training Grant Institutions**

**September, 2009**

**Grant #2008001136**

I take a definite philosophical and methodological stand: I argue for objective and relevant fact-finding, rigorous theorizing, and empirical testing, as well as for morally sensitive and socially responsible policy design. Consequently I argue against irrationalism and subjectivism ... as well as against opaque rhetoric passed off as theory and partisan sloganeering passed off as serious social-policy making.... I argue that, if the burning social issues of our time are to be tackled effectively and fairly, they must be approached in the light of serious social research together with moral principles....

Mario Bunge, *Social Science Under Debate* (1998), p. xi

The canons of scientific rationality can guide our study of the social realm. Of course, actual social research does and will continue to employ methods found nowhere in the natural sciences and vice versa. But in concrete research the different natural sciences employ diverse methods as well. Behind these different practices lies a commitment to certain basic scientific virtues. Naturalists believe those virtues can also undergird the methods of social research.... Thus the naturalism I advocate holds that some social science is good science. Note, however, that naturalism does not imply that all is well in social research.... Large parts of the social sciences do fail to produce good science. Still, that situation can be and sometimes is overcome.

Harold Kincaid, *Philosophical Foundations of the Social Sciences* (1996), pp. 3-4

What is to be resisted is the notion that the cultivation of methodology is either necessary or sufficient for successful scientific endeavor.... [But] nothing that I have said about the relation of methodology to science is intended to undermine the normative force and function of methodology.

Abraham Kaplan, *The Conduct of Inquiry* (1964), pp. 24-25

## Contents

Task Force (TF) membership

Preface

Improving research training: Complexities and constraints

The experience of the RTG institutions

The Task Force at work

Charge to the Task Force

REPORT: THE PREPARATION OF ASPIRING EDUCATIONAL RESEARCHERS

1. Introduction: Adopting a new perspective
2. The multiple universes in the life of the educational researcher
  - 2.1 Description of the four universes
  - 2.2 How the notion of “universes” brings the training issues into focus
3. Consequences of leaving a solitary and “closeted” universe
4. The importance of contextual knowledge
5. Loosening the identification of quality solely with methodological rigor
6. Introduction to the vignettes
7. Research training: Two vignettes
  - 7.1 The emergence of a researcher: Denise
  - 7.2 The emergence of a researcher: Catherine
8. The programs of Denise and Catherine assessed from the perspective of the multiple universes
  - 8.1 Analysis of Denise’s training
  - 8.2 Analysis of Catherine’s training
9. Knowledge, skills, dispositions, and practices necessary for advising
  - 9.1 Overview: The vital roles of the advisor
  - 9.2 Examples of advising practices related to the four Universes
  - 9.3 A model of proactive and integrative advising
10. Summary, and further elucidations of, the TF recommendations
  - 10.1 Recommendations concerning the initiation of ARs into Universe 1
  - 10.2 Recommendations concerning the initiation of ARs into Universe 2
  - 10.3 Recommendations concerning the initiation of ARs into Universe 3

- 10.4 Recommendations concerning the initiation of ARs into Universe 4
- 10.5 Recommendations concerning the assessment of research training programs
- 10.6 Summary of major recommendations

### **Appendices**

- Appendix 1: Biographical sketches of TF members
- Appendix 2: Annotated Table of Contents of background briefing package at outset of TF work
- Appendix 3: Blank Table for summarizing an AR's total graduate educational Experience
- Appendix 4: "Intake Survey", or guide to the analysis of student competencies during the program

### References

**TASK FORCE MEMBERSHIP**

Institutions represented: Schools/Colleges of Education at Harvard, Michigan State, Northwestern, Stanford, Teachers College, UC Berkeley, UC Los Angeles, University of Cape Town, U Michigan, U Penn, U Wisconsin.

Robert Floden, Michigan State University

Hunter Gehlbach, Harvard University

Carol Lee, Northwestern University

Judith Warren Little, University of California, Berkeley

Rebecca A. Maynard, University of Pennsylvania

Mary Haywood Metz, University of Wisconsin, Madison

Elizabeth Birr Moje, University of Michigan, Ann Arbor

D.C. Phillips, Stanford University (Task Force Chair)

William A. Sandoval, University of California, Los Angeles

Stephen Silverman, Teachers College, Columbia University

Crain Soudien, University of Cape Town, South Africa

Task Force Research Assistant: Rebecca M. Katz, Stanford University

Biographical sketches of these Task Force members are given in Appendix 1.

## PREFACE

The nature, standards, and rigor of empirical social-science oriented educational research have sparked discord for at least a century. The relationship that does, or should, hold between the findings of such research and actual educational practice similarly has been contentious. Incompatible positions have been defended with vigor: On one hand there is the perennial charge that educational research is too “ivory tower” and should be oriented much more towards elucidating the problems faced by practitioners, but on the other hand research has its own trajectory and “never has the purpose of solving human or social problems” as a president of AERA, Fred Kerlinger, put it long ago (1977, p.6). Furthermore it frequently is claimed that there is great demand for “research-based practice”; but it has also been asserted that “there is no army of educational practitioners expectantly waiting to hear what the fundamental researchers have to say” (Jackson and Kieslar, 1977, p. 13), an assertion that is fueled from time-to-time by revelations that practitioners rarely read the research literature. Analysis reveals that a large amount of educational research is ideologically-laden, and that many of its so-called “findings” are not supported by the evidence that researchers offer for them (OFSTED Report, Tooley and Darby, 1998); yet it also has been claimed that “many features of educational research are today healthy” (Cook and Payne, 2002, p. 150).

Many university schools or departments of education with doctoral level research training programs have a long history of serious engagement with such issues, for of course they aim to produce researchers who are able to navigate around the dangerous reefs and do work that is both intellectually rigorous and has social and educational relevance. And although these institutions have had some successes in producing graduates who do “good social science”, as Kincaid put it in a quotation in the Frontispiece, often they have not been satisfied that in general their research training programs have met the relevant challenges adequately.

### **Improving research training: Complexities and constraints**

Over the years, doctoral-level research training institutions have also found that the crafting of satisfactory programs is made more difficult by internal complexities and constraints: First, educational research is a thread that weaves its way through most, if not all, graduate programs in education – but not all students intend to become empirical researchers, or researchers at all,

and thus there is considerable variation in their interest in the “mechanics” of research. Nevertheless, those doctoral students who intend to pursue an applied or practitioner role still require some knowledge of the research craft, and some ability to read and analyze research reports, in order to be intelligent consumers of reported -- and purported -- findings. Students with this career-orientation usually are not well-served by being required to enroll in the opening course of a quantitative or qualitative sequence for such courses have a more specialist purpose, yet this has often been the path of least resistance. Faculty members who teach research-oriented courses face a considerable challenge here. And of course some doctoral students are adopting research-orientations traditionally related to the humanities (philosophy and history of education, for example), yet also require enough background to be “intelligent consumers” of empirical research carried out in a social science mode. Nor should it be overlooked that many students change their substantive and methodological interests during their training.

Another constraint that has been faced by virtually all institutions is that some of those doctoral students who are aiming for a career oriented towards empirical social-science type research lack first-hand acquaintance with the domains of educational practice that will be the focus of their inquiries, which puts their work at risk of being superficial and irrelevant to practice. Other students, however, possess such a background – and so the diversity among students on this important dimension makes organization of programs difficult.

An additional set of challenges arises from the fact that some students who aspire to be empirical researchers lack prior exposure to one or other social science domain upon which their doctoral training can build, and this places further demands on instructors of the introductory methodology sequences in a School of Education. Here, too, advisors have a difficult task; they need to ensure that students acquire in several years of coursework, assistantships, and dissertation writing, the capacity to produce research as methodologically rigorous and theoretically as sophisticated as that produced by students in parallel departments in other parts of the university. The total social science training of graduate students in sociology and psychology, for example, has been several years longer than that of many

students in education, because usually graduate work in those fields builds upon a related specialist undergraduate major.

Finally, many students who do have a solid foundation in one particular disciplinary approach to research lack knowledge of – and sometimes display no interest in – alternative theories and methodologies from which they would benefit. Convincing students and some of their advisors that breadth of research training is a virtue as important as depth has proven to be a difficult task for many doctoral training institutions (as has convincing some students and faculty members of the obverse, namely that depth is as important as breadth).

### **The experience of the RTG institutions**

Since the mid-1990's a dozen or so university Schools or Colleges or Departments of Education have been recipients of Spencer Foundation research training grants (RTGs) which reflected the Foundation's long-standing commitment to strengthening the quality of educational research. Clearly the RTG institutions have not been insulated from the pressures and constraints sketched above, although the generous Spencer funding has stimulated each of them in their own way to confront the challenges. Virtually all of them have been able to design and evaluate incremental improvements to their programs, to launch innovative research seminars and mentorship programs and the like, and occasionally to make even larger changes, all with the aim of increasing the capacity of their students to produce rigorous and relevant educational research. Reflecting the differences in their size and structure, there was considerable variation in the way in which the RTG institutions proposed to use their Spencer grant, and there was a Foundation requirement that most of the funds be used for student support.

However, even these RTG institutions, which are highly motivated, have not found the going to be easy – the path to improvement is, as the saying goes, “paved with good intentions” but nevertheless leads through challenging territory littered with the bones of contention. They have found that reaching faculty agreement at the “nitty-gritty” level, about the specific details of their research training programs, has been difficult, even though the majority of the faculty at the respective institutions share the same methodological and philosophical commitments as espoused by the distinguished philosophers of social science Mario Bunge and Harold Kincaid

in the passages quoted at the outset. Individual faculty members at these institutions – who hold their own research work to high standards of rigor, relevance, objectivity and social responsibility – do not always agree on what they mean by these terms, for what constitutes rigor and the rest are matters in dispute across – and even within – disciplinary and methodological borders. And even when they do agree, differences of opinion can exist about how to achieve these goals with students. It is, after all, quite a tall order to produce students who possess the kinds of skills referred-to by Bunge – that is, to produce new researchers with solid methodological skills, theoretical sophistication and precision, practical acumen, and moral sensitivity and determination.

Unfortunately there is only a limited case literature available to give faculty members detailed, specific guidance in their training endeavors, and faculty at different institutions labor in relative isolation – not only from their colleagues in other places but also from their colleagues down the corridor – and so they are unable to profit from each others' successes and failures. Furthermore, a few recalcitrant faculty members look askance at the suggestion that their doctoral assistants can be – or even ought to be – explicitly taught to be objective fact-finders, rigorous theorizers, morally sensitive, and socially responsible. Instead they have a view that their students will learn and develop by osmosis, by simply spending a number of years in close proximity to their mentor. And of course, some students actually do develop remarkably under such a regimen, but many do not.

Over the years, Deans and key faculty members from the RTG institutions have met regularly to exchange ideas and to report on progress, and it was at one of these meetings that the idea was born to establish a Task Force to pull together – and to build upon where necessary – what had been learned during the decade of Spencer sponsorship. Thus this present report was written as, in a sense, the culmination of the Spencer Foundation Training Grant program. Most of the recipient institutions agreed to nominate an experienced faculty member to serve on the Task Force (TF), the purpose of which has been to gather information about what had been achieved within the various training programs, to evaluate and reflect on this material and on the existing literature concerning research training – and then to use all this as a springboard for making recommendations about the content and dimensions of research training in two specific

empirical domains that would be likely to produce individuals who could do rigorous work, and who also would be able to navigate comfortably the complex intellectual, social and educational worlds of the new millennium. (Relevant portions of the proposal to the Spencer Foundation, that contain the precise charge to the Task Force, are reproduced in the following section.)

The leaders of the RTG institutions intend that the report of the TF shall be used initially to guide an evaluation of their respective research training programs in the two relevant empirical domains, and the recommendations also will be used as pointers to productive directions for change. Members of the TF hope that a broader range of research training institutions also will find this report to be useful.

### **The Task Force at work**

Members of the Task Force decided that this Charge required us to do more than simply reflect on past history – to do more than merely collate and then reflect upon the best practices at our respective institutions. We believed that our work needed to be forward looking, pointing towards avenues for fruitful change; but we also recognized that any specific recommendations needed to be warranted or supported in such a way that our academic colleagues would deem them to be worthy of serious consideration. We also wished to avoid hortatory prose, but nevertheless intended that our deliberations would enable us to make convincing, well-grounded, concrete recommendations. We attempted to bear in mind, as we crafted these recommendations, that the various RTG institutions varied with respect to the specific constraints and affordances that they faced, and that therefore a recommendation that was relatively easy for one institution to enact might be virtually impossible for another.

In short, members of the TF recognized that we had the extraordinary opportunity to develop *a fresh conceptualization* of graduate level research training, one that made use of what had been learned from the past, but which was not held captive to it. Dewey reminds us that “the business of education” and of learning is to liberate us from “reviving and retraversing the past”, but also that “it is a part of wisdom to utilize the products of past history so far as they are of help for the future” (*Democracy and Education*, 1916/1958, pp. 85-6).

Prior to our first meeting, the TF members familiarized themselves with more than a dozen key contributions to the published literature on issues in research training, and with the activities that our respective institutions had engaged-in under the auspices of the Spencer RTG grants they had received. (The annotated list of readings in the initial briefing package received by all TF members can be found in Appendix 2.) Over the course of eighteen months three working meetings were held, each of several days duration; at these there were vigorous whole-group discussions, and a great deal of drafting of material in sub-groups composed of members with expertise relevant to the writing tasks in hand. In between these three meetings a great deal of writing took place, and the TF chair collated and edited the drafts which were then sent to all members for further comment.

The first TF meeting was groundbreaking: It was decided that, instead of providing stultifying lists of innovations in training that had been made under the auspices of the various Spencer RTG programs, we would draw on these innovations to construct composite vignettes in which we followed the progress through doctoral training of two fictitious but none-the-less representative students who eventually became well-versed in different approaches to their research – one broadly quantitative, the other broadly qualitative, the two areas specified in our Charge. In essence, these vignettes and the accompanying comments provide an engaging account of many of the best practices across our various institutions, and they also contain in nascent form many if not all of the recommendations that later we make explicit.

But the TF went even further. Too often discussions of research training take place in a decontextualized manner, whereas the view that emerged from our initial discussions was that the training required by aspiring researchers needed to be shaped by the demands imposed by the contexts in which their research work would be carried out. And, crucially, we recognized that multiple relevant contexts exist in these opening years of the new millennium, and we discuss these in a later section. Our thinking here can be illustrated with the analogy of training members of “search and rescue” teams. Individuals who will serve in alpine terrain will not be well- prepared if their training is based on the assumption that they will be serving in the desert, and training that uses antiquated equipment will not suffice for a context dominated by

the use of helicopters, jeeps and snowmobiles, and GPS devices. To have a chance of being effective, training must be based on an analysis of the demands of the environment in which the trainees eventually are going to find themselves.

So, after precisely describing the Charge, the discussion will turn to amplifying the claim that a new approach to doctoral-level research training is needed. This is followed by our analysis of the four universes in which the emerging cohorts of educational researchers will have to function; after which there are several sections in which we discuss in more detail some of the consequences of the new orientation we have adopted. Perhaps most important here, and the most open to misinterpretation, is our amplification of the point made by Arthur Kaplan in our Frontispiece that the cultivation of methodology is neither a necessary or a sufficient condition for excellence in research – there is more to scientific worthiness than this, although we are not denigrating the importance of methodology. To say that more than this is needed is not to call for less of this important facet of research. This discussion lays the groundwork for our presentation of the vignettes about the training of two doctoral students, vignettes that together with the accompanying comments serve to summarize many of the interesting practices that have emerged in the RTG institutions. These vignettes, in turn, lay the groundwork for the detailed recommendations that we go on to make about doctoral training.

The Report that follows is a consensus document – all of the substantive points were endorsed by all TF members, although (as in any academic group) there were some residual differences of opinion about wording and emphasis in a very small number of instances.

### CHARGE TO THE TASK FORCE

The following is an extract from the proposal made on behalf of the Deans of the Spencer Research Training Grant Institutions to the Spencer Foundation, wherein the context was outlined and the charge to the proposed Task Force (TF) was laid out:

This proposal is written on behalf of the deans/chairpersons of those Schools or Departments of Education that are recipients of a Spencer Research Training Grant (RTG). These dozen institutions have not all taken the same approach to preparing students to produce *high quality, rigorous research*; and there are differences in the ways in which (and the degree to which) this preparation takes into account the extremely complex intellectual, socio-political and educational settings in which their graduates are destined to work. Although there is emerging consensus about some matters, unanimity has not been reached about all the complex issues involved. Therefore it seems timely to establish a *Task Force* with one representative from each of the participating institutions; the function of this would be, in general terms, to tease out and if necessary elaborate upon what has been learned, and to feed the results back to the respective institutions to guide future improvements to their training and assessment activities. Ultimately the goal is to produce a document that could be used as a guide for other Schools of Education endeavoring to improve research training.

In more precise terms, the members of the Task Force, informed by the insights about rigorous doctoral training gleaned from the RTG experience at their own institutions, shall deliberate and report upon, in as concrete and performatively specific a manner as possible:

(i) The key elements that need to be addressed in the doctoral-level training of high-quality educational researchers who will work in one of several important empirical traditions (outlined below). Issues to be reported upon will include:

(1) the depth of specialized methodological training in the relevant tradition, (2) the breadth of acquaintance with alternative methodologies, (3) the degree of awareness of the values and epistemological assumptions that form the intellectual context for research activities, (4) the mechanisms by which relevant contextual knowledge is or can be acquired, and (5) possible illustrative exemplars of work within the respective traditions that embody the kinds of desiderata indicated above and which will be amplified during the task force deliberations.

(ii) The ways in which the member institutions currently assess the effectiveness of their research training practices, and any recent results of such assessments.

There is a vital final stage to the work proposed here; it is not the responsibility of the Task Force but of the leadership of the participating Spencer RTG institutions subsequent to the work of the task force, but building upon the materials contained in its report that will serve as groundwork. The RTG institutions will have a commitment to use the report to guide a self-assessment. The aim shall be to determine at what precise level their students are able to function in, or show mastery of, the various dimensions of research briefly outlined above and as elaborated by the Task Force report; and also to determine whether any curricular changes need to be made in their respective training programs.

Later in the same document, two points were made that in essence elaborate upon the charge outlined above:

(i)... it is emphasized again that the RTG institutions may not (yet) be in agreement about all the pertinent issues. There is agreement, of course, that methodological training is necessary, but many difficult issues can be expected to arise in pinning down its precise details and the level of performance that should be expected of doctoral students in specific fields; achieving some level of consensus will form an important part of the work of the task force. However ... there is emerging recognition by the leadership of the RTG institutions that such training should not constitute all of – or perhaps even the “lion’s share” of – doctoral research training. But perhaps the most hotly contested issues – which the task force also will

attempt to charter – arise in the delineation of the *additional* aspects of research preparation and the depth to which each of these other constituent elements need to be pursued. (ii) ... it is recognized by the leaders of the RTG institutions who are sponsoring the present proposal, that there are so many contending frameworks or traditions or paradigms (“disciplines”), and that the issues relating to quality and rigor are so complex within and among them, that it is important for the Task Force to set out initially with modest aspirations. Thus, for the purposes of this proposal, the Task Force should clarify criteria of quality and rigor, and unearth the epistemic and other values that also help to shape inquiry, quality and rigor, for a narrow range of contrasting (but possibly complementary or compatible) major research traditions. The labels used to identify these are controversial, and are often themselves sources of confusion. *We use the following for convenience with some trepidation:*

\* Correlational

\* Qualitative case studies

It is worth emphasizing that notes from several meetings of the RTG Deans makes clear that they recognized there are many important disciplinary and methodological frameworks from which education research can be launched. The selection of the two categories above as the initial focus was tentative, and was not intended to denigrate the other research orientations; furthermore, description of the categories was deliberately left loose so as to not overly constrain the TF if its deliberations led it to broaden the focus. Nevertheless the clear consensus of the Deans who commissioned this report was that the TF should not venture too far afield.

Members of the TF have done their best to work within the constraints suggested by our Deans. The training of aspiring researchers or scholars who – for example – are historians, philosophers, curriculum theorists, cultural critics, or postmodernists/poststructuralists, is beyond even the most liberal reading of the TF charge. This is not intended to denigrate these fields, nor the training needs that exist within them; it simply is the case that, given the constraints of time and expertise, attempting to do too much would have resulted in our accomplishing too little. The complex title to our report reflects our struggle to find a way to accurately summarize the narrow but important task with which we did engage.

**THE PREPARATION OF ASPIRING EDUCATIONAL RESEARCHERS IN THE  
EMPIRICAL QUALITATIVE AND QUANTITATIVE TRADITIONS  
OF SOCIAL SCIENCE**

**Methodological Rigor, Social and Theoretical Relevance, and More**

**1. Introduction: Adopting a new perspective**

There was a time when the cultivation of plants was – at least in principle – a straightforward matter. A seedling was planted in the earth, watered, and some fertilizer was added; the gardener’s job then was to wait, perhaps fending off snails and slugs and other pests. The seedling either made it, or didn’t. Today matters are not so simple; due to increase in knowledge (a product of reflective experience and of agricultural research) we know that a seedling does not merely sit in a hole in the ground but inhabits a microclimate, one with a number of specific features – such things as a particular temperature range, full or partial sunlight, and so on. Furthermore, we now know that it is not just any old hole in the ground, but it is a hole in soil that may lack certain trace elements that the seedling needs, or perhaps it is soil that is too acidic or not acidic enough. Agricultural science has made more complex the world in which plants grow, and a competent gardener must be able to navigate around this new world, which of course is constantly becoming ever more complex.

Following in the footsteps of Rousseau, Dewey and other thinkers who have compared the educational process with plant growth, we offer the account above as a metaphor for what has happened in the training of researchers. There once was a straightforward – and what with some license might be called a linear – approach, but now a great deal of knowledge has accumulated, so that the needs of the aspiring researcher are better understood. The intellectual, professional, and socio-political contexts in which the researcher needs to be able to navigate are no longer invisible or out of focus but rather are recognized to be extremely complex, fraught with dangers and difficulties, but also rich in affordances. Finally, of course, the educational phenomena that the researcher will spend his or her career investigating are now known to be richly layered in ways that would have astounded researchers at the birth of the profession something over a century ago. Complexity has been layered upon complexity, and

because of this the urgent need has arisen for re-conceptualization of both the scope and the depth of research training.

Members of the Task Force have come to believe that the old approach to the training of researchers, which – in only a slight parody – can be thought of as based upon a linear and additive/deficit model, is no longer appropriate. In this now outmoded approach there is an uncanny parallelism with the old view of gardening: Aspiring researchers (ARs) enter graduate school with some skills and a stock of knowledge acquired during their undergraduate education, and often but by no means always with classroom or other educational experience that has solidified their professional identity as practitioners. They are then immersed in a specialized training program where they are expected to take root and start to acquire a new identity. As developmental deficiencies become apparent, a treatment of fertilizer is administered, in the form of an extra requirement or two being added to their programs. Like the poor seedlings, the ARs either make it or not.

*To put it in slightly different terms, we reject starting our rethinking of research training with the question “what competencies should the AR acquire in graduate school, and how many courses would be needed to ensure mastery of these?” We reject this way of starting, chiefly for the reason that it decontextualizes research training – it places into a secondary position any consideration of the world in which the aspiring researchers will be pursuing their craft. And crucially, just as the world of the growing plant has become more complex, and is ignored by the gardener at his or her own peril, so too with the world – or rather the worlds – of the researcher. As Lindblom and Cohen put it, professional social inquiry (PSI) “continues to uncover even more complex aspects of reality”; and quoting Thorstein Veblen, they wrote:*

*“the outcome of any serious research can only be to make two questions grow where one question grew before.” We suggest that the usual effect of PSI is to raise new issues, stimulate new debate, and multiply the complexities of the social problem at hand. (Lindblom and Cohen, 1979, p. 48)*

Inevitably, in due course we need to address in specific terms the competencies, skills, courses, workshops, and projects that we regard as necessary to prepare graduate students to work as

researchers in the complex settings that their forebears helped to create; but the point is, we cannot begin our work by focusing on such issues.

Thus, our approach to the task of re-conceptualizing research training has been markedly different from what has been traditional. Informed by experience in the Spencer RTG programs at our respective institutions, and by our sampling of the recent literature on research standards and epistemologies of inquiry, we have come to the view that assessment and re-conceptualization of training should start from the frank recognition that *the ARs who emerge from a training program in the early decades of this new millennium will find themselves situated within a number of complex, rapidly-expanding and overlapping universes that impose difficult burdens but which also offer exciting affordances*. In our judgment, graduate-level training (or rather, graduate-level education, for “training” implies a degree of narrowness that we wish to transcend) should at a minimum equip the AR to sally-forth cautiously into each of these universes; and we further believe that this can be done without jeopardizing the AR’s exposure to the technical intellectual equipment of the relevant research discipline. Later in this Report, from Section 7 on, we show that even the relatively exemplary education of two students we take as cases can be illuminated and strengthened by explicitly taking these different universes into account.

## **2. The multiple universes in the life of the educational researcher**

In the following pages we present an initial and rough characterization of these interacting universes. We stress that we view this as a framework or analytic device, one that we found useful and that provided us with some direction during our re-thinking of doctoral level training programs. We recognize that there are other ways to conceptualize the realms in which the AR of the early 21<sup>st</sup> century will be working, and we recognize also that the details of our analyses of the various universes – such as they are – may be contested. But, to put it candidly, our interest has not been to fine-hone the tool, but rather to see what we could carve-out with its assistance. With these provisos, the four universes that we see as important are as follows:

## 2.1 Description of the four universes

1. *A universe of research frameworks or paradigms*: This is the universe the AR typically enters in graduate school, initially via coursework; it is within one or another framework from this universe that the AR will carry out his or her research. It is extremely complex, being made up of a large number of social science theories or conceptual frames, which are often closely intertwined with specific research methodologies, data analytic techniques, and the like. Furthermore, it is a constantly expanding universe; the actual number of theoretical frameworks is increasing, and there is also growth within many of the frames, and new statistical and other research tools and approaches constantly appear. Different “epistemologies of inquiry” are associated with the various theories and methodologies. Most ARs explore only a small portion of this universe during their training.

Thus (to mention only a few examples) this is the universe of Piaget and Vygotsky, of Newell and Simon, of Lave and Wenger, of Marx and Freud, of Geertz and Mead, of Chomsky and Labov, of Greene and Gilligan and Noddings; it is the universe in which many citizens are “bowling alone”, and where students may be members of “communities of learners”; it is the universe of social stratification and of “loose coupling.” It is the universe of descriptive statistics, of HLM, of meta-analysis, of experimental design, of observation and interview protocols, of qualitative data reduction techniques; it is the universe of Campbell and Stanley, of Miles and Huberman, of Glaser and Strauss. It is the universe where positivism and post-positivism and post-modernism face-off, and it is the universe of Popper, Kuhn, Lakatos, Dewey, Lyotard and others.

Many contentious issues arise within this universe, but often they are debated in the fourth universe described below. Nevertheless, a body of widely-accepted findings has emerged from the research carried out under the umbrella of the theories and methodologies in this universe; these findings constitute the

background knowledge that ARs need to have mastered, at least in part, and for convenience we locate this body of material in Universe 3.

**2. A *discordant social universe*** that is the context in which researchers do their work, for this universe is the setting that houses the educational programs and processes that are the focus of most (if not all) educational research. (We have more to say about *context* in Section 4 below.)

Thus this is the universe inhabited by practitioners and policy-makers, and it is where students attend school and learn or fail to learn. It is the universe where struggles for social justice occur; and where there are scarce resources for which often there is fierce competition. It is the universe of clashing ideologies; it is where political pressure and interference are rife, where groups of stakeholders supporting rival social and educational programs are active and extremely vocal in pursuit of their interests (e.g. teachers' unions; parents' groups, religious sects), and it is where interest groups make selective use of research findings to strengthen their cases and where work that is deemed to be unfavorable to their cause sometimes is suppressed.

**3. A *universe of educational and related substantive knowledge and research findings*** that constantly is being added-to, due to the work of skilled practitioners who have found ways of communicating their insights, and due to the efforts of researchers who are working within one of the research frameworks in Universe 1. These insights and research findings may be contested by researchers and practitioners from different frameworks; for example, what research tells us about learning to read is contested by supporters of the "whole language" and "phonics" programs. Much of this contestation takes place in the professional venues that form part of Universe 4, but it can also spill over into the broad social universe described above. This universe contains the background knowledge that ARs require to frame their research; they become acquainted with this material in the substantive courses they take, and

when preparing literature reviews.

As a result of research that has accumulated over the past few decades, it is realized now that educational phenomena that once appeared to be simple are actually much more complex and are more interrelated with other social forces and social structures than we had previously realized. Thus, to cite a well-known example, if Kohlberg was now doing his work on moral development, it would be impossible for him to overlook gender, social class and ethnicity as being relevant factors. Additional complexities arise from the fact that groups whose voices were previously silenced – and hence whose interests and welfare easily could be ignored by researchers – have now to some degree been enfranchised. Education now has to be seen as integral to struggles for social justice and economic development, and thus the requisite knowledge-base has been greatly expanded.

**4. *The universe of the professional infrastructure of educational research.*** This contains the field's professional associations (AERA, and specialist groups for economists, historians, sociologists, evaluators, and so on), the annual and the "one-off" conferences, and a variety of journals both national and international. Failure to navigate successfully in this universe can have dire consequences for the AR's career – from not obtaining research support to not being granted tenure.

But in addition to its career-enhancing potential – providing venues to publish and opportunities to make professional contacts (including potential referees) – this universe is important for being a domain in which the AR will be forced to confront a smorgasbord of contentious issues about research that have arisen in the first universe described above (and probably in the second and third as well) and about which he or she reasonably can be expected to take a stand. Nor should the classroom context be overlooked here, and the importance of learning to teach and talk about research findings and methods – for example, when

serving as a teaching assistant the AR may well find that fundamental issues come to the surface about which his or her students need guidance.

A Task Force member provided the following example of a situation where ARs who were present could hardly fail to have had their suppositions about research and design of program evaluations put under stress: Several years ago, alongside more than two hundred others, he attended a session at the annual meeting of the *American Evaluation Association*, at which several panelists discussed both their work as editors/referees for the "What Works Clearinghouse" website, and the research values and epistemological beliefs that underscored their overwhelming commitment to the view that well-conducted true experiments served as the "gold standard" in research design. The discussant strongly criticized the underlying view of the nature of scientific research that he believed was held by the panelists, and he also argued that other research designs can uncover causal relationships satisfactorily. A leading evaluation theorist/methodologist, who also had a significant grounding in philosophy of science, joined the discussion from the floor, in general supporting the discussant; very soon emotions were running high, with assertions and counter-assertions about research and evaluation methodology and epistemology being tossed back and forth with increasing vigor, and with many audience-members being drawn in. Everyone in the auditorium seemed engaged, and judging from comments afterwards many had been forced by the exchanges to rethink some of their deep-lying suppositions.

Almost certainly reading a journal such as the *Educational Researcher* is a less-dramatic experience than the one described above, but nevertheless over the course of a year or two enough probing essays appear to force the attentive reader to do some soul-searching. It seems obvious to members of the Task Force that the AR should be able to profit from such experiences – should be able to follow the debates, if not to participate actively in them, and should be aware enough of his or her foundational suppositions about research so that these can be revised (or perhaps defended) in the aftermath.

## **2.2 How the notion of “universes” brings the training issues into focus**

If the premise is accepted that the work of ARs will be located in, or will bring them into contact with, these four universes (or some subset of them, depending upon the specific project), the issue then arises as to how their graduate programs can best prepare them. It seems clear to the members of the TF that if an aspiring researcher is immersed only in a small subset of the research methodologies and theoretical frameworks that are found in Universe 1, and is relatively ignorant with respect to the epistemological underpinnings of the chief theories and methodologies in that universe (a situation that some TF members believe to be too common at present), and if this AR only has superficial acquaintance with Universes 2 and 3, and is naïve or innocent about the workings of Universe 4, then several adverse consequences are likely to follow. The AR's work will be substantively and methodologically narrow; questions that ought to have been asked might not have been pursued, for these might arise from literature with which the AR is not familiar and the necessary wider investigation might have required theoretical and methodological acumen that the AR does not possess; the AR might well be devastated both by the criticisms that come from adherents of rival methodologies and epistemologies, and also by attacks from displeased stakeholder groups and the like. Furthermore, methodological purity or narrowness will not protect the AR from technical criticism – grounds can always be found upon which to mount an attack upon aspects of study design and data analysis, and indeed it is a sad fact of life in the research world that “political” or ideological differences with the researcher can surface in a thinly-disguised technical form. It is worth noting that there is now a category of consultants for whom “doubt is their product” (Michaels, 2005); that is, these experts can be hired to debunk on technical grounds a piece of research the findings of which go counter to the interests of a client. (The phenomenon of debunking is well-known in the field of educational program evaluation, and wise evaluators quickly learn to “guard their backs”.) It seems to be a counsel of wisdom to know something about the frameworks from which such criticisms can be launched, and this broad background knowledge might also serve to strengthen the initial piece of research.

Perusal of the descriptions of the four universes reveals the extent of the challenge to doctoral training. How deeply in each should the AR be immersed? Are there some topics or areas

within each universe that are absolutely essential, both for survival and for future professional growth as a researcher?

Consider as an example the disputes about the nature and quality of empirically-oriented research in education, briefly touched-upon in the *Preface* and outlined somewhat more fully below (for documentation, see Phillips, 2009). Most of the issues in contention have arisen within the first universe described above, but the debates themselves have taken place in the fourth universe. Only the most insular and narrowly-read researchers would have failed to make some contact with the warring ideas; and indeed in recent years the issues at stake have been featured prominently in journals and at conferences. And they are issues with which the AR could be expected to personally have grappled, for they have bearing upon the professional role of the researcher.

The example is as follows:

As sketched earlier, for a century or more the general corpus of educational research has been subjected to the recurring criticisms that, on one hand, it is too theoretical and “ivory tower”, and on the other hand that it lacks theory and is too focused on practical matters. Critics also suggest that educational research is infected by ideology; and it has been documented that the evidence presented by researchers often fails to support conclusions drawn. Furthermore, during the closing decades of the twentieth century the so-called “paradigm wars” between supporters of experimental/quantitative approaches and advocates for ethnographic/qualitative modes of inquiry were waged at professional gatherings and in such widely-read journals as the *Educational Researcher*; however, in the new millennium there are indications – such as the rise in advocacy for mixed-methods research – that some scholars at both ends of the methodological spectrum are retreating towards a compromise. (The key issues in this “warfare” have been described and analyzed in Howe, 2003.) Debates about “truth” and “relativism” multiplied in educational research journals and conferences from the 1980s down to the present (see, for example, Siegel, 2006), and during the same period alternative research approaches such as narrative

research, portraiture, and connoisseurship became popular among researchers who argued these were based upon epistemologies that were less problematic than the positivism that under-girded dominant approaches to empirical research (for discussions and references, see Phillips, 2000).

In recent years, too, the international push for so-called “evidence based policy” has inevitably focused attention on the quality of evidence available in the education research literature, and has led to unflattering comparisons with the (reputed) quality of evidence available to policy makers in the public health field (for one example, see Mosteller and Boruch, 2002). Recently the National Research Council report (NRC, 2002), which attempted a modest dampening of enthusiasm for the randomized field trial that was being promoted by some as the “gold standard” methodology for education research, was strongly critiqued (on diverse and often incompatible grounds) in symposia in special numbers of *Educational Researcher* (2002), *Qualitative Inquiry* (2004), *Educational Theory* (2005), and *Teachers College Record* (2005).

A multi-faceted controversy such as this forms part of the intellectual environment into which doctoral students following a research or policy path after graduation inevitably will be thrust; to what degree should the training of these ARs familiarize them with the pertinent issues, and how precisely should this be done? Would training in Universe 1 be complete if the AR was not able to face such issues competently? Does mastering a research methodology of necessity involve understanding its epistemological foundations?

And there are numerous less rarified issues: Should an AR who is going to carry out research in a qualitative mode have at least a nodding acquaintance with significance testing? What about the more advanced contemporary tools, such as hierarchical linear modeling or meta-analysis? How much of an overview would make sense and also be helpful, and would it be possible to teach to ARs without a technical background? Should a quantitatively-oriented AR be familiar with the epistemological backing for certain modes of qualitative inquiry, and also be familiar with the “threats to validity” of qualitative research and with how these can be countered? How

much familiarity is enough? How deeply should an AR whose specialty is the design and conducting of tight randomized, controlled experiments understand the complexities in the settings in which this type of work will be conducted when the insularity of the laboratory is left behind? And, for example, should this AR be able to respond to the criticism – or even simply be able to comprehend the point that was being made – that he or she is using “your father’s paradigm” (a charge that was leveled at the authors of the 2002 NRC report; see Lather, 2004)?

### **3. The consequences of leaving a solitary and “closeted” universe**

The point of the discussion above is straightforward: the AR does not, and will not, live in a closeted world where he or she is protected from sharp questioning about all levels of a piece of research; aspiring researchers – together with those who are well-established – will not be able, as Jerome Bruner once put it in the context of psychology, to “seal themselves within their own rhetoric and within their own parish of authorities” (Bruner, 1990, pp. ix-x). Instead, the AR will be living in hotly-contested terrain where the universes described above interact and overlap, where the flow of questioning and criticism will not abate. Researchers are likely to be called to account about the social or educational relevance of the research; about the way in which the questions driving the research had been formulated; and about the design, the measuring instruments, and the data collection techniques. There are likely to be questions about whether or not (and why or why not) intangible or not-directly-observable factors were taken into account, such as the beliefs and motivations of those individuals being studied – and whether or not the stance adopted here makes one a “positivist”. The researcher may also be challenged about whether or not a confirmatory or a Popperian disconfirming orientation had been adopted; and about the social and ethical values that lay behind the way the research had been conceptualized and executed.

It needs to be made clear that the Task Force is not making the unreasonable suggestion that the AR, by time of graduation, should have become a Renaissance person who is able to function with remarkable facility in a number of disparate domains. It *is* being suggested, however, that in order to produce work of sufficient quality to withstand critical scrutiny from many directions – work that cannot easily be dismissed as “missing the point”, “simplistic”, “badly

conceived”, “flagrantly biased”, “riddled with ideology”, “based on a philosophy long-since outmoded”, or “technically naïve if not incompetent” – the AR needs to have been *educated* rather than trained. For being educated implies that the AR, in addition to rigorous training in his or her core theories and methodologies, has acquired a degree of intellectual breadth, and has developed the disposition to probe issues to some depth and with critical facility.

The point also needs to be made that disciplinary compartmentalization is intellectually stultifying. Much ground-breaking work has come from researchers and others who received stimulus from across boundaries. Jerome Bruner pointed out that Chomsky was greatly influenced by Descartes, Piaget by Kant, Vygotsky by Hegel and Marx, and the classic learning theorists by John Locke (Bruner, 1990, pp. x-xi). He could have added that Donald Campbell and Paul Meehl were influenced by Popper, Clifford Geertz by Weber and Wittgenstein, John Dewey by William James and Charles Darwin, and B.F. Skinner by the logical positivists.

#### **4. The importance of contextual knowledge**

There is an alternative terminology available to describe what happens when the AR leaves his or her “closeted” world to enter the realm of multiple jostling and overlapping universes – namely, the *context* for the ensuing research obviously changes and acquires complexity. But it is not only the context of the phenomenon that is being studied that is important – the researchers themselves are situated in a complex theoretical/methodological context, and also in a professional one. Why introduce the new terminology of “universes” when the old seems satisfactory? The answer is simple: It is a matter of emphasis.

We make several points here. First, members of the TF wanted to acknowledge in a striking way, and also to build upon, the growing importance in the social science literature of “contextual understanding”, “socio-cultural context”, and the like (for representative discussions see Flyvbjerg 2001; LeCompte, Millroy and Preissle, 1992). Often these notions enter educational research by way of the recognition that educational programs and policies, and of course actual episodes of classroom teaching, are affected in many ways by the multiple contexts in which they are located – such as the socio-economic and ethnic composition of the community served by the school; the gender composition of the classroom; the rules and

procedures of the particular school; the composition, training, and experience of the teachers; the school maintenance budget; and the subject being taught in the particular classroom. In essence, these contexts make up Universe 2. Research that ignores these contexts faces the danger of being simplistic at best, and at worst irrelevant to the actual problems of education – hardly the hallmarks of quality work. But we also stress that the contexts of the other universes are of great importance, although it has not been common for these to be discussed *as* “contexts”.

An additional point is that if an understanding of the context of an educational program or phenomenon is lacking, the concepts being used by researchers run the risk of being ecologically invalid. An example suggested by a TF member is that poor performance on a test might be interpreted by the culturally-naïve researcher as an indication of lack of relevant ability, or of low intelligence – but knowledge of the context, that reveals the full range of consequences of doing well on the test, might suggest the interpretation that performing badly on the test was an extremely smart thing to do. For example, doing well might result in being placed in a “stream” at school where the individual is separated from important peers, a consequence undervalued by school authorities but perhaps of major importance to the student.

Our second point, touched upon briefly above, is *that what is true for educational programs and classroom phenomena also holds true for educational researchers themselves* – a researcher working in the context of a closed closet, particularly one labeled “methodology”, will only see the interior walls and may not be aware of how much his or her work has been constrained or even compromised. In research, as in life, what one finds depends upon *where* one looks, and *how* one looks – and the tools and the methods that are used are determinative of these things. The report will return to methodology in the following Section, and the topic will be important for our discussion thereafter.

However, at this stage more needs to be said about the importance of the researcher having some understanding of that portion of Universe 2 that forms the context of the program or phenomenon that actually is being investigated. This “case for context” has been made in several slightly different ways.

Lee J. Cronbach and Associates, in a book that had an important impact on evaluation research in the early 1980's, argued that one of the key reasons that up until then the field of program evaluation was having such little impact on educational policy and practice was that it attempted to give very precise and technically unassailable answers to the wrong or irrelevant questions; these authors insisted that evaluators should have deep knowledge of the context in which the program to be evaluated was operating or else should use consultants who possessed this expert knowledge (Cronbach and Associates, 1980). The point, of course, is that being able to ask the "right question", or being able to select the telling problem, are contextually-sensitive activities; a researcher who lacks deep background understanding is unlikely to fare very well, and is thus unlikely to produce work of enduring quality. Members of the Task Force are well aware, of course, that the same can be said about the researcher with great contextual sophistication, but who is methodologically weak – we are not discussing "either/or" here.

A powerful account – that has hardly been bettered over a period of thirty years – of the contextual factors that are at play in educational settings has been provided by Parlett and Hamilton (1977); the passage illuminates why it is often an advantage for an AR entering a graduate research-training program to have had successful experience as a teacher or in some other practitioner role. A successful practitioner often can "read" complex educational situations (or the "learning milieu" in Parlett's and Hamilton's terminology) in a way that may be impossible for a researcher who lacks such experience:

The learning milieu represents a network or nexus of cultural, social, institutional and psychological variables. These interact in complicated ways to produce in each class or course, a unique pattern of circumstances, pressures, customs, opinions and work-styles which suffuse the teaching and learning that occur there.... For instance, students' intellectual development cannot be understood in isolation but only within a particular school or college milieu. Equally, there are phenomena of crucial educational significance (such as boredom, interest, concentration, 'floundering' and intellectual dependency) ... which

customarily arise as responses to the total learning milieu, not to single components of it. (Parlett and Hamilton, 1977, pp. 11, 13)

### **5. Loosening the identification of quality solely with methodological rigor**

If the points made above – about the overlapping universes of the researcher, the importance of contextual knowledge, and the impossibility of remaining in a closeted intellectual environment – are taken seriously, *they point to the incompleteness of identifying research quality solely with methodological rigor*. For undoubtedly, during the century-long debates about the complex and interrelated issues outlined above, there has been a pronounced tendency to approach the issues of rigor and quality solely in terms of methodology, and this has led to the giving of short shrift to such factors as depth of understanding educational contexts or classroom milieu, sensitivity to issues of social justice and access to educational goods, and epistemological sophistication. “Is this a piece of high quality research?” for many in the research community, has been regarded as synonymous with “is this piece of research methodologically sound?”

This identification of the two was abundantly clear in the opening years of the new millennium, in the rhetoric used initially in US Federal circles to justify the use of one specific research method – the randomized controlled experiment or field trial (RFT), the so-called “gold standard” methodology, and closely related designs such as tight quasi-experiments -- as the key criterion for making funding decisions, and for making decisions about what pieces of research should be posted on the Federally-funded *What Works Clearinghouse* website as being of sufficient quality to serve as a basis for setting policy or for selection of treatments.

Many members of the research community, however, have moved in a different direction, the one we have been pointing to in the preceding discussion: They recognize that no one research tradition or methodology has a monopoly on the production of fruitful educational research. The reason is not difficult to discern – educational phenomena are complex, and complex socio-cultural contexts and forces oftentimes have marked impact upon the educational processes themselves and thus on their results. A variety of rigorous – and vigorous – research traditions, with differing methodologies, have evolved to study different aspects of these complexities; and

it is becoming ever clearer that epistemological and value issues also need to be taken into account when designing research.

But whether it is with respect to one favored methodology, or with respect to some type of eclectic or multi-method approach, it is *methodology* that often receives the emphasis. Faculty in the leading research-oriented Schools of Education have not been immune from the seductive appeal of this identification of methodology with quality; and they often revel in discussions of technical matters concerning design, data collection, and statistical analysis. Over the years many have expressed a sentiment similar to that put pithily by Ann and Hubert Blalock in the "Preface" to a classic volume they edited forty years ago, *Methodology in Social Research*: "The impression that all our contributors wish to convey is that methodological issues *are at the core* of the social sciences" (Blalock and Blalock, 1968, p. ii, emphasis added). Another strong statement linking quality with methodology was made by Light, Singer and Willett in *By Design* (1990); the authors italicized the key portion to emphasize its importance: "Our basic tenet is that *your study's design is the single most important factor that determines whether your findings will be scientifically first-class*" (Light et al., p. 7).

Members of the Task Force, however, have adopted a slightly more nuanced view. Without decrying the importance of a study's methodology, issue is taken with the narrowness implied by the terms "core" and "single." But we were careful above in saying that the identification of quality with methodology was *incomplete*, not *wrong*. In our view, (i) it is too much of a simplification of a complex set of desiderata to hold that methodology stands alone as the "core" of research in education and the related social sciences, and (ii) it also is problematic to see methodology, or that sub-set of it dealing with design, as the "single main determinant" of scientific quality. Important as it is, methodology is only one aspect of a piece of research. It is worth noting that the history of science reveals that a study that is methodologically impeccable can nevertheless be deficient, or irrelevant, because it lacks other virtues; moreover, some major advances in science have come from work where the methodology was confused or deficient. Presumably this was the point of Kaplan's remark, cited in the Frontispiece, that "what is to be resisted is the notion that the cultivation of methodology is either necessary or sufficient for

successful scientific endeavor". *In short, when considering the criteria of excellence in research, the part should not be taken for the whole!*

The need to incorporate other criteria can be supported by numerous selections from literature on the nature of science. Thus, Sir Henry Tizard (at one time scientific advisor to the UK Ministry of Defense, who chaired the committee responsible for recommending the pre-WW2 development of radar) remarked that the "secret of science is to ask the right question, and it is the choice of problem more than anything else that marks the man of genius in the scientific world" (cited by Thomson, 1963). Further illumination comes from a memorable analogy used in Abraham Kaplan's classic work, *The Conduct of Inquiry* (1964). Kaplan remarked that a pervasive trait of American culture is "the overemphasis on what methodology can achieve" (1964, p. 24). He went on to state that he was objecting to the image of the methodologist as baseball commissioner, or perhaps as an umpire; instead, Kaplan argued that the methodologist should be regarded as a coach, the merit of whose "recommendations rests entirely on what the play of the game shows to be effective" (1964, p.25). Building on this analogy, our point is that a player or coach can be perfect methodologically or technically, but still not play or direct a game of high quality, because for this latter task much more than technical expertise is required. The winning coach, and the great player, must also have the ability to analyze and comprehend the context at crucial times in the game, and the operative constraints and affordances. In short, they must have deep knowledge of baseball! We hope the moral for educational research is clear.

By way of summary the Task Force offers the following, which applies to both of the empirical research traditions we are considering in this Report:

$$\text{QUALITY} = \text{fn}[\text{METHODOLOGICAL RIGOR} + ? + ? \dots]$$

Thus, while quality is a function of methodological rigor, it is also a function of the important other factors -- such things as scientific or educational/social relevance or fruitfulness of the questions being pursued, contextual understanding of educational settings and ecological validity of the concepts and data-gathering techniques that are used, epistemological sophistication,

sensitivity to value issues such as equity or fairness, clarity and completeness and validity of the case that is made to warrant the conclusions that were drawn.

The point we are stressing here is that whether or not a piece of research is of high quality is not determined by methodological rigor or depth *alone*.

Members of the TF have been gratified to learn that a group of experienced researchers working independently of us have reached a very similar position. The Spencer Foundation Exemplary Dissertation Award Committee reported in 2007 on the criteria it used to make its recommendations; the following are short extracts that make points that we strongly endorse:

- “What is an exemplary dissertation in education research? ... Five qualities stand out: originality, rigor, relevance, contribution to disciplinary knowledge, and clarity of expression.”
- “In short, rigor is not just a matter of design and analysis, but a quality of mind in which evidence and insights are carefully specified and intertwined.”
- “Relevance and rigor are sometimes cast as opposing qualities, but in an exemplary dissertation both are present. By “relevance” ... [we] signify a problem or question that is important, whose answer *matters* to those who care about education.”
- “While each of these [exemplary] works makes evident its original contribution, what they contribute *to* is as important as their originality. Each is grounded in and responsive to a disciplinary body of knowledge....”
- “Perhaps more than any other quality, the dissertations the committee identified as exemplary stand out for the clarity with which they were written. Each lays out its argument in plain terms, using technical language as appropriate but without excessive use of jargon that clouds its meaning. While the ideas are often complex, the language is pointed and concise.” (Gamoran, 2007)

## **6. Introduction to the vignettes**

We commence filling-in the details of how ARs can be educated in accord with the vision sketched above, by considering key events in – and key influences on – the intellectual development of two aspiring educational researchers Denise and Catherine. While these

individuals are fictitious, the intellectual histories that we depict are based upon our experiences in the RTG programs at our own institutions.

*At the outset we need to make several important points about these vignettes.* The first of these is obvious: no two individuals are representative either of the varied backgrounds of our graduate students, or the trajectories they take during the years of their programs. We do sketch the backgrounds from which these two students come (after all, the vignettes need some context), but they are only sketches and are not offered either as being “typical” or being complete intellectual biographies. However, the two programs we describe incorporate many of the current practices at the RTG institutions; indeed, the TF was expected to summarize and discuss what the RTG institutions have been doing to improve doctoral research training, and rather than providing unwieldy lists of innovations the TF members decided that these vignettes accomplish the same goal but in a more readable and informative way.

Second, the students we have created here in many ways are “too good to be true”. They never take “incompletes”, they never fail to meet deadlines, they always take the advice of their advisors, they eagerly listen to visiting speakers and they are inveterate participators in brown bag sessions. They never succumb to the flu, and never have to take time off to attend to an ailing parent. Not only, then, are they fictitious – they are perfect. Why do we depict them this way, instead of adopting the strategy of “warts and all”? One subsidiary reason is that if we were to give more details of the lives of our graduate students, our discussion necessarily would become so lengthy that it would tax the patience of our readers. The main point, however, is that these two vignettes are intended to serve as a record of some of the best RTG practices, and the account would become murky if we were to include even a subset of the myriad ways training can be derailed. We stress that we were not charged with producing an accurate portrait of all facets of graduate student life; and we are preparing the ground for a set of recommendations about training. In short, we see little point in documenting horror stories or in depicting “grim reality”, and in fact we assume that our readers are well-acquainted with this latter domain. As an aside, we also draw attention to the fact that we have been quite even handed, as (for the same reasons) we depict the faculty members with whom these students

come into contact as paragons of virtue, with none of the frailties that plague flesh-and-blood advisors and mentors.

Third, the vignettes serve as concrete starting points for our detailed deliberations about appropriate doctoral training for the new century. We make comments about the programs of study of Denise and Catherine in “boxes” along the way, and then, in subsequent sections, we analyze these programs from the perspective of the four overlapping universes we described earlier and highlight relevant features; finally, we pull-together our recommendations.

Fourth, although our charge constrains us to focus upon research training itself – as pointed out above, we were not constituted to write a treatise on all aspects of graduate life – we feel it necessary to emphasize the point that quality research is difficult to carry out (and a normal life is nearly impossible to lead) when one is faced by a host of near-debilitating problems. The life of a graduate student is not an easy one, and among the trials and tribulations that need to be faced are the following: lack of funding – or inadequate funding – and the economic hardships that ensue; availability of health care, for oneself and dependants; shortage of child day-care facilities; availability of affordable housing near campus; job and educational opportunities for spouses which is particularly a problem for foreign students; finding a professionally-suitable job upon graduation, the difficulties of which are multiplied for two-career families. There also are problems of integration into a research community, and these are particularly severe for women, members of minority groups, and for overseas students if there is no appropriate cohort or support group. The faculty of graduate programs often do not recognize these problems, and they need to be more than sympathetic – they need to be actively supportive.

## **7. Research training: Two vignettes**

### **7.1 The emergence of a quantitative-oriented researcher, Denise**

#### **Phase 1: Recruitment**

Denise, who had a strong undergraduate record together with a high degree of enthusiasm for pursuing a career in education research, applied for

Clearly Denise was exemplary in the way she researched her graduate program options. Unfortunately, experience at several of our Spencer RTG Schools of Education indicates she is not typical – many applicants are unfamiliar with key faculty members and their research interests, their statements of purpose are flowery and vague, and some applicants even apply to specialties that are not represented at the institution.

admission to a research-oriented doctoral training program. Although her interests were not yet formulated with any specificity, she indicated during her campus visit prior to submitting an application and in her “Statement of Purpose” that she had been stimulated by an undergraduate project on the influence of politics on school programs, which had led to concerns about the “achievement gap”.

Denise was a comparative rarity among the applicants in that she expressed having quantitatively-oriented methodological interests, and initially was attracted to the program by its recruitment literature that indicated that the goal was to prepare students for careers as faculty members at leading Graduate Schools of Education or as senior researchers at premier research institutes. She had also found appealing the recruitment brochure’s stressing that all students would be given an overview of the main alternative “research paradigms”. Furthermore, her browsing on the website reassured her that several faculty-members were working on topics in which she had some interest. However, her lack of practical experience as an educator was of concern to the selection committee, and led to a lengthy discussion.

Her undergraduate record – grades and internships, plus test scores – was strong enough that, on balance, she was judged to be a “promising prospect”. Indeed, she received offers from several programs, so to aid her in making a decision arrangements were made for a faculty member and a third-year doctoral student with roughly similar interests to call her. She found the conversations to be extremely helpful in developing a reasonable set of expectations.

The major part of their discussion, then, concerned what earlier was called “Universe 3” – they focused upon Denise’s substantive interests. Perhaps the advisor held the view that interests should determine the research methods, not vice versa!

### Phase 2: Joining a research community, and the shift from broad interests to research questions

Shortly after arriving on campus Denise had a brief initial meeting with her advisor, followed a few days later by a much lengthier one where they discussed in greater detail the tentative interests she had outlined in her “Statement of Purpose”, and they developed a plan for her first year of graduate study. They agreed that – as a major goal for the year – she

Denise was fortunate that her advisor was such a paragon, who in addition to other virtues recognized the importance of the early shaping of expectations. Evidently she was also modeling excitement about, and commitment to, research. Unfortunately, some advisors like this become inundated by “drift” from their less effective colleagues.

should gradually become immersed in the literature on the overlapping factors that influence student learning and achievement; and she should start keeping files on related matters such as how achievement is measured in schools, the connections between policy and practice, who is and is not succeeding at the level of the school and at the level of the child, what levers are available to the federal government to influence district and school level practices, and what we know about the relationship between federal money and student outcomes.

During this lengthy discussion her advisor, whose scholarship was related to Denise's focal topic, noticed that the proposed program of study for the winter quarter probably had a couple of unallocated units, so she offered to supervise a "directed reading" for Denise – specifically she could read articles that analyzed the different ideological and political pressures that shaped the allocation of federal money in education. At one stage in their discussion her advisor showed Denise some preliminary data that she had just collected on school achievement, and drew her attention to the fact that scholars from different research paradigms were likely to interpret these data in divergent ways. At the end of this discussion they decided that, as a part of Denise's research assistantship, she would participate in the project from which these early data had been taken.

During the past decade, campus-wide surveys at several institutions housing RTG Schools of Education have indicated that the quality of advising is a major concern to a small but non-trivial number of graduate students. There have even been students who felt "trapped" in an unhelpful relationship about which they could not complain.

As a result of both this meeting and the social activities during orientation, Denise was excited by the prospect of attending the (mandatory) first-year doctoral seminar, for she sensed that she had much to learn from the other members of her cohort -- who came from all points of the methodological and ideological compass, and who already had impressed her with their range and depth of educational experiences. This seminar, which was co-taught each year by two professors who were versed in different research traditions, was listed in the catalogue as "Spencer first year seminar: Introduction to research methods and research traditions". Designed to provide a springboard for further coursework and research internships, a number

Many Spencer RTG Schools of Education have mounted a special first-year seminar along these lines (one institution has offered this seminar in the second year). One problem that emerged was that instructors changed frequently, and with each change there was a change in the syllabus. Longer-term faculty commitment seems highly desirable. Some course descriptions may be found on the websites of the Spencer institutions (which are listed on p.5 of this Report). It should be noted that some TF members are skeptical about the usefulness of seminars that are *solely* "show and tell".

of faculty guests appeared during the quarter to present a piece of their research and to face lively questioning.

In the Autumn quarter Denise also made a start on her “distribution requirements” by registering for a class in the history of American education, that had a politics of education focus (this was an intermediate class, since she had a class similar to the introductory class in her undergraduate studies). And as she had taken some low-level statistics classes as an undergraduate she felt it important “to keep her hand in while the iron was still hot” (as she put it to her advisor, who uttered a soft “tut tut” at her mixed metaphor), and she found an appropriate level course taught – of all places – in the Psychology Department. (This appealed to her because it used illustrative examples taken from psychological research, which she expected to be more relevant for her than the examples her experience had led her to believe would be used in a parallel course in the Statistics Department.)

As the first quarter progressed Denise found herself – in addition to meeting regularly with the team on the research project where she was an RA – becoming more active in an informal five-person “collaborative reading and writing group”. The members came from across the intellectual map, and their coming together initially was fortuitous – they had been randomly assigned together as a “one shot” group to work on a classroom assignment for the Spencer seminar. However, their discussion carried over later to the campus coffee shop, and soon they all had a firm commitment to meeting at least weekly. Gradually the members shared drafts of their writing, and they distributed for discussion occasional journal articles that they had found stimulating. Later in the year the girlfriend of one of the members joined the group; she was a third-year student in the School’s “Program and Policy Evaluation” area, and late in Spring presented to the group a draft of a dissertation proposal that received vigorous critical scrutiny from the group.

In some universities there are administrative, and even “ideological”, barriers to cross-School enrolment. In places where the boundaries are permeable, education students have expanded opportunities to receive rigorous training and even mentoring – many “outside” faculty are happy to serve on dissertation committees and the like. Clearly such types of cooperation are to be encouraged.

Besides Denise, the other members of the group were doctoral students from developmental psychology, economics of education, secondary teacher education, and anthropology of education. This latter student had, as well as background in ethnography, a grounding in feminist theory, and like others in the group she was politically active. Over time her views stimulated Denise’s intellectual development, although they never saw things completely “eye to eye”.

As the first two years of Denise's graduate studies progressed she took more classes in her content area, some further statistics and research design classes, and she met periodically with her advisor. But she was well-aware of the central role played in her transition to being a researcher by two groups: the team on the research project where she worked as an RA; and her co-operative reading and writing group (of which she rarely missed a meeting). Thus there was a variety of ways Denise was able to get honest and sometimes surprising feedback on what she was reading and writing, and she was regularly exposed to material in which other students were interested and with which she would not otherwise have come into contact.

One Spencer RTG Institution planned to make membership of a writing group mandatory, but it was discovered that groups had formed spontaneously and the majority of doctoral students already were members of them. Doctoral students highly prize them. Another example of a student-initiated program at a different institution – later taken up by the faculty – was journal reading club in which students critiqued empirical and theoretical articles across an array of topics; this earned academic credit.

Her advisor, from the beginning, had stressed that Denise read research papers with an eye toward understanding the results and conclusions, but also toward scrutinizing the warrant the researcher offered by way of justification of the design, the data-gathering and analysis, and the conclusions that had been reached. (In fact, she had become so interested in the notion of “warrants” that, on the recommendation of one of her writing group members who had a strong grounding in the humanities, she took an introductory class on “Epistemology of Science” offered in the Philosophy Department; she was pleased that her advisor, after a short period of bemusement, became quite supportive and even asked to see the class reading list.)

The TF identified seven broad categories of skills required for doing quantitative research, in addition to the body of necessary substantive knowledge about the relevant area—(1) defining the purpose of the research; (2) framing research questions; (3) developing defensible study designs; (4) conducting statistical analysis; (5) interpreting and writing about results; (6) presenting findings for various audiences; and (7) protecting human subjects.

In addition to these activities, through her apprenticeship as an RA Denise began gaining experience in collecting data in schools. She took on more responsibility and worked with a faculty member and senior graduate student on gaining access to schools and working with principals and teachers to implement their research. The PI (who also was her advisor) held frequent research meetings and stressed the importance of respecting

Some RTG Schools of Education offer such a “research epistemology” course “in house”, and some have given this topic a central place in their Spencer seminar. This course is ideal for team teaching, pairing an empirical researcher with a philosopher of education.

and communicating with the stakeholders in schools—administrators, teachers, students, and parents—and as a part of this training Denise and the other RAs wrote and rewrote, and practiced and critiqued all communication with schools. In addition, in discussing how to gain access to schools a number of issues can come up, so Denise took a mini-course on research ethics that was offered on Saturdays.

Thus, during this early phase of her training Denise developed a knowledge base on her topic of interest and began to develop research skills and knowledge through courses, individual reading, and through her apprenticeship. She began thinking about research from a variety of perspectives and came to realize that understanding these perspectives would help her to more fully develop her own research. A milestone was reached when her advisor/PI asked her to take the lead role in giving the whole research team an overview of some new data that she and Denise had collected.

### Phase 3: Developing expertise and research skills

At about the end of the second year Denise moved into a new phase of her doctoral program. She and her advisor, through meetings and discussions, refined and focused her interests. Her initial interest in asking questions about why federal policy hasn't or can't fix failing schools changed into how educational policy is implemented in schools and classrooms and how this affects students of different backgrounds. Her reading became broader as she looked at related research in child development, teaching and learning, and measurement, and she realized that this issue could be investigated from different methodological approaches (including "mixed methods" research).

Understanding that she needed to be trained in depth in at least one approach to research (she, and her advisor, were hesitant about referring to "research paradigms", especially after her class in philosophy of science), she continued to take statistics classes; she had already worked on correlation and regression, and also had taken an experimental design

Denise's substantive interest in the policy formation process and policy implementation could well have been developed with more sophistication had she been able to take either an MA or a doctoral minor outside Education, in for example the Political Science Department. "Outside" study of this sort is required at a few RTG Schools of Education. This practice also fosters closer relations between Education and the Social Science and Humanities departments.

Several RTG institutions are trying to break the expectation that all training be done via traditional courses. "Just in time" intense workshops (sometimes held in conjunction with a large research project when RAs need to be trained), and short summer workshops prior to the start of the academic year, are among the innovations. Instructors should get credit for conducting such modules.

course in the Psychology Department, and now she took multivariate statistics with a small group of other advanced students. She attended an informal workshop for non-measurement specialists on reliability, validity, and equating, to assist her in understanding testing in school environments.

Although Denise was committed to using quantitative methods in research, she clearly recognized the benefit that would accrue from her having at least some background in the theory behind qualitative research, but also in its data-gathering techniques, and methods of data analysis and reduction. She was able to attend some “just-in-time” workshops on these topics that her School had recently introduced on a trial basis (one of these was held over a long weekend).

Several things had sensitized her to the need to have at least a minimal familiarity with qualitative research methods. First, she had come across some studies in her areas of interest that were qualitative and even ethnographic in nature, and which revealed aspects of phenomena that her own methods were not able to pick up. She had been particularly excited by McDermott’s study “The acquisition of a child by a learning disability” (McDermott, 1993), which had been carried out largely in a family’s home. Second, the comments of her ethnographic friend from the writing group, about this article and about Denise’s own work that she had shared with the group, convinced her that rigor was not

In conversations leading to the establishment of this Task Force, the deans of the participating Schools of Education agreed that doctoral students should be trained to become rigorous developers of knowledge who were steeped in one of the major research traditions – but, crucially, that they should also be educated to the “informed consumer level” in one or more major alternative traditions. Members of the TF share this view.

confined to the use of tests and statistical analysis. Third, she realized that much of her future research would be carried out as part of a team (hopefully later in her career she would lead a team of her own), and therefore she needed to understand the contributions that individuals with different research competencies could make to the team’s work. And finally, she had been stimulated, overwhelmed, and disturbed by her first visit to an AERA convention the previous Spring. She did not make a presentation at this meeting, and mainly invested her time in attending keynote addresses by researchers whose work she knew but whom of course she had never seen “in action”. At one of these she had a “peak experience” -- she had initially been horrified by one famous scholar’s very fluent advocacy for the acceptance of a novel as a

doctoral research dissertation; however, on the flight back to campus she decided that being horror-struck was not a good-enough response. This led her to do some deep reflection (without, alas, coming to any firm conclusions) on the differences between quantitative and qualitative research, and between qualitative research and fiction. As a result of all these influences, in addition to the “just in time” workshops mentioned above, she took “Introduction to qualitative research” (which, in addition to its theoretical orientation, involved doing a short observational study in an elementary classroom). She audited a course on ethnography in the Anthropology department. In addition her collaborative writing and reading group spent several sessions discussing the classic book by Clifford Geertz, *The Interpretation of Cultures*.

In this phase of her graduate study she also took on more responsibility in her assistantship, and obtained well-mentored experience in data cleaning and linking, and in data analysis. She also participated in meetings where the work she was doing was discussed in the context of the larger studies, and she was responsible for situating her work in the scholarly literature.

All students in her graduate program were eventually required to plan, implement, and write-up a pilot research study that would provide a base for their future dissertation work. Denise, who had been reading widely, began a more formal literature review that focused on how her questions fitted into the larger literature and which methods were appropriate for answering these questions. During this process she realized that her ease of access to her advisor was a godsend, and she also profited enormously from the comments of her co-operative writing and reading group. As her formal course-work was essentially completed, she was able to devote substantial blocks of time to data-gathering and analysis, during which a number of unexpected problems arose that caused her (after numerous consultations) to revise her design, and to lengthen her research timetable for the final full study. She was consoled by the realization that this is what pilot studies were for!

Many of the RTG institutions offer a “dissertation proposal writing” seminar-workshop; indeed often there are several, offered by discipline or program because the specific issues differ across fields. Although student numbers in these workshops are small, the work (and the teaching) is always extremely demanding.

After seeing the CVs of several students who were a year or more ahead of her, she became aware that in order to find a suitable job she needed to make a few conference presentations

and begin a program of publishing. So she decided to “kill several birds with the one stone”. She wrote-up a version of her pilot study that was suitable for presentation at her School’s annual “graduate student research day”, and she also crafted another version which she submitted for the forthcoming AERA convention. (She was slightly nervous about the latter, but here her experience at the previous conference proved valuable – she realized she could do at least as good a job as many of the panelists she had seen there.) She also started working on a related article that she planned to submit to a new journal she had come across in the library; it seemed to be of good quality, but because it was new she thought her chances of acceptance would be increased. She also asked her reading group members to act as an audience for a trial presentation of her material after she learned it had been accepted for the graduate student conference. She would not hear about AERA for several more months.

Student research conferences, with other students and faculty members acting as respondents and advice-givers, are now held at most RTG institutions. Presenters not only get advice on substance and methodology, but also get helpful input on their modes of public delivery and question-handling strategies.

#### Phase 4: Working on an interdisciplinary project and writing a dissertation

As Denise was looking for new opportunities to work as a research assistant (her previous project had come to a conclusion), she learned from one of the students in her cohort about a small team of faculty and graduate students who were part of a national study of distributed leadership. She had first learned about research on distributed leadership in one of her courses; she learned more about its application to school improvement during a student’s presentation at an informal research seminar.

At a meeting with her advisor, Denise talked about how work with this research project might give her opportunities to study in greater depth the implementation of school reform, and also to gain more experience working with teachers and school administrators. Her advisor suggested that the project might also offer a setting for carrying out her own dissertation study.

After an interview with the principal investigator of the study, Denise was offered a position on the project team. She was asked to take a leading role on a component of the project that would construct and administer surveys for describing the social networks within and between

schools. Survey results would be used in a broader analysis of the role that social networks played in school reform, and ultimately in changing patterns of student achievement.

Denise recalled hearing something about social networks in her early coursework, but knew little about recent work on the topic. The PI of her new project suggested that she enroll in a four-week workshop on the measurement of social networks, which was being offered over the Internet by a leading expert. He also recommended some recent articles on underlying theories of social capital and its role in school learning.

As Denise began reading the theoretical literature, she could see the value of the studies that built on survey data, but she was skeptical about the conclusions being drawn in studies that used field observations in schools. She worried that the investigators might have a “confirmatory bias”, and would (wittingly or unwittingly) select examples that supported their opinions, while counter examples were being ignored; she also had the realization that bias also could infect the early stages of survey development. She brought up her concerns in a project meeting, which included a few faculty members and several other graduate students. The project team had some people from her doctoral program, but also included faculty and students from programs in anthropology, school leadership, and curriculum. Her comments spurred a lively give and take, which surfaced a variety of issues to consider in judging the validity of the social capital studies, but also brought in issues related to other research that had shaped the project, including analyses of some national surveys, descriptive accounts of other reform initiatives, and a couple of experimental studies comparing the effects of different approaches to school leadership. Denise came away from the discussion with a better sense of what she could take from the theoretical work and from the prior empirical studies.

As the quarter progressed, Denise worked both on the social network survey to be used in the project and on a proposal for her doctoral dissertation. Her dissertation would build on those survey results, but also would make use of student achievement results collected elsewhere in the project to test hypotheses she was developing about the relationships among teacher networks, principal leadership, and student achievement. During this period, working closely with the chair of her dissertation committee, she applied for IRB approval of her project. Then,

to pilot test the survey, which was an adaptation of an instrument used in other studies, she went to several of the schools participating in the study, trying to persuade school staff to take the time to fill out the survey and to talk to her about what they thought the items on the survey were asking. The initial school visits were rather tense, because some of the teachers were suspicious about the purposes of the research and how the results would be used. To get help in making these visits go more smoothly, Denise talked to her advisors, project team members, and some of her fellow graduate students who had more experience working in schools.

To make a long story short, the dissertation finally emerged and was successfully defended. Denise experienced the alternating stages of exhilaration and depression that accompany sustained effort over many months being spent on a single (and all-consuming) project. She also made the common mistake of not allowing sufficient time for polishing and editing her chapter drafts and as result, after her oral, she had to spend another three weeks on these activities. While all this was going on, she had a potential publication to work on (it had been returned from a refereed journal with the recommendation “revise and resubmit”), and she had to prepare for several job interviews.

By the time she had her oral exam, Denise was a competent quantitatively-oriented researcher, who had received an impressive training in depth as well as breadth. *Her competencies, and how she obtained them, are summarized in the table that follows the second vignette; this table also facilitates a comparison between the training received by the two students featured in these vignettes.*

Some RTG institutions brief their ARs on the need to publish, and on the intricacies of the submission, review, and publication process. But many students remain ignorant and misconceptions abound, for example about the large time lag between initial submission and final publication.

#### Phase 5: The first job

Denise felt fortunate that eventually she received two job offers – one in an education leadership program in a research-oriented School of Education, and one in a policy research firm. As she considered which offer to accept, she thought about what each would allow her to accomplish and what she would have to learn to succeed in each. The education leadership job would involve teaching and working with graduate students, some of whom would be preparing for research careers and some who aspired to be school administrators. It would also allow her the freedom to shape her own research agenda. On the other hand, in the policy

research position she would devote all her time to research, with greater emphasis on writing proposals for funding and more frequent meetings with policy makers. Her work would call for getting quick answers to questions about the consequences of competing policy options. After deliberation, and a few discussions (over a drink) with her advisor, she chose the university job.

## 7.2 The preparation of a predominantly qualitative researcher, Catherine

This vignette, like the one dealing with the career of Denise, describes a program that incorporates many features found at RTG institutions; as explained in Section 6, Catherine (like Denise) is idealized and is depicted as being extraordinarily receptive to the opportunities that arise during her graduate career. We have given her some distinctive background characteristics and experiences as a reminder that students do not enter graduate school as blank slates, and to challenge some stereotypes.

Members of the TF believe that at admission time, faculty should be proactive in analyzing the intellectual profile – the particular combination of strengths and weaknesses – of each candidate. We address this in a little more detail in Section 8 below.

### Catherine's Background

Catherine was a first generation college student; good at math and science, she qualified for the main campus of her local state university. She chose to major in chemistry, chiefly with an eye to obtaining employment after graduation. However, some other aspects of her undergraduate coursework had great impact, and sensitized her to the insights into human motivations that come from literature and other disciplines; and in her junior and senior years she enjoyed the very different pedagogical styles she encountered in several elective courses in literature. These varied classroom experiences sparked her initial interest in teaching.

At the end of her college career Catherine found the prospect of working full time in a chemistry lab unappealing; and after exploring options she was accepted in "Teach for America". When she was assigned to teach general science at a middle school in a large urban center, she found herself in a new world. Many of her students were first or second generation immigrants from Mexico and Central America, and there were a few from Asia; most of the remainder were from a local area that was almost completely African American. She quickly realized that the brief training in pedagogy and classroom management that she had received was inadequate for the daily challenges she faced in such a diverse setting.

After three years, and the masters level training she pursued at night and in the summers, she began to feel somewhat more competent in the classroom. But at the same time she was becoming aware of problems that were generated by the structure of schooling itself. Eventually she decided to apply to Schools of Education to see if she could better understand the issues that were of growing concern to her.

### Phase 1: Introduction to Graduate School and finding a direction

Catherine was accepted to the Educational Studies Department at Flagship State University, a public university with a large and prestigious school of education. The admissions committee was impressed by her teaching experience, but saw her lack of background in social science and related educational theories as a problem. However her high college grades and her GRE scores, and strong letters from colleagues who praised her work in her school and from the professors in her masters program, convinced the committee that she would be a quick learner. It was noted that although she had substantial training in natural science, she would likely need significant social science research training.

Catherine was assigned to an initial advisor, a sociologist of education whose own research used mixed methods to understand classrooms serving varied student populations. He drew on large data sets, did his own surveys of students and teachers in fairly large numbers of schools, but, with a team, also used qualitative data to look in depth at some classrooms and the schools in which these classrooms were embedded.

Catherine and her advisor emailed about her program in the late spring after she was accepted. Since she was coming to campus for the summer, he advised her to start the three course sequence in statistics with the first course, which was the only one given in the summer. Taking a course would allow her to get accustomed to being a student again and give her some connections while she searched for work. Her advisor was away for the summer, but they were able to email and then to meet before classes started in order to map out not only her

Many education graduate students have backgrounds similar to Catherine's. She has impressive school experience, but her academic training was light with respect to social science research-related skills. By comparison, some students in social science departments are weak with respect to education-related experience. Given the disparity in background of education students, teaching of research is a pedagogical challenge, and use of realistic education examples is a necessity.

courses for the first semester but an initial plan for her work that encompassed the first year and somewhat beyond.

Catherine's advisor stressed that she was about to undergo a change in roles – from being a classroom teacher to being a researcher who will bring social science theories and methods to bear on educational issues in a sensitive way. This new role necessitated that she acquire a good deal of knowledge of the research findings about schools, classroom relationships, and the daily lives of children in poor urban neighborhoods. It also demanded an appreciation of the theoretical perspectives through which social scientists seek to develop a systematic understanding of these issues. Because her background largely was in natural science, she would need to learn the varying ways that social scientists think and do research.

Her advisor suggested that because she had not been immersed in either the systematic study of schools or the social sciences, she should start by taking courses that would increase her substantive knowledge but also expose her to an array of disciplinary approaches. Although Catherine wanted to get straight to work on a program that would train her to do a specific kind of research, and although she was concerned that her program would be lengthier than she had imagined, and thus more financially draining, she acquiesced.

Catherine, like Denise, is blessed with having a conscientious and sensitive advisor. Many students with backgrounds like Catherine's do not yet know enough to select an appropriate methodological frame within which to pursue their substantive interests (which of course also are subject to development).

Together Catherine and her advisor selected courses that would develop her research-based knowledge of educational institutions and their surroundings and also would introduce her to several relevant disciplinary approaches. Thus, in her first year, in addition to the statistics sequence she took introductory courses in History of Education, the Political Economy of Education, Sociology of Education, and Anthropological Studies in Education.

Students' financial situations often impact their intellectual decisions. Thus, we see Catherine preferring to move ahead more quickly with her training, but ultimately ceding to her advisor's recommendation. Advisors should remember that students accept their advice for reasons other than their understanding of, or agreement with, the advice. The acceptance of advice may instead reflect a power differential.

Both she and her advisor thought that a position as project/research assistant (RA) on a project headed by a prominent educational psychologist on teaching science and math to English

language learners in elementary classrooms would be a good fit for her, and she needed the financial support. Catherine's advisor encouraged her to apply because of her background in teaching science and work with English language learners in her middle school classrooms. It also seemed likely that this project would introduce her to psychological approaches to education – a discipline in which she was not initially taking a course – and to issues of second language acquisition. Further, it would introduce her to quasi-experimental designs, and she would learn many relevant research tools.

In the fall, Catherine enrolled in a one credit course required of all first year students in her department/division, in which the group read a sample of work by every member of the department, wrote individual comments and questions about it to be read by the facilitator and each week's author, discussed it, and met with the professor to discuss both the particular piece and his or her wider research and teaching interests.

This course helped to acquaint Catherine with the approaches to the study of education that were available in her intellectually diverse department, and also gave her a chance to become intellectually and socially acquainted with a cohort of students having a wide range of interests. It was the faculty's expectation that some members of the group would have interests that would re-enforce hers, while others would have challengingly different views.

In their first conversations, the advisor had strongly recommended that Catherine take the full three semester statistics sequence offered in the Sociology Department, which used many examples drawn from education (including the study of classrooms). The first course was quite elementary and thus possible to absorb in a summer course, especially for someone with Catherine's strong mathematical background.

The TF members believe it is a serious mistake to overlook the educative potential that comes from having a cohort with diverse backgrounds and interests. However, instructional methods need to be designed to capitalize on this diversity, by fostering fruitful modes of interaction.

At many universities it makes more sense for certain students to go outside of Education to take statistics because they need to match the type of statistics they learn to the type of research questions they are likely to investigate (e.g., Catherine took her statistics courses in sociology in order to focus on groups).

As she enrolled in the second course in the statistics sequence, Catherine did ask whether she should wait to invest three semesters' of effort into learning stats until she knew how much she would be using quantitative techniques. Her advisor replied that whatever the methods she eventually embraced for her own work, she would need a strong background in statistics in order to be able to read research in the dominant traditions in educational research both with deep comprehension and an ability to be critical. Her advisor also suggested that she hold off on other research methods courses because she would get more out of them after she had read more social scientific research in her initial courses, and after she had a better sense of the research questions that she would be pursuing.

A number of RTG institutions now require that *all* doctoral students (even historians and philosophers!) take statistics at least to the "literate" level; some also require that at least introductory-level work in qualitative inquiry be taken by all students. The justification for these requirements includes the argument that students need to be literate and able to follow the research literature.

Catherine found her first year exciting, but initially she felt pulled in many directions. The departmental pro-seminar opened her eyes to many exciting possibilities for further study, and, despite (or because of) occasional heated debates, the conversations with her peers provoked her to much new reflection. As the semester went along the two substantive courses in History of Education and Political Economy of Education, which rounded out Catherine's fall semester, gave her a new way to think about the behavior of groups of people – less as a matter of individual characteristics and decisions and more in terms of social patterns and social influences that become part of individuals' perspectives.

Catherine heard of a course in the Philosophy Department on philosophy of social science that treated some relevant issues, but timetable clashes prevented her from taking the class.

It was in the spring semester, as Catherine was taking courses in Sociology of Education and Anthropological Studies in Education as well as continuing her work with the psychologist's project on second language acquisition that the disparate threads in her graduate education began to come together. Catherine came to appreciate the power of the social patterns that statistical analysis can reveal. At the same time, she found herself sometimes frustrated with the approximations of important social concepts represented in the indicators available through large data sets.

Catherine was somewhat unusual in her ability to question and to put together her diverse experiences with a high degree of independence. Some students might have felt more overwhelmed than invigorated by the diversity of intellectual experiences that Catherine had. Advisors need to monitor student capacities to process the experiences they are given and tailor their advice accordingly.

Catherine's assistantship on the project about math and science teaching with ELL students also challenged her to think further about the interplay between question, method, findings, and the larger universe of research questions as well.

Although she learned a lot from the spring course in the sociology of education, it was in her Anthropological Studies in Education course where Catherine finally discovered the intellectual home that she had sought. The class read and critiqued a series of ethnographic studies of educational settings; Catherine read these with mounting excitement, for they explored the political, economic, institutional and social processes that shape educational processes and situations, but they also showed how groups of people make collective sense of the social and educational settings in which they find themselves. She was excited that the course provided her with a disciplinary lens through which to view these issues.

But despite her excitement, as she placed these ethnographic studies next to what she had learned about precision and consistency of measurement in her work as a project RA, and next to questions of sampling and descriptive generalization learned in her sociology and statistics courses, Catherine was concerned. She questioned whether her excitement over them was valid – whether they were subjective narratives rather than social science. How could one trust the accuracy of the description or the logic of causal connections drawn? How could one generalize at all beyond this particular time and place?

Through the course of the semester, the professor allayed Catherine's fears and showed the class how ethnography had its own standards that were just as rigorous as those she had encountered in her statistics class. For example, she demonstrated how researchers could discipline their own subjectivity and be systematic in the search for disconfirming evidence – both of which increase reliability. She also addressed questions of validity with small samples; she argued that ethnographic generalization is more often theoretical than descriptive.

Catherine's change in advisor was painless, taking place as it did in a context where everyone knew her former advisor was "initial". Where this understanding is lacking, change of advisor can be painful, even though no one profits if the intellectual interests of advisor and student do not match. Students with advising problems are hesitant even to use their institution's ombudsperson service, for fear of reprisal.

As a result of Catherine's newfound affinity for anthropology, after further thought she approached her anthropology of education professor who agreed to be her new advisor. Fortunately, her old advisor was completely supportive, agreed to remain on her committee, and, in essence, functioned as a co-advisor. In consultation with both her new advisor and her previous one, she planned her program for the following year. For the summer following her first year, Catherine accepted the invitation of the psychologist leading the language acquisition project to take a position in the summer to work on a jointly authored article from their work during the previous year.

In summary, Catherine's first year experiences were the following:

1. Substantive courses: history of education; political economy of education; sociology of education; anthropology of education.
2. Methodological courses: year-long sequence in statistics.
3. Department-wide pro-seminar: pro-seminar introducing members of the entering cohort to each other and to the intellectual work of the faculty.
4. Research assistantship: assistant on English language learning project directed by a psychologist.

Co-advising is a practice at some RTG institutions, one that may benefit certain ARs by allowing them to better balance the expertise of two important advisors. This practice also may provide intellectually rewarding partnerships for the advisors.

### Phase 2: Advanced work

Catherine and her advisor planned out a program for her final years of course work that would advance her substantive knowledge, increase her sophistication in the fundamental premises underlying various kinds of research, and hone her qualitative and anthropological research skills.

Clearly, in support of her decision to approach educational phenomena through an anthropological lens, she needed to take more advanced courses in anthropology or sociology that leaned in a qualitative direction. She took an external minor— required of all doctoral students at her university— that was “distributed”, meaning it included courses from more than one department outside her own. She took several courses in anthropology and two in sociology. As Catherine pursued this

Her university required all doctoral students to develop a minor in a discipline or disciplines outside their home department.

minor in anthropology and sociology, she was growing increasingly interested in the effects of social class, and race/ethnicity, and to a lesser extent gender, on classroom relationships and on the ways that students understand themselves as learners, so as much as possible she directed her coursework toward these topics.

In her own department, she took a course on ethnicity and education and she also took a course on non-traditional methods of teaching science jointly taught by professors in the departments of educational psychology and curriculum and instruction.

Most of her second year course work was taken up with two year-long courses. The first of these examined the assumptions underlying the variety of disciplinary and methodological approaches – it was a course roughly similar to the philosophy of social science course that she missed in her first year. It began with inquiry into the philosophical underpinnings of research, then asked students to study ontological and epistemological assumptions that often are at issue in social science inquiry. It then gave an introduction to specific research approaches frequently used in social scientific research in education, laying bare their assumptions and asking what is revealed and what hidden by different approaches. Finally, it asked students to read exemplary work on related substantive topics undertaken through different research approaches, and to discuss the underlying philosophical assumptions in each and the gains and losses for the study of education that seemed to follow from those assumptions.

The discussions were lively as students questioned each others' tacit philosophical assumptions in the early part of the course and then each other's favored research paradigms in the later part. Serendipitously, the discussions in this course helped to prepare Catherine for the experiences she had upon attending her first AERA conference in the spring – for at this conference, she heard a panel of international scholars argue that educational research was a mistaken enterprise that could only turn up trivial results and therefore should be abandoned; and she

Many Schools of education, even if they have a philosopher, are unlikely to have one who is specialist in philosophy of science. However, some top-flight researchers are well-read in the relevant issues and have even contributed to the published literature on the “paradigm wars” and other controversial issues, and could direct a reading group.

The TF believes it is vital for young researchers to be introduced to the live, often heated, fundamental debates concerning the nature of research (Universe 1). They cannot remain with “closeted” beliefs, as Bruner put it; but there is also a substantial motivational benefit that accrues from appreciating that one's field is intellectually vibrant!

attended a session where a presenter and the respondent had a lively argument about whether agreement between observers was an indicator of truth, and also whether consistency within an account of life in a classroom was such an indicator.

In her second year, too, Catherine began her formal training in qualitative research methods with a second year-long course. It began with study of the fundamental presuppositions of such work and the kinds of research questions and research findings to which qualitative work is suited; the course then progressed through data collection in field settings, to data reduction and analysis, and finally, writing. The course culminated with the students individually conducting a small-scale study of their own devised in a setting such as a local school, community family services center, or a dormitory on campus. (The professor had discovered that this assignment was a crucial one for the course, and that students also learned a great deal from discussing their project with the class.) Catherine found that locating a site where she might see the processes she was interested in proved to be difficult, but after two or three tries she found a school fifteen miles away that was willing to let her do her study.

The professor also explained human subjects review in a class session, and read drafts of the applications they were submitting to the IRB committee, on which she had served. Catherine submitted her own IRB application and, like the majority of her classmates, had to revise it in response to queries from the committee. Once in the field, Catherine and her classmates kept discovering new substantive issues they had not anticipated. The class-members also shared dilemmas that arose in making observations and conducting interviews.

Catherine discovered that her class-work in the methods class and the philosophy of science course, along with her assistantship, led to deep discussions with her

Students are often each other's best resources and to the extent that schools can structure situations to allow more senior and more junior ARs to interact, it is likely to be beneficial for all.

The skills of qualitative research are best learned through as much actual student research as possible. While faculty do need to have student read and hear general expositions of methodological points, generally students only really comprehend and internalize these principles when they practice them amid the complexity of actual data gathering and analysis. The statement, "I did not really learn it until I went through the process." is commonly heard.

At some institutions taking a concurrent MA either internal to the School of Education or outside it, may not be possible. The point however is that Catherine had to undertake several short research projects before starting her doctoral dissertation, to ensure that it was not her first piece of independent work. In fact, Catherine had a preliminary trial run through her methods class, then a second attempt for her MA thesis, and then she proceeded to her dissertation.

fellow-students that forced her both to question and to defend her emerging ideas about research. She came fully to the realization, initially presented to her in her first anthropology course, that the researcher's own social positioning and social assumptions have to be considered – and disciplined – in any kind of social scientific study.

Separately, Catherine's program required that she be responsible for undertaking a small scale original study of her own and writing it up as an MA thesis. Satisfactory completion of this master's study was a prerequisite to formal admission to the doctoral program, although the work on the two was co-terminous. While this work was ideally to be completed by the end of the second year, Catherine, like many other students pursuing the doctoral degree, did not find time for it by then. She planned it and wrote a formal proposal for it while doing her study for the qualitative methods course—and in fact she was able to incorporate some of the field work from that study into the masters study. She drafted the introduction and literature review over the summer, returned to the same classrooms in the fall, and with an enormous push, finished the Master's early in the spring semester before her course papers started to come due.

Catherine's MA thesis was, as her advisor predicted, an imperfect product in which, with the wisdom of hindsight, she could easily see that she had made several mistakes in planning and execution. But simply seeing this taught her a lot and made her realize that her insight and research skill had improved as a result of this educative experience. She also found that she became both more critical and more appreciative of the scholars she read. Further, despite the flaws in the work, she was able to shorten it into a paper to be given at a roundtable at the American Educational Research Association in the spring of her fourth year.

Catherine's course work included an advanced course in qualitative methods in the Sociology department, specializing in varieties of interviewing. The course included discussion of studies using interviews of different kinds—such as life history interviews, interviews with children, and interviews seeking to

get quantifiably comparative answers to open-ended questions. Students chose a single broad subject area, crafted interview guides for these different kinds of interviews and subjected them

Many ARs go through a phase where they are exceptionally critical of most of the research that they read (and perhaps end up "throwing the baby out with the bathwater" a little too frequently). Others are insufficiently critical at the outset of their training. One benefit of some of these early research projects is that students often become more balanced consumers of others' work.

to group critique, then conducted practice interviews, wrote up transcripts and initial analyses, and subjected these also to group critique. The course closed with treatment of different schemes for coding and analysis of bodies of interviews, and the management of qualitative data.

Concurrently with her coursework, for the fall of her second year of graduate work, Catherine accepted a position as an RA on a two-year project under the leadership of her original advisor, in which she was to be one of two students responsible for producing short case studies of classroom interaction in several large city schools that had student bodies with different ethnic compositions. The larger study planned to use a quantitative data set to examine a variety of variables and to select some schools for closer examination.

As Catherine became increasingly intrigued with, and skilled at, the kind of analysis that requires ethnographic data, she also saw that her science background and advanced math skills, along with her experience on the psychology project, were good preparation for conducting mixed-methods research. Her work on paid projects also gave her a chance to make research presentations at professional meetings. In the winter of her second year she attended, as a non-speaking co-author, a conference where the paper from her first-year project was presented. She also was present as contributing member of the staff when her first advisor gave an initial paper out of their project at the American Sociological Association at the beginning of her third year. She gave a short part of the presentation when he presented a more substantive paper at the beginning of her fourth year.

Finally, also in her second year, Catherine started participating in a DAG, a Doctoral Apprentice Group, in which her advisor brought together on a weekly basis all the students on whose committees she served. They took turns reporting on their progress and sharing their dilemmas. The more advanced students had many questions and suggestions for the less advanced students like Catherine as she moved into her Master's work. Catherine came to see that this group extended the work of the qualitative research course, giving her a chance to

In learning qualitative methods, group discussion of research dilemmas (such as those in the DAG, and to a lesser extent those in research classes) is very important to developing advanced skills. Good practice is often less a matter of following rules, and more about good judgment in complex situations that are never fully replicable. Discussing how others solve research dilemmas and discussing one's own dilemmas is invaluable in developing good judgment. The DAG was thus a critical part of research training.

learn more from the other students about question formulation, fieldwork dilemmas, and the complex process of analysis.

Catherine continued to participate in the DAG even after her Master's thesis was complete as she finished coursework and prepared for her comprehensive examinations. In her third year of graduate work, she also worked for a second year on the project with her initial advisor, completing field work in the fall and participating in the writing of their final research report and working with him and another graduate student who studied different schools, as second author on a paper about the case study schools.

Thus, after her initial broad intellectual survey of the field of education, she sought more substantive intellectual breadth, even as her own specific interests became more defined and research oriented.

By her third year, although her time was limited, she attended most of the monthly lunch-hour Brown Bag sessions that her department sponsored, where faculty and advanced graduate students talked about their research. She also made it to some talks in Anthropology when the topic seemed reasonably close to her interests.

The situation at many RTG institutions is that informal venues – brown bags, writing groups, “fireside chats”, colloquia by visitors – are a major resource for students, who often learn things from these sources that they can't learn elsewhere.

In her third and later years she also became a regular participant in Brown Bag sessions offered by a research center that provided an umbrella for most of the research in the School of Education. It consequently provided support services for a wide variety of research projects and for members of most departments in the School of Education. These institutionally-supported sessions dealt with professional development issues for budding academic researchers, running the gamut from how to write a paper proposal for a professional meeting, to delivering the paper effectively, to tips for effective networking at professional meetings, to how to get access to your advisor's time and to use it efficiently, to the peer review process at academic journals, to the Institutional Review Board process for research on human subjects.

RTG institutions have found a variety of ways to formalize the transfer of this information about the practical skills of being an academic – a matter we will return to later. It is easy for the faculty to assume students know how to do these things or that they will pick them up by watching. But for many students they remain a mystery.

Catherine found these sessions eye-opening. After attending the first one or two, she made a point of attending as many as possible. She had had no idea about this “hidden curriculum” of becoming an academic, except as individual topics had arisen informally through conversations about other things with her advisor or sometimes with more advanced students. Some of her classmates and co-workers dismissed these as covering obvious matters; Catherine felt differently.

Graduate students “don’t know what they don’t know” as one RTG student put it. It is important that faculty make students aware of the four universes they will be working in, and the venues through which they can become prepared.

Catherine also attended sessions sponsored by the School of Education Information Services Center (an expanded term for what was once just the library) on various software packages of use to students and faculty. A two-session presentation on N Vivo software for storing and coding qualitative data was particularly useful, but she also was helped by sessions on SPSS, by a session on voice recognition software, and by one on making more efficient and effective web searches. The Brown Bag series included a session on the mechanics of how to apply for IRB approval, and the dilemmas that researchers should be alert to in dealing with vulnerable human subjects, among whom minors clearly rank.

Thus, during the two and a half years of her second phase of graduate school, Catherine had a wide variety of experiences that are summarized in the following table:

Substantive courses: anthropology course on interracial interaction; sociology course on race and gender in work settings based on qualitative methodologies; a course on social stratification (with an emphasis on relevant theory); sociology of ethnicity in education

Methodological courses: year-long course on philosophical foundations of quantitative and qualitative research; year-long course on qualitative methods of research; advanced seminar in qualitative interviewing methods in sociology

Informal academic sessions: brown bag sessions on faculty and student research in her own department; talks and presentations in anthropology department and elsewhere

Doctoral apprenticeship group (DAG): Continuing group of students, at multiple stages, on whose committees her advisor serves presenting their research and research dilemmas to each other. This group concentrates on the challenges in producing research, though also on its presentation at professional meetings. In a field of guidelines rather than rules, it gives students a chance both to get advice about fieldwork and analysis dilemmas and to see how others encounter, define, and resolve them. It also builds a network of developing scholars with similar interests and methods who support work together outside as well as inside the group.

Professional presentations across campus: Catherine attends School of Education Information Systems sessions on software programs, campus-wide presentations on protection of human subjects, and presentations on the skills of presenting and publishing one’s work.

Presentations at professional meetings: Catherine participates in presentations of research from projects in which she was a staff member, with an increasingly active role in her second, third and fourth years.

Research assistantship and experience: Independent research for masters thesis, including responsibility for development of the question, research design, data gathering and analysis, and writing a scholarly paper. Assistant on a mixed methods project studying instruction, classroom interaction, and school practices in urban schools serving different ethnic groups. Catherine's own work is in observing classrooms, but she is part of discussions about all aspects of the project.

### Phase 3: Work on the dissertation

Catherine took her comprehensive/preliminary examinations which in her relatively small and intellectually diverse department took the form of twenty-five page essays in two to three areas that she and her oversight committee identified as appropriate. She had two weeks to write these essays.

After passing the comprehensive/preliminary examination and before writing her dissertation proposal, Catherine polished her AERA paper (which was based on her masters thesis) and sent it to a journal. She received comments on drafts of the manuscript from her major professor/advisor and two members of her DAG whose interests were close to hers.

Catherine's advisor insisted on meeting with her advisees at least once a month while they were writing their dissertation proposals. Catherine found this to be an intellectually difficult period and she felt like she was floundering, trying to find a way to shape her interests into a sufficiently narrow, researchable question. However, having gone through these steps for her MA thesis turned out to be helpful, and gave her confidence that she could do it again.

Catherine decided to do a study that would look at the differences in teaching that occurred in classes of differing demographic composition. She eventually presented this proposal idea to the DAG. She made up her mind to narrow the question in light of feasible limits and in light of the available demographic mixes at schools that would not require undue amounts of travel. She ended up selecting two middle schools that served areas that differed markedly in social class; and in these she looked at science and math classes, which are most tracked, and at English classes which are likely to be less tracked and which directly call upon more knowledge

acquired in the community. She also was able to do an informal pilot study. The DAG made several helpful comments, and Catherine moved ahead to drafting her formal proposal.

She wrote her IRB proposal before she had completed her literature review for the departmental proposal, because approval can take several months and fieldwork can not begin until approval is final. While she was waiting to hear from the IRB she completed a fifty page proposal for the department and submitted it to her committee, which included her present and past advisors and the professor from her sociology class on Social Stratification. (Flagship University allows, even encourages, faculty from other departments to serve on dissertation committees.)

Catherine spent most of a year in the field – after negotiating access and IRB approval – and all of another year in writing her dissertation. She attended the DAG when she could during her fieldwork; and after her field work, she attended more regularly. It was of great assistance. She also continued to attend the professional development Brown Bags when not in the field.

When she started feeling isolated from the university and peer support in the middle of her fieldwork, she set up a writing support group with three other students from her entering cohort. They were not all doing the same kind of work and were sometimes surprised by the different expectations in their disciplines, but they served effectively as friendly critics. When they started writing, they agreed to trade chapters by appointed dates, which helped to keep them all on schedule.

During the writing process, Catherine’s dissertation committee members were supportive. Her advisor was prepared to read small “chunks” of her work, while the others preferred to read two or three chapters at a time, after she had responded to her advisor’s initial comments about them.

In the last semester that Catherine was taking classes, and the one while she was writing her proposal, her initial advisor’s project had wrapped up. She needed teaching experience and so she turned to a

Members of the TF stress the importance of some form of teaching experience at the graduate level, even for ARs who do not intend to take a teaching position – not only does this help in forcing one to clarify ideas, it also helps improve the clarity with which one formulates material (and grading student work teaches a lot about poor argumentative and expository practices).

position as a teaching assistant for salary, tuition remission, and health insurance, not to mention for some teaching experience at the higher education level. She had the opportunity to serve as Teaching Assistant for an undergraduate course for students preparing to teach, that broadly introduces them to the social context of education. She found that she learned more than she could have imagined possible from preparing for her discussion sections and especially from helping to grade student assignments.

Catherine wanted to be able to concentrate on her field work and writing during her two final years in the field and writing, but she needed financial support (and she felt that some limited time each week away from her dissertation would be beneficial), so she took an assistantship with the program for urban high school students where she had worked in the summers. Her responsibilities included administrative duties and working with groups of students who came to campus two afternoons a week for tutoring and enrichment. She appreciated being able to keep most of three days a week free for her own work, and she found that her conversations with the students on campus and with school representatives as part of her administrative duties enriched her thinking about the dissertation work.

Students can draw from many sources for intellectual growth and support during their dissertation writing. In Catherine's case she benefitted from:

- Her DAG group
- A peer dissertators writing group
- Working with her advisor and other committee members
- Brown Bag professional development sessions
- Her administrative position working with urban students

#### Phase 4: Moving into a job

In her writing year, Catherine also was applying for jobs. The Brown Bag at the research center had sessions on searching for jobs, on resume writing, and on negotiating when given an offer, all of which Catherine found very helpful. More advanced peers, including some whom she could email who were holding their first jobs, also were helpful. Her advisor sent out notices of position openings to members of her DAG and also sent personal notes to Catherine (as to others) about positions that would particularly suit her.

The DAG had a couple of sessions where they discussed some of the special issues that qualitative researchers face in a job hunt. How do they choose chapters to send as writing samples, when each tends to depend upon the whole for its full meaning, and when the

conclusions have not yet been written? How does one present one's work persuasively to an audience that comes from different traditions, especially ones where research follows clear rules that are inappropriate in ethnography? How does one articulate the purposes and strengths of ethnographic work for such audiences?

There was also a session on writing syllabi, since some institutions requested them. One of Catherine's peers who was to be interviewed in a liberal arts college, practiced a job talk centered around a draft syllabus that she had worked up, with a discussion of the ways in which it instantiated her teaching philosophy. Catherine realized that she needed to be prepared for some serious questioning about courses that she would be prepared to teach and her approach to teaching, even at a research university.

Catherine interviewed at two universities while writing her dissertation. She practiced her job talks in front of the DAG for critique and discussion of the sort of questions just mentioned. They pooled knowledge of the work of faculty at the universities where the interview would take place. The professor also advised them all to go to the department's website prior to the interview, study the page of each faculty member and read at least one thing written by as many as possible, but certainly those with the interests closest to one's own.

<p>In the conception of advising held by the TF, the role involves more than shepherding the student through the dissertation writing stage. Getting a job, and continued guidance about professional development, lead often to a lifelong bond.</p>
---

Catherine did not get offered a job the first year. She took a position lecturing for one course per semester at Flagship, and her advisor told her to spend time turning her dissertation into a book. This strategy proved helpful, as in the following year – with doctorate in hand and a stronger record of teaching and publication – she was successful in finding an appropriate university position.

*The following table summarizes the graduate training and relevant research-related experiences of both Denise and Catherine. A blank form of the table can be found in Appendix 3; members of the TF believe this might be a useful tool for many faculty research advisors/mentors.*

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
<b>Identifying and refining important research questions</b>								
Understand the education landscape (problems and issues that require research)	Denise Catherine		Denise Catherine					
Locate one's own research interests and questions within the education landscape	Catherine		Denise Catherine		Denise Catherine			
Gain familiarity with various disciplinary traditions, perspectives and methods of research in education	Denise Catherine		Denise Catherine	Denise Catherine	Denise Catherine	Denise Catherine	Denise Catherine	
Understand definitions of and distinctions between major <i>quantitative</i> approaches to research	Denise Catherine			Catherine				
Understand definitions of and distinctions between major <i>qualitative</i> approaches to research	Denise Catherine	Denise			Catherine			
Situate and refine one's own questions in relationship to prior education research and theory	Denise Catherine		Denise Catherine	Denise Catherine	Denise Catherine	Denise Catherine		Denise Catherine
Formulate researchable questions	Denise Catherine		Denise Catherine	Denise Catherine	Denise Catherine	Denise Catherine		Denise Catherine

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
<b>Developing research design and strategy</b>								
Develop understanding of the assumptions, strengths and limitations of alternative research designs	Denise Catherine		Denise Catherine	Denise Catherine	Denise Catherine	Denise Catherine		
Create and justify a research design <ul style="list-style-type: none"> <li>• Establish fit between given research question(s) and design</li> <li>• Determine unit(s) of analysis</li> <li>• Define data needs and critically assess different data types and sources</li> <li>• Justify the nature of a sample and sampling design</li> <li>• Select and explain analytic tools and processes</li> </ul>	Denise Catherine	Denise	Denise Catherine		Denise Catherine			Denise Catherine
Develop or select and pilot appropriate instrumentation or processes (observation protocols; attitudinal measures; surveys; etc)	Denise Catherine			Denise Catherine	Denise Catherine			Denise Catherine
Evaluate and protect against threats to validity (e.g., selection bias)	Denise Catherine		Denise Catherine		Catherine			Denise Catherine
Develop understanding of	Catherine	Denise		Denise Catherine	Catherine		Catherine	

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
ethical issues and protection of human subjects								
Complete IRB requirements	Catherine			Denise Catherine	Catherine		Catherine	Denise
<b>Prepare for quantitative data analysis</b>								
Understand concepts and methods related to probability, hypothesis testing and alternatives to hypothesis testing; error rates and power analyses; null hypothesis testing & significance levels	Denise Catherine			Denise Catherine				
Become familiar with the assumptions and processes associated with major analytic methods	Denise	Denise		Denise Catherine			Denise	Denise
<b>Prepare for qualitative data analysis</b>								
Become familiar with the assumptions and processes associated with major analytic methods (e.g., taxonomic analysis, domain analysis, thematic analysis, discourse analysis, constant	Catherine	Denise			Catherine			

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
comparison ("grounded theory"), narrative analysis, artifact analysis, proxemic & kinesic analysis								
Gain in-depth knowledge of and experience with one or more of the major approaches and associated analysis techniques	Catherine	Catherine	Catherine	Catherine			Catherine	Catherine
Negotiate access to sites and/or data sources	Catherine			Denise Catherine				Denise Catherine
Obtain informed consent documents from subjects if needed	Catherine			Denise Catherine				Denise Catherine
Establish and maintain field research relationships	Catherine			Denise Catherine				Denise Catherine
Conduct initial observations to understand immediate context and to map the units of analysis to be studied most intensively.	Catherine							Denise Catherine
Do observation using structured protocols, checklists, or rating scales				Catherine				
Conduct observations, creating fieldnotes and/or audio and video records	Catherine			Catherine				Catherine
Conduct interviews with varying degrees of	Catherine			Catherine				Catherine

	Regular Courses	Mini-courses work-shops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
structure								
Administer surveys or assessments				Denise				Denise
Collect administrative record data	Catherine			Catherine				Catherine
Develop new indices or scales in existing data sets Denise Catherine								
Collect relevant documents and artifacts	Catherine			Catherine				Catherine
Keep field research journal to track data collection activities, decisions, issues	Catherine			Catherine	Catherine			Catherine
Write methodological and analytic memos	Catherine			Catherine	Catherine			Catherine
Identify, manage and document ethical issues and dilemmas	Catherine			Denise? Catherine				Denise? Catherine
Systematically seek disconfirming evidence	Denise Catherine			Denise Catherine				Denise Catherine
Organize, clean and manage data; determine level and type of missing data	Denise Catherine			Denise Catherine				Denise Catherine
<b>Conduct statistical analysis</b>								
Choose and apply the appropriate analytical methods/techniques:								
Descriptive statistics	Denise Catherine			Denise				Denise Catherine
Correlational and	Denise			Denise				Denise

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
Regression analysis	Catherine							
Determining reliability and validity	Denise	Denise		Denise				Denise
Multivariate Analysis	Denise			Denise				Denise
Advanced techniques (multi-level analysis; growth modeling; social network analysis)		Denise						Denise?
Testing of means (ANOVA, related techniques)	Denise Catherine ?							
Modify analysis plan based on data and new questions that emerge			Denise			Denise		Denise
<b>Conduct qualitative analysis</b>								
Prepare data for analysis (transcription, content logging, etc.)	Catherine			Catherine				Catherine
Confirm or select analysis method appropriate to the research question and the available data (e.g., narrative analysis, taxonomic analysis, discourse analysis)	Catherine		Catherine	Catherine	Catherine			Catherine
Conduct coding/analysis; document decision rules; prepare coding glossary	Catherine			Catherine				Catherine
Determine reliability or consistency of coding method and codes	Catherine			Catherine				Catherine

	Regular Courses	Mini-courses workshops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
Assess validity or trustworthiness of emerging assertions and claims (linkage charts; assertion matrices; member checks; peer coding sessions)	Catherine		Catherine	Catherine	Catherine			Catherine
Evaluate and protect against threats to validity or trustworthiness of analysis (e.g., selection bias; confirmation bias)	Catherine		Catherine	Catherine	Catherine			Catherine
Prepare appropriate data summaries and displays (e.g., matrices, taxonomies, interaction maps, category charts)	Catherine			Catherine	Catherine			Catherine
Write theoretical and analytic memos that trace analytic decisions and progress	Catherine			Catherine	Catherine			Catherine
Modify analysis plan based on data and new questions that emerge			Catherine		Catherine			Catherine
<b>Interpret, write about, present and publish results</b>								
Complete a research paper: using the results of analysis; decide the focus for a particular paper; frame the argument; present methods and	Catherine		Catherine		Catherine			Catherine

	Regular Courses	Mini-courses work-shops	Faculty advising	Research apprenticeship projects	Doctoral Study/Rsrch Group (faculty led)	Peer Groups (student formed)	Professional Activity (e.g. brownbags, conferences, publishing)	Students' Own Research including Dissertation
results; distinguish between findings and conclusions; make a clear chain of argument in the discussion								
Become familiar with journal submission, review, and revision process							Catherine	
Practice writing for different audiences (specialist, non-specialist; research and professional)	Denise Catherine		Denise	Denise Catherine				Denise? Catherine
Develop and practice conference presentations	Catherine			Catherine	Catherine	Denise Catherine	Denise Catherine	Catherine

## **8. The programs of Denise and Catherine assessed from the perspective of the multiple universes**

As explained earlier, the two vignettes are amalgams in which we have incorporated some of what we believe to be the fruitful practices that have evolved at many of the RTG institutions. It is the view of the TF that the two students received good training, and were well-equipped by time of their graduation to become active members of the educational research “community of practice”. We wished to document these training practices, as the institutional memories at our respective institutions are fallible; in particular, as faculty participating in RTG activities phase in and out of core seminars and courses and the like, the offerings change and successful innovations made by faculty members in one year can be modified or even forgotten in the next. We are hopeful that, with the vignettes as an aid, our colleagues will be motivated to preserve the worthwhile features that we have built into the graduate programs of Denise and Catherine.

The TF members are aware, as well, that some at least of the experiences that proved to be so important in the development of Denise and Catherine as researchers, are unlikely to have been a part of the graduate history of perhaps the majority of students. Ensuring that all ARs in our institutions have the same opportunities to grow – the same stimuli to prompt intellectual development – as Denise and Catherine have had, is a major challenge facing doctoral-level research training programs. To borrow an expression that sometimes appears in curriculum debates, “the best education for some should be the best education for all”.

But our main focus in the present Section is the question of how the training of our two heroines fares when examined in the light of the analytic tool we developed – that is, from the perspective of the four universes they will find themselves inhabiting after graduation. It is clear that initially they are positioned differently with respect to these universes, and their graduate training can be thought of as attempting to build upon their respective strengths and to remediate their deficiencies (and the ultimate degree of success of their programs can be assessed in terms of these dimensions). It is noteworthy, however, that neither of their institutions had a formal mechanism for establishing a “baseline” of these strengths and deficiencies with respect to the four universes – for recording the academic profile of their entering students; therefore the TF recommends that an “Intake Survey” instrument be used,

and an example of what we have in mind is included in Appendix 4. We believe that this instrument could be useful not only at time of intake, but also during the annual reviews of student progress.

So, with respect to Denise and Catherine, although we do not have the guidance that such a survey could provide, our preliminary judgment is that their training fares quite well, but nevertheless there are some lacunae to be filled and some lessons to be learned. However, before we become more specific in our evaluation, we note several vitally important aspects of both the vignettes that leapt out at us during the course of our deliberations:

First, some of the experiences that were so important in the development of Denise and Catherine had been planned, and indeed had arisen in situations that had been institutionalized – such things as the first or second year multi-disciplinary “Spencer seminar”, membership of small multi-disciplinary writing support groups, brown-bag sessions and “fireside chats” and the like that focused on professional development issues. *Crucially, however, a number of the “peak experiences” occurred by happenstance* – they were not systematically planned but arose as happy accidents. We have in mind such things as attending a controversial session at AERA or at other professional meetings, the contacts made in an AERA session with students from different institutions who had similar interests, the spur-of-the-moment decision to take a course across campus in the Philosophy or Anthropology Department, the interaction in a small group with a fellow student having diametrically opposed views about methodology, the chance opportunity to join a group carrying out research in an area that ended-up being incredibly fruitful. *Such important stimuli to development should not be left entirely to chance*; the TF believes that monitoring each student’s program through the lenses of the four universes, and taking account of the broad range of both formal and informal opportunities potentially available to students (rather than merely focusing upon courses being offered in the School), will likely lead to more widespread incidence of these crucial experiences.

Second, it is noticeable that throughout the vignettes, and also in the discussion above, that the advisors play an absolutely central role. For reasons described in Section 6, we have depicted the advisors of both Denise and Catherine as being quite exemplary; these advisors were

considerate, generous with their time, and ever-vigilant in pointing the ARs towards experiences (courses, projects, and the like) that would foster their growth. Even so, the efforts of the advisors seemed slightly haphazard – even when discussing core methodological or substantive courses, their advice sometimes seemed to come on the spur of the moment, in response to some remark from Denise or Catherine, or in response to something that had just come across the advisor’s desk. If they had been guided more explicitly by a set of principles, and by a vision of the universes they need to introduce their advisees to, they would have been even more effective. And of course, all who read the vignettes wonder about the fate of students whose advisors are not of the same caliber as those depicted here. Also, as we noted earlier, we are quite aware that we have depicted Denise and Catherine as being extremely compliant; most graduate students – for good or ill – have wills of their own, and although they often *consider* the advice of their mentors they do not always *follow* it. We take up the issue of advising in more depth, and offer some suggestions, in Section 9 below.

Third, we note a weakness in the preparation of Denise and Catherine at a particularly crucial stage of their graduate histories – one that, we believe, is also to be found in the vast majority of programs of doctoral students across our institutions: We refer to what happens, or rather to what does not happen, at the stage in which these ARs are selecting a dissertation research topic. To make our point, we need to revisit the conceptual discussions that occurred prior to the presentation of the vignettes – it will be recalled that in Section 5 we made the case that quality research is a function not only of methodological rigor, but also of factors such as social and scientific relevance, and sensitivity to issues of social justice and equity; and when outlining the nature of “Universe 2: A discordant social universe” (this universe essentially is the context in which researchers will be doing their work) we made the point that a prominent feature here is the struggle pursued by many groups for social justice. The weakness, then, is this: It is not clear that when engaged in the difficult task of delineating a research topic our students were made explicitly aware of these desiderata. Perhaps there was some discussion of what constitutes quality in research in the “dissertation proposal” writing class, although such issues are not commonly found in the syllabi of such classes; and perhaps the students were asked to address these issues during their proposal defenses. But again the TF believes that these matters should not be left to happenstance, or paid only “lip service”.

The final point is one that pertains to an aspect of doctoral training that members of the TF believe to be of great importance, and in which our ARs might not have been adequately prepared. It was not clear from the vignettes what opportunities there had been for Denise and Catherine to reflect about – and to discuss with mentors and peers – the ethical issues that arise in the course of empirical educational research. It now is standard practice, of course, for doctoral students doing empirical work to become familiar with the process of IRB approval, and our ARs certainly did so. But we stress that the real ethical dilemmas that often face researchers, particularly when working in field settings, go far beyond the issues that are likely to have been discussed when the ARs were being prepared to submit the paperwork for IRB approval. Discussion of these issues probably belongs in Universe 4, but the issues themselves arise in the planning/designing, and conduct, of the actual research project – we are thinking of course, of the principle that no harm should be done to individuals who volunteer to participate as subjects in a study; the right to “informed consent”; the right to be treated with courtesy and consideration; the importance of lack of deception; the need to consider not only those individuals who profit from a treatment or program that is being studied, but also the degree to which those who most need the assistance of the program have gained or suffered a loss; and balance-of-power issues between rights of students, their parents, and teachers and school administrators.

While every type of empirical study in education is likely to face one or more of these issues, qualitative research faces an additional subset of issues – for example, the need to write qualitative reports in such a manner that specific individuals cannot be identified, and the need to keep confidential things that might happen to be learned during fieldwork where the researcher is a “participant observer” in the life-world of the people being studied. There is the additional complexity here that there may be legal requirements, such as the need to report cases of child abuse that are discovered. Members of the TF with experience in conducting qualitative research were fairly confident that Catherine would have been sensitized to issues such as these; one of the Task Force with a great deal of experience teaching qualitative research methods wrote as follows: “IRBs can pick up only the most egregious ethical issues in research design. But most ethical issues pop up unannounced and often urgently in the midst of

fieldwork or analysis. In qualitative work they get discussed in many venues *because the students bring them up*. The students feel them keenly, as ethical dilemmas rear themselves abruptly and often without warning as obstacles in the midst of daily relationships or of analytical reflection." Putting aside the details of Catherine's training, the TF agreed that this general domain of issues is so important that faculty members should carefully monitor the training of all ARs to ensure that they obtain adequate exposure.

Preliminaries over, the discussion now turns to review the programs of the two ARs from the following perspective: How well did their training prepare them to work within, and deal with the problems and crises that arise within, the four universes? Although, for convenience, our discussions are set out sequentially in terms of the universes, we do not wish to imply that the actual programs were or ought to be arranged in this order; rather, work in each of the universes is, and must be, spread over several years. We give a succinct summary of the relevant training of Denise and Catherine in each universe, following each of which is a brief commentary.

### **8.1 Analysis of the training of Denise and Catherine**

#### Universe 1: Research frames and their epistemologies, and basic social-science theories

When all of Denise's work in the domain of Universe 1 over the years is collated, the result is impressive. She had the advantage (from this Universe's perspective) of entering graduate school with some relevant training in social science theories and research, and with an already-developed interest in quantitative methods. She both broadened and deepened her knowledge of theory and of quantitative methods – these were the chosen bases of her research orientation – through her coursework, workshops, interactions in her writing group, and of course through her research apprenticeships and the growing responsibility she accepted therein; and – importantly in our view – she conducted a pilot study so that her dissertation was not her first major, independent piece of research. The development of her knowledge of qualitatively-oriented research was somewhat less systematic; her interest was aroused by interactions with other students, some course-assigned reading, and sessions she happened to attend at AERA. Eventually she acquired "consumer-level" knowledge of this domain of research from interactions with students in her writing group, from workshops, and from an introductory

course. In parallel, she also gained some understanding of the different epistemological foundations of the major research frameworks – enough, probably, to serve as a base from which to do more reading and thinking.

One point we wish to make here, that is of significance above and beyond Denise's training, pertains to the importance of the admission process into graduate programs. One way to circumvent certain problems that we discussed briefly earlier, is to select students for admission who have a background in those areas of professional activity, or academic disciplines, which are regarded as important and which the particular programs at the relevant institution wish to assume as the base upon which they will build; or alternatively these types of backgrounds might be desired because the program is itself unable to provide training in them. (The intake survey mentioned earlier might be of some use here.) We pointed out that one such problem in many Schools of Education is ensuring that ARs achieve the same deep level of methodological training as students in roughly parallel programs elsewhere in the university. Seen from this perspective, Denise was an appealing candidate to faculty in the program to which she applied, because she had an undergraduate training that served as a solid foundation upon which to build her subsequent theoretical and methodological training. However, clearly there are tradeoffs here – her strengths in one dimension or universe were offset by weakness in others; for example, as we comment later, she was relatively weak in her detailed knowledge of educational settings. The TF believes that evaluation of each candidate for admission using the four universes as an analytic tool brings into focus the individual's relative strengths and weaknesses, and sharpens the tradeoffs that need to be considered explicitly.

Catherine, on the other hand, came to her doctoral program with an undergraduate background in chemistry, which had been followed by teacher-training and a number of years of successful teaching; but essentially she had no familiarity with the social-science disciplines, their epistemologies, and their associated research methodologies. However, her undergraduate background no doubt gave her a basic understanding of the role of experimentation in science, and also provided her with a mathematical launching place for her later work in statistics. She took a basic statistics sequence in her first year, and worked on a research project that gave her a variety of important experiences in different phases of the research and which prompted her

reflections about the limits of various research methodologies. Coursework in sociology and anthropology served as an entry into social science theories, and further sensitized her to the differences between qualitatively and quantitatively-oriented research; and the tension between the two different approaches led her to think deeply about underlying epistemological and methodological assumptions. Later she took a minor outside Education that focused on the more qualitative aspects of sociology and anthropology (theory, methods, and epistemology seemed to have been covered here), and in the midst of these efforts she worked on an MA project. In addition, in her second year she took both a year-long sequence that gave in-depth experience in qualitative research methods and in applying them in the small-scale study that she conducted, and a course focusing on the logic of inquiry. Her development in Universe 1, then, was quite extraordinary – she moved from being quite ignorant to being very well-informed about the whole universe, with strong interest in the deep underlying issues, and she had become familiar with an array of methodological skills from the universe’s main constellations.

We note, in reflecting upon Catherine’s training here, that her initial advisor was exemplary in alerting her to the need to become literate in statistics which she would need either as a researcher or as an informed consumer, but also in encouraging her both to obtain initially a broad overview of the field of educational studies and also to become involved in research from the outset. We also see reflected in her program the fact that acquiring qualitative research skills follows a somewhat less clearly defined path than acquisition of quantitative skills; consequently training can be more drawn out. But what stands out, of course, is how impressive a student Catherine really is – in the degree to which she reflected very deeply about issues raised in her courses, and how she thought about these issues in terms of their consequences for her own research work, and how she sought appropriate courses and the like when she needed to probe even further. Many graduate students have great seriousness of purpose, of course, but not usually to the same extent as Catherine; and the TF is at something of a loss with respect to giving advice on how this degree of reflection and analysis can be fostered in all students – although we do offer some tentative suggestions in Section 9.

Reflecting on the experiences of both Denise and Catherine, it is striking (as we noted in the opening paragraphs of Section 8) how their decisions about what to explore in Universe 1 were made contingently; to put it figuratively and perhaps a little strongly, both of them, but more particularly Denise, were like sailboats that moved in one direction or another at the mercy of the winds and the tides (the affordances on the lake) – one set of forces moves the vessels to one position, where other forces soon operated and took them further on their journey but not necessarily in the same direction. There was continuity of movement, but not necessarily a pre-planned course and clear destination. Thus, in Denise’s case, the chief decisions about her coursework in her substantive areas of interest, and also with respect to the methodological toolbox she was acquiring, were made in discussions with her advisor, but without an explicit “template” or framework or rationale giving direction; this also was lacking in Catherine’s case, although there are some hints that her advisor was guided by a well-thought-out set of principles that were only partially conveyed to her.

It is, of course, a good thing for an advisor to be able to work closely with a student to individualize the details of her training; but there are some downsides. First, without an underlying institutionally-endorsed rationale about foundational courses, developmental sequences and the like – for the first phases of a program at least – there is the consequence that different advisors working with comparable students can recommend different courses as being of basic importance, or as being prerequisites for further work. This can be confusing to students, who often compare notes about their courses and requirements. We are not suggesting that there should be a single, invariant pathway through the coursework, but it seems advisable for the faculty in a program to have reached agreement about what skills and bodies of knowledge are basic, and which are necessary as foundations for further development.

Second, institutionally designed guidelines or pathways allow students to see – from the beginning – the overall direction in which they will be moving; such guidelines help provide coherence and lessen the chance of students becoming “lost” in a welter of coursework. These guidelines or pathways or requirements (we have no stake in the form these take – it is the

function that is of present concern) also often provide a window into the underlying educational philosophy of the institution.

To cut a long story short, most members of the TF would have been happier with the advising that both the ARs received – careful and helpful though it was – if the faculty members of their respective programs had reached broad consensus about course sequences and the like, so that early on the advisor was able to point to these in the “courses and degrees bulletin”, and also to say something along the following lines:

“In addition to fostering development of your substantive interests, our program is designed to get you to the methodological cutting edge, which at the moment involves use of techniques X, Y and Z. You might not need to be expert in use of all of these (it will depend on how your research interests develop), but certainly you will need to master one or two.... But to reach this target, you need first to build a foundation, and we have found that courses A followed by B are very effective at accomplishing this; these can usually be followed by C and D and E taken in any order, which then open the door to X and the others! Since you already have undergraduate work that was equivalent to A, we can waive taking that course and you can go straight to B.”

A number of the RTG institutions, after considerable work and discussion, have been able to develop a sequence of courses that their faculty expects will provide new doctoral students with a solid and broad foundation within Universe 1. We offer the following example not necessarily for emulation, but because it presents an interesting path that can be taken to deal with some of the problems discussed earlier. Clearly, not all institutions will want to be as directive as this, and undoubtedly there are alternative viable ways to organize the methodological content. We also realize that training institutions differ with respect to resources that are available; in particular smaller Schools of Education might not be able to offer the same range of courses as larger institutions.

The material in the example here has been extracted from the Stanford University School of Education website at <http://ed.stanford.edu/suse/programs-degrees/program-phd.html>; and it is to be noted that over the past ten years or so – the period covered by the Spencer RTG grant – there has been considerable variation in the organization of the School’s methodological training and in the theory-oriented courses that have been offered; the example, then, is a snapshot of faculty thinking at one moment in time. Other examples may be found at <http://www.gse.harvard.edu/academics/doctorate/curriculum/index.html>, at <http://www.sesp.northwestern.edu/ls> and at <http://www.educ.msu.edu/te/phd/> (and click on “current students”).

**The Core**

The School of Education requires all first-year doctoral students to enroll in a set of foundation courses in the first two years of their program. In the first year, all students enroll in EDUC 250A: Inquiry and Measurement in Education (Autumn Quarter) followed by EDUC 311X: First-Year Doctoral Seminar-Introduction to Research, in the Winter and Spring Quarters. Every doctoral student is expected to complete the methodology core by the end of the second year of studies. EDUC 250A is the first of a series of courses in this core which is taken in the first year, followed by EDUC 250B: Statistical Analysis in Educational Research, and EDUC 250C: Qualitative Analysis in Education.

In addition, doctoral students develop competence in each of the following four dimensions of education during the course of their studies:

Curriculum, instruction, administration, and special services.

Behavioral and social sciences. Courses in organizational theory, psychological foundations of learning, and social processes familiarize students with basic concepts in the behavioral and social sciences applicable to understanding and resolving educational problems.

Normative studies. Students should understand the ideological, historical, and philosophical contexts from which educational values and aims arise in order to develop effective educational policies for the future.

Inquiry skills. Historiography, ethnography, statistical analysis, and computer programming are a few of the tools students will use to complete their dissertation research. The methodology core courses may be used to satisfy this requirement. Specific course requirements for each of these four dimensions differ within each program area.

Often the requirements may be satisfied by courses taken in other academic departments at Stanford.

**The Minor**

Doctoral candidates minor in a discipline relevant to the field of education such as anthropology, economics, philosophy, psychology, or sociology. A prior Master's degree in a relevant field can often be used to satisfy this requirement. Such approval is granted if the

student's Area Committee, after consulting with faculty in the relevant department, is convinced that the Master's degree meets the objective of providing the student with scholarly and methodological skills equivalent to those required of a PhD in that field. Typically, professional Master's degrees, such as those in business and nursing, do not satisfy this requirement.

### **COURSE DESCRIPTIONS**

**EDUC 250A.** Inquiry and Measurement in Education. Part of doctoral research core. The logic of scientific inquiry in education, including identification of research questions, selection of qualitative or quantitative research methods, design of research studies, measurement, and collection, analysis and interpretation of evidence.

**EDUC 250B.** Statistical Analysis in Education: Regression. Primarily for doctoral students; part of doctoral research core; prerequisite for advanced statistical methods courses in School of Education. Basic regression, a widely used data-analytic procedure, including multiple and curvilinear regression, regression diagnostics, analysis of residuals and model selection, logistic regression. Proficiency with statistical computer packages.

**EDUC 250C.** Qualitative Analysis in Education. Primarily for doctoral students; part of doctoral research core. Methods for collecting and interpreting qualitative data including case study, ethnography, discourse analysis, observation, and interview.

Later, in Section 10, we shall provide the specific recommendations of the Task Force about the topics an AR should have mastered, either as a prospective producer of research or as an informed consumer; we focus on topics rather than courses, for different faculty members and different institutions may well bundle these topics into actual courses in quite different ways. The issue we have been pointing to here is that the faculty in a program should have reached agreement about (i) what topics a student needs to have taken to be considered well-trained; (ii) in what sequence these topics normally should be met, if indeed some are pre-requisites for others; and (iii) what the end-point is, that is, how near to the cutting-edge are students expected to move? To add to the burden of these faculty members, it is important to remember that they also need to decide what courses or topics are needed for adequate “consumer-level”

competence – for example, what precisely should a qualitatively-oriented student such as Catherine study in the quantitative area, and what qualitative skills and methodological precepts are appropriate for a quantitative person like Denise? It is not clear from the vignettes that their advisors had well-thought-out positions on this latter matter.

### Universe 2: Social and educational problems

Denise entered her graduate program with a specific educational issue that she wanted to work upon, but her knowledge of – and interest in – this issue seemed to be purely academic. She broadened her academic understanding of education by way of a history of education course; and later she worked with a team that was communicating with a school, and she had structured practice in how to communicate effectively in this kind of context. She still faced challenges in dealing with schools later in her program.

The overall impression is that Denise still has a way to go with respect to developing her understanding of Universe 2. And of course she is not alone here; earlier we pointed to the tradeoff involved in admitting a technically-qualified applicant to a program that has a methodological/analytic focus – a strong technical background in an individual newly-emerged from the baccalaureate is not commonly accompanied by a background of practical educational experience of one type or another, and of course vice-versa. Most if not all RTG institutions have doctoral students who were admitted under each of the possibilities offered by this tradeoff.

Denise was aware that her lack of educational experience made it difficult for her to communicate effectively, but we wish to stress that there is a deeper deficiency, one that Denise seems not to have clearly recognized. This was touched upon earlier in Section 4, in our discussion of contextual knowledge; as we put it there: “being able to ask the ‘right question’, or being able to select the telling problem, are contextually-sensitive activities”. A number of RTG institutions attempt to resolve this situation – or at least to partially alleviate it – by engaging students as much as possible in the field-data-collection phase of a research project; this requires careful monitoring of the research-related experiences all students are having, and sadly there are not sufficient field-based opportunities at any one time for all of the students to profit in this

manner. This is not an entirely satisfactory resolution to the problem, however, as field data collection does not necessarily mean the AR has experiences that increase understanding of the often complex educational setting; another possibility is one that was presented in Catherine's vignette – she worked in summer in a school district, doing administrative work and tutoring.

In contrast to Denise, Catherine came to her program with a teaching background which had given her an appreciation of social and educational problems, especially the difficulties faced by language-minority children in school, and also the stereotypes faced both by them and by female students in math and science classes. Initially her practitioner's perspective was context-bound, but starting with her coursework in history and political economy of education, as well as in anthropology and sociology, her understanding was expanded. She was, never-the-less, more solidly grounded within Universe 2 than most entering students. But again we note a caveat: former practitioners often struggle for extended periods to acquire the ability to view familiar educational settings through the eyes of a researcher.

### Universe 3: Substantive educational knowledge

Denise came to her graduate studies with some undergraduate background in political science, but without knowledge about how this material played out in education. In graduate school she built upon this foundation, taking a variety of courses; but she seemed to broaden more as a result of discussions in her writing group and her interdisciplinary project. Her advisor discussed with her the importance of reading in overlapping areas, and indeed she did read on her own to get a deeper understanding of some of the issues and related research findings. Thus, over the span of her program, she moved from being a complete novice in this Universe to being moderately well-acquainted with it.

Again, what strikes us is the relatively haphazard or contingent nature of her immersion into this domain; her advisor gave her advice at very general level, but did not suggest a core literature for her to study – and if her writing group had possessed a somewhat different membership she might not have developed the way she did. Some, but not all, RTG institutions deal with this situation by having comprehensive examinations at about the end of the second year of the program, and these “comps” usually require students to read in at least the areas

(known beforehand) that will be covered. Other institutions have specified a set of distribution requirements (DRs) that can be satisfied by taking two courses in each of the areas identified in the DRs, as in the example from Stanford that was included above.

On the other hand Catherine's teaching background, including the coursework she took while earning teacher certification, had given her some knowledge of the research-based literature, but only to a limited extent. However, the courses she took in her first doctoral year were admirably suited to provide her with a solid overview of educational scholarship (history of education, sociology of education, political economy of education, and anthropology of education), and she built on these with further coursework and with a wide range of experiences as an assistant on research projects. Her comprehensive exams also pushed her deeper into the literature. She was not at a loss – compared to the state of many students – when it came to generating interesting and important research questions.

#### Universe 4: Professional

As mentioned earlier, Denise entered her doctoral program with some understanding of what it is like to do research, and of course this deepened as a result of her assistantships and independent projects. She attended conferences, and found that these can be stimulating – in particular some conference sessions alerted her to the issues that currently were provoking controversy among well-established researchers; and she learned something about the mechanics of making effective presentations. Also, she developed a sense of the parameters of a good CV by examining those of more advanced students.

Several things strike the members of the TF about this dimension of Denise's training. First, most students heading for a research-oriented career will obtain a position in a university, where they will be expected to teach; a major weakness in Denise's work in Universe 4 is that she did not have the opportunity to acquire even elementary teaching skills. She did not get serious experience as a teaching assistant, where she would have made a few class presentations (and received formative feedback), graded papers, helped to prepare the syllabus, and given one-on-one tutoring and remedial assistance during office hours. And we stress the point that being a TA not only develops these skills, but also can promote deeper

understanding – by the TA – of the material she is responsible for teaching and grading and tutoring. (We have often heard senior, experienced colleagues remark that they really learned a topic by teaching it.) Second, again it is striking to us that Denise’s initiation into this universe was largely unplanned and thus undirected; it seems not to have occurred to her advisor to work through, for example, the AERA conference program and draw her attention to sessions that held promise of being extremely stimulating. And Denise’s experience here is quite typical; by-and-large faculty members ignore the need for ARs to become knowledgeable about this universe – if they think about it all, they assume the AR will learn by osmosis.

Catherine’s exposure to this universe came by way of her avid attendance at brown bag programs, the discussions in her DAG, and from attendance and co-presentations at professional meetings such as AERA, American Sociological Association, and the like. She also submitted work for publication, and gained insight into this process. Thus, during her program she built a serviceable body of professional understandings. And although she was an experienced teacher, she obtained experience with tertiary students by working as a TA.

Our comment is that her university, in common with most if not all, relies on informal means for initiating research students into professional life; but there is a difference between intermittent or spur-of-the-moment informal avenues and more-or-less institutionalized ones. Catherine’s School of Education had a number of institutionally supported or sponsored venues or structures devoted to professional development; these seemed to be a regular part of the life of the institution, although they were “informal” in the sense that attendance was not mandatory. In addition to the various experiences Catherine had, we note that most RTG institutions encourage students to attend “job talks” and presentations by candidates for faculty positions; there is much to be learned by watching someone else – unfortunately, often the most educative sessions are those that do not go very well for the candidate! Some ARs also obtain invaluable experience through appointments as “student representative” on search committees or on program standing committees and the like.

The TF notes that at many if not all RTG institutions, pre-conference practice sessions are held, where students who have a place on the program rehearse their presentations and other

students (and sometimes faculty as well) act as critics and coaches. Often these practice sessions are organized by students themselves, but they are so valuable that consideration should be given to institutionalizing them. Some RTG institutions have gone further, and as we indicated in some “boxed” comments in the vignettes, they have brown-bag sessions devoted to professional matters that students can attend on a voluntary basis. One institution called these “Fireside chats – without the fire!”; for a period of about six years from six to eight sessions were held each year, covering such provocatively titled topics as “how a journal article gets published”, “what is tenure, and how do you earn it?”, “how to write a funding proposal that is convincing – and is likely to be read”, “on how to make – and how not to make – a conference presentation”, “applying for a professional position – the CV, the covering letter, and the campus visit”, “how a university is organized: structure, function, and more”. The fact that these sessions are no longer held at this School illustrates how difficult it is to institutionalize informal training activities.

## **9. Knowledge, skills, dispositions, and practices necessary for advising**

### **9.1 Overview: The vital roles of the advisor**

The TF recommends that doctoral research-training programs adopt the model of proactive and integrative advising discussed in some detail in the final part of this Section (9.3 below). In our collective view, this would be the single most important step that could be taken to improve the capability of ARs to produce research that is of the high quality that we have highlighted throughout our report.

In the opening of this Section, however, we wish to make another point. In the vignettes of Denise and Catherine we depicted the advisors as mentors and guides, pointing to opportunities and helping the ARs develop the appropriate bodies of knowledge, skills, and experience, as they navigated through their programs. But in addition we see them playing another vital role, one that did not emerge so clearly in the vignettes – namely, a role that is analogous to that played by personal trainers who help athletes make the most of their abilities by giving strong formative feedback on their performance and who push them hard – sometimes extremely hard – to hone their skills to a degree these athletes initially might not

have thought possible. Many of us on the TF recall the honest, detailed and sometimes painful feedback that we received on our work from our own advisors, and how this eventually paid off in improving the quality of both the research we carried out and the papers and reports that we were able to produce. Some researchers, of course, have developed well without such input, but our belief is that most individuals will not reach their full potential absent this type of help. Eventually all of those who devote their lives to research build up a cadre of colleagues from whom they regularly seek honest comments about work in progress, but ARs often are not yet in a position to form such a group – and it falls to advisors and other mentors on the training faculty to play this absolutely essential role.

We do recognize, however, that there is one venue within many graduate training programs where the ARs are able to get honest and probing feedback, though it is not as informed, perhaps, as that which comes from faculty advisors, and where they can also hone their own skills at giving such feedback to others – namely, the interdisciplinary collaborative writing groups that have sprung into existence at many of our institutions; and we recommend that these become institutionalized – programs should foster these, monitor them, and even if possible support their activities with small amounts of funding (a small amount of “pizza money” goes a long way).

Finally in this preamble we also recommend that Deans and program chairs be proactive in monitoring the quality of advising/mentoring; it is noteworthy that although now it is *pro forma* for regular classroom teaching to be evaluated, advising is rarely if ever subjected to the same discipline. The TF recognizes the difficulties here, for students are oftentimes reluctant to provide negative information about faculty members with whom they are working closely. But training institutions must find a way around this problem. In addition to lack of evaluation of advisors, there also is in general a lack of training, and programs at a very minimum should have annual meetings where faculty can discuss relevant issues, exemplary practices, institutional expectations about advising, and the like. Such meetings and discussions are likely to be particularly valuable for new faculty members.

In what follows, we draw from the vignettes presented in Section 7, together with our own experiences at Spencer Research Training Grant (RTG) institutions, to present examples of advising practices that might support students' development within the four universes identified in Section 2. Because we have argued for a future-oriented research preparation that develops researchers who can navigate these universes, we conclude this section, true to our vision, with a description of how advisors might not only support, but also encourage and model, movement across universes. More important, we provide examples of the kind of advising structures and practices that can foster the integration of the skills, knowledge, dispositions, and practices associated with the various universes. In each case, we want to acknowledge that advising is not a unilateral or always peaceful and fluid relationship. We have already noted that the vignettes depict dedicated faculty advisors who work diligently to provide their students with opportunities, and Denise and Catherine are stellar students who blissfully accept each opportunity and point of feedback provided. As faculty advisors in the real world, however, we recognize that some advising relationships are fraught with tension and struggle. It is, in fact, those relationships that we think can benefit most from the proactive moves we suggest, and we take the time to highlight both the strong advising moves in the vignettes (which it will be recalled are based on RTG institution practices) and also in other examples from our own institutions.

## **9.2 Examples of advising practices related to the four universes**

### **Universe 1:** Examples from the vignettes and beyond

In the vignettes:

- Despite her interest in and skill with statistical analyses, Denise's advisor nudged her to take some course work in qualitative methods; similarly, Catherine's advisor encouraged her to diversify her disciplinary background, thus opening her to new perspectives on teaching and learning.
- Both Catherine and Denise had different roles on various research teams throughout their programs of study, and their advisors explicitly drew from their experiences and helped them connect their learning in different paradigms and across paradigms.

Some other examples from the RTG institutions:

- Most RTGs are fortunate to have multiple research projects running simultaneously, and just as in the cases of Catherine and Denise, most RTG faculty members run research teams. One challenge, however, is in providing those students who are offered teaching assistantships or fellowships with the opportunity to participate in research apprenticeships. Some RTG Schools are now requiring both research and teaching apprenticeships of all students, over and above their paid assistantships. The apprenticeships provide the students with opportunities to observe and learn from research and teaching activities. A challenge lies, however, in monitoring these apprenticeships to ensure that all students have equally useful opportunities and that the opportunities are duly recorded in students' files.
- A number of institutions have established regular interdisciplinary seminars during which students present developing ideas, projects, analyses, and research reports. In some cases, these are initiated at the level of program or department (often prompted by the RTG work), and in other cases, these meetings are sponsored by individual faculty members or teams of faculty members within programs. Institutionalizing such activity has both positive and negative consequences. On the positive side, institutionalizing ensures that faculty members do not carry an unequal burden in offering such activities, and that all students have the same opportunity to participate. At the same time, institutionalizing interdisciplinary seminar opportunities can diminish the innovation and creativity that result when people come together more spontaneously to meet a joint need or purpose.
- Several RTG institutions (Stanford and University of Michigan, among others) require a number of course credits be taken outside of the School of Education (i.e., in a "base" discipline, or distributed over a combination of fields outside education); at some institutions, students obtain an M.A. degree outside of Education because of this requirement, thereby gaining further knowledge of the theories and methodologies related to Universe 1. A somewhat different model is provided by Northwestern University, which obtained two successive training grants to support interdisciplinary cohorts of students who undertook extensive training not only in Education, but also in Sociology, Psychology, Economics, and Communications.

## Universe 2: Examples from the vignettes and beyond

In the vignettes:

- Denise’s advisor encouraged her to go out and get some experience in the field, but did not push her to get the kind of experience that would prepare her for designing and conducting research in the field. Denise could have further developed her understanding of the applicability of large-scale research studies by examining actual practice on the ground.
- As a student in a year-long qualitative research course, Catherine was able to negotiate the demands of institutional review board approvals and conduct research in an area school. This experience allowed her a direct opportunity to engage with the social universe about which her research was being conducted and to the individuals and institutions to whom she would have an ethical responsibility.

Some other examples from the RTG institutions:

- Students at many RTG institutions serve on research and professional development teams based in school settings, which provides those students who bring little prior knowledge of classrooms and schools with firsthand experience, while also providing those students who have prior school-teaching experience the opportunity to see the classroom or school through a research or theoretical lens. These opportunities, however, tend to be unevenly distributed in Education programs or departments, with the greatest number falling in areas that explicitly engage in classroom research. The Task Force suggests, however, that any student engaged in education research—even analyses of large-scale data bases, study of historical archives, or policy analysis—could benefit from direct experience with classroom or school-level research because of the unique vantage point such experience provides. This is not to deny, of course, that ARs interested in educational policy issues have much to gain from internships in a branch of government or in policy “think-tanks”.
- Some RTGs offer doctoral professional seminars or core courses that bring students together who are heterogeneous by research interest, background, and experience. In at least one of these core course experiences in an RTG institution, students are required to conduct a first-year research project in a nearby school district and to report their findings back to the high school principal and other district administrators. This work

has obvious implications for students' learning about how to be responsible to and to communicate with the social universe to which their research is responsible. Some variation of this activity could be encouraged as a core requirement in all education research institutions.

### **Universe 3:** Examples from the vignettes and beyond

In the vignettes:

- To augment her knowledge of social science disciplines related to education, Catherine's advisor encouraged her to take substantive course work in history, sociology, and anthropology of education.
- Similarly, Denise's advisor suggested that she immerse herself in a particular literature (i.e., on the overlapping factors that influence student learning and achievement).
- In both cases, the students also built knowledge in practice as they participated on research assistantships, in lab activities, and in teaching experiences.

Some other examples from the RTG institutions:

- Some of the RTG institutions have vigorous brown bag speaker or discussion series for faculty and/or students that focus on matters pertaining to Universe 3; one example is Harvard, which currently offers a discussion series on, "the future of teaching and learning" as well one on "civic and moral education". Aspiring researchers may need guidance in knowing in which of the many possible activities they should engage.

### **Universe 4:** Examples from the vignettes and beyond

In the vignettes:

- Both Denise's and Catherine's advisors encouraged them to present at conferences, join research groups, co-author papers for publication.
- Advisors reviewed position descriptions, job letters, and curriculum vitae prior to a student's launching a job search.
- Catherine's advisor set up a Doctoral Apprentice Group, in which all students of that advisor came together to report on progress. Advanced students were expected to participate and served as additional mentors to newer students.

Some other examples from the RTG institutions:

- Many RTG institutions run student research conferences, most often scheduled so that they can serve as a “warm-up” for AERA or other conference presentations in Spring (i.e., the conference is usually in mid-March). In at least two cases, Northwestern and Michigan, this conference is open to all graduate students across the university.
- Many of the institutions also run brown bag series that focus on professional development matters; some examples were given earlier in the boxes accompanying the vignettes.
- At a number of RTG institutions, faculty members form working groups that bring together graduate students, post-doctoral fellows, and other faculty on a regular basis. These groups allow students (and faculty members) to try out data analyses, practice conference presentations and job talks, work on papers for publication, and learn about the review process (both ends, as reviewer and reviewed) – including how to revise papers and respond to reviewer concerns. At some RTG institutions, faculty members receive instructional credit toward their teaching loads as a result of such activity.

### 9.3 A model of proactive and integrative advising

Our discussion here embodies the assumption that to foster the outcomes discussed throughout our report, advisors need to engage in proactive and responsive advising. We assert, in fact, that one of the greatest barriers to producing meaningful, high-quality research is the lack of guidance in existing doctoral programs to help students synthesize the knowledge, skills, and dispositions that are central to the different universes for which they must be prepared. Many advisors engage in these activities, but as we have stressed they often do so idiosyncratically, as needs arise with students. Few of us are proactive or systematic in giving relevant advice. Even a smaller number of us are *synthetic* in our advising efforts, not because we are unaware of the need to synthesize, but because the syntheses may often seem so obvious to us as experienced scholars as to be virtually invisible. It is this explicit attention to synthesis that Task Force members believe needs attention in our work as advisors. Thus, we suggest that doctoral advising needs to build on the discrete activities given as examples previously, and move toward a model of advising that supports the integration of these universes for students.

To be proactive and responsive, an advisor has to be able to assess an incoming student’s knowledge, skills, and dispositions. Proactive/responsive advising engages students in

reflection on and development of questions that matter within the universe of social and educational problems, to make sense within and be articulated to broader research traditions, perspectives, and methods. Proactive and integrative advisors, we assert, do some or all of the following (and more):

- Provide opportunities for students to engage in cross-disciplinary and cross-methodological conversations;
- Provide opportunities for students to engage in systematic reflection on different epistemological traditions, research questions, problems of education practice;
- Organize a set of experiences (including courses) to promote a deep level of competence within a specific research tradition;
- Model intellectual openness and curiosity;
- Cultivate the disposition of questioning;
- Model problem solving, particularly when trying to acquire new knowledge or skills;
- Consult with colleagues about students in the interest of supporting students' development;
- Engage students in the process of writing in a way that models how to seek multiple perspectives on and critiques of one's findings and arguments.

For example, at a number of different lab meetings a year ago one Task Force member developed a coding system with his students for one of his research projects. He had no previous experience with developing coding systems, but it was what the research question called for and the question was important enough that he was willing to work on it in spite of the fact that it was beyond his methodological expertise. He asked the students for ideas about systems they had seen others use, advantages or disadvantages of those systems, and what they had read on the topic. Together, they built an approach that worked for them, which was an important developmental experience for the students in the sense that they not only learned a strategy for doing such work themselves, but they also had the opportunity to see how a faculty member worked through a new research task. In other words, instead of conveying the sense that as a faculty member one must have all of the processes and practices of research honed at the outset of one's work, involving students in this learning process demonstrated that research (and teaching) is open-ended and faculty members themselves are always learning.

What could have made the learning opportunity even more effective for the students, however, would have been the explicit articulation of the subtext of what they were learning in this process. Perhaps most notable was the fact that these students were learning how to use multiple resources to solve novel research problems, rather than dismissing the task as impossible or relying on someone else to answer the question. Another productive learning opportunity might have been a discussion about whether the choice of methods should essentially determine or constrain the choice of research question, or *vice-versa*. The activity was proactive in the sense that the researcher contemplated trying to figure out a good coding system on his own before meeting with the students, rather than just trying to just get them to do the coding – which probably would have moved the project along more quickly – but decided that the “figuring it all out” part of the experience might be helpful to novice researchers.

Another example of proactive advising is setting up initial conversations with incoming doctoral students to serve as informal “intake” interviews; it will be recalled that earlier we also recommended the use of an “intake survey”. These intake interviews can be used to gauge student goals, determine type of institution they might like to work at upon graduating (teaching, research, both), and to map out a developmental checklist of activities (conferences, publishing, being a teaching assistant) in a flexible sequence that will help position students to achieve those goals. One simple method might be to open all meetings with some variant of the question, “How have your goals changed since we last met for a discussion?” to give students openings to refine and shift their goals. A follow-up question, building on the “possible selves” literature, could be: “And what are you doing/have you done to reach those goals?” Asking students to prepare written responses to these questions, and to be prepared to discuss them in regular meetings (beginning of fall term, end of winter term, for example), might be one way of ensuring both systematicity and rigor in the advising process. This kind of advising is proactive because it does not wait for a student crisis, or even a question from a student, to encourage reflection and planning. At the same time that it is proactive, it is not a restrictive process because the developmental advice is shaped by the student’s own goals.

Finally, proactive work on advising should be supported at the program, department, or school levels. Indeed, the Task Force suggests that to some degree advising of each AR should be seen as a communal responsibility and not as the sole responsibility of individual faculty members -- in part because such an arrangement is bound to produce disparities in quality of advising, and in part because advising loads can vary dramatically among faculty members. Furthermore, some faculty members may have more challenging advising situations than others. Supporting advisors at the School, program, and departmental levels, then, is necessary to building capacity in faculty members to be proactive advisors. This support can take a number of forms, but a vital one was alluded-to earlier: it will be recalled that in the opening part of this Section (9.1) we stressed the importance of regular faculty-wide discussions about advising and the setting of norms.

#### **10. Summary, and elucidations, of the TF recommendations**

In this section of our report, the TF fulfils its obligation to make specific, concrete recommendations about the training of aspiring researchers who work in either the broad quantitative or qualitative empirical traditions. (We remind readers of the problems inherent in labels here, which were discussed at the end of the excerpts from the original proposal that were reproduced in the opening pages of this document.) We stress the narrowness of our focus; essentially we are discussing the training of the subset of researchers who are appropriately regarded as working in an empirical social-scientific mode – and clearly there are many researchers whose mode of work does not fall under this rubric, and who are contributing valuable insights about educational processes and practices, and more.

Our recommendations about initiation of the AR into Universe 1 require relatively lengthy exposition and justification; the remaining recommendations require less by way of commentary. At the end of this Section, in 10.6, we provide a summary of the major recommendations made throughout this Report.

### 10.1 Recommendations concerning initiation of ARs into Universe 1

The table of “basics” that appears below – that summarizes what we regard as the basic level of understandings that serve as initiation into Universe 1 – was derived from a variety of sources. The vignettes of the graduate careers of Denise and Catherine provided some of the inspiration, as did the contents of some of the first-year doctoral seminars mounted by RTG institutions over the years since the mid-1990s and also the way in which these institutions have from time-to-time framed their introductory methodological training. The table was also influenced by an analysis of the major theoretical and epistemological issues that lie behind the disputes in some of the leading education journals of the past decade or more, together with some of the more lively and better-attended sessions at AERA and other conferences including some of the Presidential Addresses. Some work of prominent contemporary philosophers of social science has also been drawn upon; as have reflective statements by major educational researchers themselves.

The ever-present temptation when thinking about the ideal training for doctoral students is to aim high – to aim for mastery, or for the kind of command of the issues that we imagine ourselves to have achieved. We have attempted to resist this temptation. We have made the point earlier in this report that aspiring researchers need to have reached a level of competence in the theoretical and methodological frameworks within which it is expected that their research will be conducted, but also should be familiar enough with the other main tradition of empirical work to appreciate how it complements their work or answers questions that their own approach is less effective at pursuing. We suggested, in addition, that having a basic understanding of the epistemological and related issues will not only enable aspiring researchers to respond to the criticisms and questions they might face in Universe 4, but will also result in their own research being strengthened. Clearly we subscribe to the view that researchers who have not thought about the presuppositions underlying their own research methodology do not fully comprehend it – they have been trained, but not educated; to put it aphoristically, *it is better for research to be conducted self-consciously rather than unconsciously*. Or as the philosopher of social science Harold Kincaid put it in his book analyzing controversies in “social research”,

Since my ultimate hope is to contribute to better social research .... I hope to show that philosophy of science and philosophical analysis can help clarify important social science controversies .... *Broad philosophic theses about science are in the end part of science itself.* (Kincaid, 1996, pp. xvi, 2, emphasis added)

We stress again that in the paragraph above we used the expression “basic understanding” when referring to the ARs’ initiation into the epistemological and related controversies about educational research, and into other research traditions. We are not arguing that these prospective researchers need to become competent philosophers and expert in all methodologies – but our hope is that they will become *reflective researchers*, who have sufficient understanding to be able to follow the arguments and who are poised to undertake further reading should they be so motivated. (The members of the TF, of course, had to use their own professional judgment about what, specifically, is required for this basic level of understanding.) Throughout our deliberations we also were cognizant of the objection that too much reflection, too much self-consciousness, can be stultifying – the centipede who was asked to reflect about how it was able to move its legs without tripping up became incapable of movement.

Finally in this preamble we stress a point that we made in “boxes” in the vignettes, namely, that we are committed to the view that training programs should move away from a model based entirely on undergraduate education – where all or certainly most training takes place within a traditional course structure. We believe that at the graduate level, and even when the “basics” are being considered, there is great value in training that occurs in conjunction with faculty research projects, and in the offering of “just in time” modules and the like. In these types of learning situations the importance and relevance of the “basics” is likely to be more apparent than is the case in traditional classroom settings. We also believe that institutions should find ways to give faculty members conducting this type of training adequate course-credit.

The table summarizing our recommendations about Universe 1 is as follows:

**BASICS WITH WHICH ALL ARs IN EMPIRICAL, RESEARCH-ORIENTED  
QUANTITATIVE OR QUALITATIVE FIELDS SHOULD BE FAMILIAR**

**Justification for inclusion of each of these is given below, together with a specific account of what, precisely, each entails.**

1. METHODOLOGICAL RIGOR IS AT BEST ONLY ONE FEATURE POSSESSED BY QUALITY RESEARCH: Need for social/ educational relevance, and scientific relevance.
2. EPISTEMOLOGICAL ISSUES AND ASSUMPTIONS OF DIFFERENT RESEARCH "FRAMES"
3. GUIDING VIRTUES OR PRINCIPLES UNDERLYING RIGOROUS EMPIRICAL INQUIRY IN ALL FIELDS, AND RELATED CONTROVERSIES
4. SUCH CONSTRUCTS AS: A WARRANT, THREATS TO VALIDITY, CONFIRMATORY VERSUS DISCONFIRMATORY ORIENTATIONS, ETC.
5. PROS AND CONS OF "TRUE" EXPERIMENTS, Campbell and Stanley, ESTABLISHING CAUSAL INFLUENCE; OTHER FOCI OF RESEARCH APART FROM "Did this treatment/intervention cause this effect?"
6. THE CONCEPT OF "EDUCATION"; THE ROLES OF VALUES IN EDUCATION AND ED. RESEARCH; ARE VALUE-NEUTRALITY AND OBJECTIVITY POSSIBLE?
7. THE RESEARCH PROCESS IN OVERVIEW, AND SOME EXPERIENCE ACTUALLY DOING THE FOLLOWING: FORMULATING A RESEARCHABLE QUESTION, DESIGN AND ITS PURPOSES, COLLECTION OF RELEVANT DATA, AND ANALYSIS
8. THE ETHICAL TREATMENT OF HUMAN SUBJECTS, AND THE RIGHT TO PRIVACY
9. CONSUMER CONCEPTUAL UNDERSTANDING OF RESEARCH METHODS
10. IN-DEPTH TRAINING IN QUANTITATIVE METHODS
11. IN-DEPTH TRAINING IN QUALITATIVE METHODS
12. FAMILIARITY WITH THE MAJOR SUBSTANTIVE THEORIES (APPLICABLE TO EDUCATIONAL PHENOMENA) AVAILABLE IN THIS UNIVERSE

Note: the TF suggests that these "basics" should have a place in the education of all graduate students – philosophers and historians of education, policy and curriculum scholars, etc.

*Each of the items in this table requires further discussion.*

(1). Quality research is marked by more than methodological rigor. Earlier we made the point that, in our experience, most ARs (including Denise and Catherine) are not explicitly familiarized with the characteristics that delineate pieces of quality research (the ones discussed in Section 5), and only by happenstance are their dissertation topics evaluated in the light of these characteristics. Most, if not all, ARs are of course aware that their work must display methodological competence, but the need for it to possess scientific or social and educational relevance is often not stressed. We therefore urge that venues be found where ARs can be held accountable with respect to such desiderata; basic research design courses are probably a likely venue, as are the dissertation proposal writing seminars or workshops, and the proposal hearing. In addition, many members of the TF believe that familiarity with Donald Stokes' book *Pasteur's Quadrant* (1997) would be beneficial, for it stresses that important scientific work can have its origin in an investigation of problems that arise within some realm of practice – in essence it makes the point that the theory/practice dichotomy is a false one (and this dichotomy is one that often plagues discussion of educational research). Pasteur is an example of a researcher whose work had *both* social/practical and scientific importance.

(2). Epistemological issues and assumptions of different research frames. Across the sciences – including qualitatively and quantitatively oriented social science and educational research – major research traditions are underwritten by a foundation of assumptions or hypotheses about such matters as the nature of knowledge, about the types of evidence that are relevant, about the nature of the causal agents or causal processes that operate in the broad domain that is the focus of the research frame, about the kinds of generalized claims if any that in principle can be made. As Kincaid pointed out, these epistemological premises “are in the end part of science itself”, but what he meant of course was that, in final judgment, they must be seen as part of science *from the beginning*. Skinner's influential behaviorist program of research could not have begun without its foundation of views about such matters, and neither of course could Freud's, or Pasteur's, or Newton's.

The TF therefore recommends that all ARs in both of the empirical domains that fall within our charge should become familiar with the foundational assumptions of both these traditions; and the TF further points out that numerous accessible journal articles and monographs are available that cover this material and that could be used with profit as a basis for classroom discussion and argumentation and application to research reports.

The broad quantitative tradition – which label covers for us the experimental, quasi-experimental and correlational modes of inquiry – focuses upon observable/measurable factors or entities, and the aim of inquiry frequently is either the production of generalizations (which can be of probabilistic nature) that link these factors or the testing of hypotheses about the nature of the relationship between such factors. The broad qualitative/ethnographic tradition is committed to the view that human behavior, while it clearly can be influenced by social forces and socio-cultural norms and practices, is also deeply influenced by such things as personal intentions, individual understandings and meanings and beliefs and situational appraisals (which is not to deny that some aspects of these can be studied by way of quantitative experimental and quasi-experimental means). Adherents of both these traditions believe that they can offer good warrants for the validity of the data that they collect about such factors or entities, and also for the claims they make about the relationships that hold between these. These warrants should be assessed, and also worth pursuing is the issue of whether or not these different traditions are incompatible paradigms or are complementary, or whether they are focused on answering different questions.

(3). Guiding virtues or principles of scientific inquiry in all fields. We open our discussion of this item with a homely analogy: An individual can be a serviceable cook without giving consideration to a set of principles or to underlying scientific or even aesthetic theories; however, it would be difficult to find a professional chef who did not bring these perspectives to bear on his or her work. And for good reason: production of high-quality artifacts belonging to a particular genre, particularly novel ones, basically is blind if it proceeds without the guidance provided by a deep understanding. Members of the TF feel that ARs are in a comparable position; if they have thought about the principles of, or virtues associated with,

good science, they will have deeper insight into their craft – and if they ignore or contravene one or more of these principles, hopefully they only will do so consciously and for good reason.

We recommend, then, that graduate students should be exposed to the principles and virtues listed below, which it has been argued apply to empirical work in both modes. These should be discussed (perhaps in the first-year interdisciplinary seminar), and it would be beneficial to apply them in an analysis of several pieces of research drawn from across the two empirical traditions that form our focus in this report and that are highly regarded by program faculty.

*We also stress that these items are not sacrosanct; several have been the subject of considerable controversy, and the analyses and classroom discussions mentioned above might well reveal that some of the principles apply more readily (or even only) to one of the two major research frames rather than to both. In other words, the claim that these principles or virtues are crosscutting should itself be a focus of discussion.*

We provide several listings of the guiding principles or virtues; later we make some brief comments about each set. The first of these lists is taken from the NRC report (2002) referred to above, that has served as a lightning rod for much of the recent debate about the desiderata of high quality educational research. The charge to the NRC committee enjoined it to focus upon (and only upon) “scientific educational research” – which the committee interpreted to mean “empirical educational research” (as the first of its principles, summarized below, makes clear). Thus the NRC committee and our Task Force have had a similar focus. The opening of their report asserts:

The scientific enterprise depends on a healthy community of researchers and is guided by a set of fundamental principles. These principles are not a set of rigid standards for conducting and evaluating individual studies, but rather are a set of norms enforced by the community of researchers.... (NRC, 2002, p.2)

These six “guiding principles” are as follows (NRC, 2002, pp.3-5):

- i. Pose significant questions that can be investigated empirically
- ii. Link research to relevant theory
- iii. Use methods that permit direct investigation of the question

- iv. Provide a coherent and explicit chain of reasoning
- v. Replicate and generalize across studies
- vi. Disclose research to encourage professional scrutiny and critique

Harold Kincaid, in his discussion of good science that we examine below, suggests another principle that we believe should be added here:

- vii. Use fair tests, independent tests, and cross tests (e.g. triangulation) (Kincaid, 1996, p.51).

We also take the liberty of adding a further principle:

- viii. Take steps to counter, or if not possible then to document the effects of, threats to the validity of the study

Kincaid also usefully clarifies an ambiguity in usage of the term “science” that can be troublesome; he distinguishes between what he calls “the symptoms of good *processes*” and “the symptoms of good *products*” (Kincaid, 1996, p.49). The NRC report in essence attempted to provide a list of the former, as is made clear when it is stated that the focus is on principles (which the report also calls “norms”) for “conducting...studies”; whereas Kincaid himself provides a list of the latter. A good scientific product has several kinds of “symptoms”, but it is his remarks on only one of these – the evidential virtues – that we draw upon below; these virtues pertain to how well the data or evidence that is cited supports the theory that is being proposed or the claim that is being made. “Good science”, he argues, “typically has at least the following evidential virtues” (Kincaid, 1996, p.50). For ease of reference we continue the numbering from above:

- ix. Falsifiability
- x. Empirical adequacy
- xi. Scope
- xii. Coherence
- xiii. Fruitfulness
- xiv. Objectivity – science is objective, he states, “when our beliefs reliably indicate the way the world is rather than the way we want the world to be” (p. 51).

Our general comments about these lists are as follows; we turn first to the NRC list as supplemented above, where the focus is upon *the process of research*. While we generally endorse the items presented, we are more enamored of some and regard it as an overstatement to claim that others are “fundamental principles” of all scientific inquiry. Insofar as both qualitative and quantitative researchers are doing empirical work, it seems that the first principle is uncontroversial, but the second – linking research to theory – is too strong, for even in the most “hard-nosed” sciences some puzzling phenomena are investigated for which no theory is available at the outset. It rarely happens, if it happens at all, that an investigator approaches a problem as a *tabula rasa*, without some conceptual apparatus and presumptions; but nevertheless the state of the knowledge base may be such that there is no specific theory available that he or she explicitly can link the research to (see the discussion of these issues in Clement, 2000). A clear example is provided by the work of Roentgen, who undoubtedly was doing science when he investigated puzzling phenomena discovered by accident in his laboratory, and indeed the name he gave to the unknown causal factor that he was pursuing – “X radiation”, where of course X signified the unknown – was a clear indication of the lack of a specific theory to which he could link his work. This of course is not to deny that he was working within a framework of background assumptions about the kind of factors and mechanisms that operate in the physical universe. But the objection can be made even more strongly: some areas of rigorous empirical inquiry do not aspire to produce theory at all, but rather are focused upon gaining deeper insight into the factors at work in a specific setting -- in short they are descriptive, a process that might, or might not, lead to the production of a specific theory in the long run. Thus, we also believe item (v) on the list (replicate and generalize) is questionable. And while researchers should reason coherently as they work (iv), providing a coherent and explicit chain of reasoning is probably a principle that applies to the written product rather than to the actual process. The other items are ones we happily endorse; we postpone until a little later our discussion of the item we added concerning threats to validity.

Turning briefly to the list of evidential virtues provided by Kincaid, where the focus is upon the *products of research*: We endorse the items on this list, and believe they apply to work in both the traditions of empirical research that we are focusing upon. Thus, whatever the research mode, if empirical inquiry results in the putting forward of claims that in principle are not testable, then these must be regarded as scientifically dubious. While Popper is the philosopher who is

most closely identified with testability/refutability as the demarcating criterion of science, many practicing scientists also advocate it – witness the debates within the field of cosmological physics over “string theory”, which is controversial because it is not clear that it can produce testable predictions. The second virtue, empirical adequacy, seems difficult to dispute; a research product such as, for example, a theory or hypothesis that can only account for some, but not all, of the relevant evidence that is available, seems inadequate – although we note that occasionally products get modified over time and sometimes acquire the capacity to account for the evidence that had previously been cast aside. Items (xi) and (xiii) are often overlooked, but it seems clear that, all things considered, research products that have wide scope and which open up fruitful lines of future inquiry, are much to be preferred over those that are narrow and sterile. As signaled in the paragraph above, we also would merge item (iv) from the NRC list (provide a coherent chain of reasoning) with Kincaid’s item about coherence (xii). Finally, we point out that much has been written about the connection between several of the items on both lists; thus, one way (some would argue it is the only way) to ensure a research product is objective (item xiv) is for it in principle to be refutable (item ix), and for it to actually have been tested (item vii) and exposed to professional scrutiny (item vi).

(4). Warrants, threats to validity, confirmatory versus disconfirmatory orientations. Members of the TF believe that an understanding of the concepts listed here is of vital importance for both qualitatively and quantitatively oriented researchers – they open the door to rigorous work in both modalities. Furthermore, these ideas can be introduced in venues such as the first-year doctoral (“Spencer”) seminar, and they can provoke lively discussion when used to analyze real cases drawn from the research literature. We offer a brief remark about each of these three:

John Dewey (1938) popularized the notion of an epistemological warrant, but its home is in the legal sphere. Essentially a warrant is the argument or body of evidence that is marshaled (i.e. the case that is made) to support or justify either a conclusion or knowledge claim or judgment that is being advanced as a result of a research project, or a course of action or policy that is being recommended on the basis of research findings. (It should be noted that in his classic work *The Uses of Argument* (1964/2003), Stephen Toulmin developed a different concept of “warrant”.) Warrants (in the Deweyan sense) need to have the evidential virtues listed earlier,

in particular they need to embody coherent, explicit, and valid chains of reasoning. There are grounds for believing that research reports are frequently deficient in this respect, putting forward conclusions that *are not* warranted by the evidence that has been gathered or by the argument that has been presented (see Tooley and Darby, OFSTED, 1998).

The classic treatment of threats to validity is, of course, Campbell's and Stanley's *Experimental and Quasi-Experimental Designs for Research* (1963), and there can be little doubt that all ARs who are adopting a quantitative research orientation should be familiar with the common types of threats to their work. The TF emphasizes the point that qualitative modes of inquiry also face threats to validity; in Maxwell's words: "Qualitative researchers deal primarily with specific threats to the validity of particular features of their accounts, and they generally address such threats by seeking evidence that would allow them to be ruled out. In doing this, they are using a logic similar to that of quasi-experimental researchers..." (Maxwell, 2002, p. 56; on this general topic see also Eisenhart and Howe, 1992). Thus, when *any* type of empirical research is being planned, the rigor of the study can hardly fail to be improved if the likely threats to its validity are identified, and if the design or data collection then is modified in such a way as to allow them to be nullified or deflected so far as is possible.

It is human nature to look for, or to more readily notice, evidence that we are right rather than that we are mistaken. But, as Popper once remarked, any fool can provide evidence that he is right – even the most wrong-headed of viewpoints can be "confirmed" with some grains of evidence; what is important is whether or not disconfirming evidence has been actively sought. And, after all, the disconfirmatory orientation goes hand-in-glove with the principle that testing, criticism, and openness to refutation are major scientific desiderata; and it is also worth noting that objectivity inheres primarily in this disciplining of subjectivity through the dogged pursuit of disconfirming evidence.

(5). "Pros" and "cons" of the randomized field trial (RFT); establishing causal influences; and other foci of research. The TF does not need to emphasize the fact that the RFT or true experiment – the so called "gold standard" research design – has, for more than a decade, been the focus of the most vigorous and sometimes intemperate debates within the research

community (that is, within Universe 4). We believe that all educational researchers need to know where they stand with respect to the issues at stake – not only does the future direction taken by the overall research enterprise depend upon how the disputes are resolved, but in clarifying his or her stand on the issues here, the researcher will also gain clarity about the place of his or her own work in the research terrain.

The logical underpinning of the true experiment (which derives from J.S. Mill's "methods of logic"), and a sense of the various forms that experiments can take, and familiarity with related matters such as threats to validity, are all part of the basic literacy that should be expected of all researchers; parts at least of the classic foundational monograph by Campbell and Stanley (referenced above) should be required reading. But the AR should also have considered whether there are other ways to warrant causal claims, especially in qualitative work – after all, in legal matters and in everyday life, we make many successful causal attributions, some with major consequences, without conducting randomized experiments. In addition the AR should have reflected about whether causal knowledge has a high priority in his or her own field of work, and about the tradeoffs that are involved on settling upon a research design-- whether experimental, quasi-experimental, or qualitative and observational. Other possible aims of scientific inquiry, apart from establishing causation, could well be discussed in this context (see Phillips, 2005).

(6). The concept of "education", and the role of values in the research process. At first sight, this item seems to be moving the focus away from the contents of Universe 1, and even away from the life of the researcher. The TF believes, however, that a multi-pronged case can be made in support of this item. First, just as it is difficult to see how researchers working on such topics as acquisition of analytic ability, or the relation between intelligence and creativity, could make much headway if they lacked an understanding of (or held a stunted concept of) "analytic ability" or "intelligence" or "creativity", it also is difficult to see how a researcher inquiring into educational processes and programs and the like could do effective work absent an understanding of the ramifications of the concept of education.

While this topic could well fall within Universe 3, it becomes of pressing concern in Universe 1 in the following way: education is a value-laden activity, in the sense that when we educate a student, we are making *changes* in that individual which are judged to be *valuable* within the particular socio-cultural setting in which we and the student are located, which is not to deny that different subgroups – for example those that exist in Universe 2 – make different judgments about these values. Furthermore, it is argued by many philosophers – and crucially by many researchers themselves – that these, and other values, inevitably leach into the work of the researcher and sometimes subtly influence such things as what is measured or observed, how these things are measured or observed, how the data is analyzed or codified, or how the conclusions are presented. In other words, the whole contentious topic is opened-up of whether or not objective inquiry is possible – for objectivity and value-neutrality are closely related.

At stake here are so-called “non-epistemic” – socio-political and religious – values; there is little if any dispute that the knowledge-producing activities of science do, and must, embody “epistemic” values such as intellectual honesty, precision and clarity, openness to negative evidence (see Phillips, 2000, ch.13). Quantitative researchers are often accused of being too dogged in their commitment to value-neutral inquiry, whereas qualitative researchers frequently are charged with being too open to the influence of values; but of course there are some individuals in each camp who hold a different view from that of their colleagues. The AR needs to be reflective about the issues here, and to have a well thought-out position, and certainly needs to be able to respond to interlocutors he or she will meet in Universe 4.

(7). Having an overview of, and some hands-on experience, of the research process. Members of the TF regard this item as being quite uncontroversial; early in their graduate careers the ARs should gain an overview of the various stages of research, and preferably should have some direct hands-on experience. In the vignette concerning Denise, the TF identified seven aspects of research that the quantitative AR should be familiar with; we also believe that a broadly similar but not quite identical list would be applicable to a qualitatively-oriented AR, and so we offer the following more-encompassing breakdown of the research process, although we also stress that some of these stages are not always temporally discrete:

Defining the purpose of the research\*

Framing the research questions\*\*

Selection of research populations or study sites

Developing defensible study designs

Collecting data

Conducting statistical (or other) analysis of data; or data reduction

Interpretation of, and writing about, the results

Presenting an overview of the study and findings for various audiences

Protection of human subjects during all phases of the research

\* We note here that the TF accepts the fact that inquiry is initiated by a problem or puzzle (as Dewey and Popper, among many others, have emphasized); the first efforts of the inquiry itself are directed at – or should be directed at – clarifying the precise nature of the problem, which of course directly relates to the purpose of the research and to the research methods that are appropriate.

\*\* We note here that in some modes of qualitative inquiry, the research questions actually evolve or become more refined during the research process itself, as the researcher becomes more deeply familiar with the setting being studied.

Experience in these aspects of research could come by way of assistantships, or via class-based projects, or in a pilot study prior to the dissertation, but the TF prefers the experience to come earlier rather than later. It has been surprising to TF members to discover that familiarity with the unfolding research process is far from universal. Thus, for example, some TF members report having met students who, when almost at the end of their dissertation research, have inquired about how to do a literature review – and, even more disturbingly, who seem not to comprehend its purpose; and they report that it is not uncommon to meet students who, at the beginning of their dissertations, have never been into the field to collect data.

(8). Ethical treatment of human subjects, and the right to privacy. We noted earlier our concern about the degree of exposure that Denise and Catherine had to the issues in this domain – although we believe that, given the qualitative nature of her research, Catherine together with

her peers almost certainly must have discussed the dilemmas that their field work had revealed. On the other hand, we also note that sensitivity to this important item has grown in recent years, and most ARs become acquainted, as a minimum, with their institution's procedures for obtaining clearance to enlist human subjects in their research. But we stress the point we made earlier – acquaintance with IRB procedures is a rather sparse minimum. But even with respect to the IRB, we have concerns that some students might “fall between the cracks”; for example not all students working on a large project may have had direct experience of obtaining IRB approval; and students who do not need to directly collect human data for their research may – but should not – remain in ignorance. Furthermore, not all students fully appreciate the ethical issues at stake here (such as the ones we discussed briefly in the opening portion of Section 8) and may regard seeking IRB approval as “red tape”, and may have little inkling of the full range or depth of the ethical issues that can arise in the course of research that involves human subjects. We therefore recommend that this general item be included as a topic in research methods and design courses, or perhaps in the first year seminar or in a brown-bag session; furthermore, we recommend that classes in which cases of empirical research are studied, attend to the kinds of ethical issues that we alluded to in Section 8 that are entailed by the specific designs of those studies.

(9). Consumer Conceptual Understanding of Research Methods. All ARs need to be able to read, understand, and critically evaluate both quantitative and qualitative types of research. This conceptual and consumer understanding is necessary, but not sufficient, for the conduct of research and demands a different type of skill than that required for designing, implementing, and analyzing research. Students at this level should have conceptual understanding of basic issues in how and why studies are designed and how and why data are collected and analyzed. In addition, students at this level should also have a conceptual understanding of issues related to reliability, validity, trustworthiness, credibility, ethics, and warrant. A number of texts are available that focus upon developing what is often called a “reading level” understanding of statistical techniques or of a broader array of research techniques; see for example Huck, 2000; Locke, Silverman and Spirduso, 2010.

Some RTG institutions have approached the need to build consumer and conceptual understanding by offering a course that covers the concepts described above. Such a course is then followed by a series of courses or experiences that provide the opportunity to learn the technical skills associated with various methods. These more advanced experiences and courses allow students to specialize in concepts and techniques described in more detail in the following sections.

(10). Quantitative Methods. It is obvious that those ARs who have chosen to work within a quantitative framework in order to pursue their substantive educational research interests, must start with a basic level of statistical literacy and then build upon this with more advanced study; many such students will enter the graduate program already at, or beyond, an introductory level. The TF recognizes that there are institutional differences in philosophy concerning the way to approach technical training, and also it is clear that there are differing sets of constraints that affect what courses or experiences it is possible to offer. But with respect to content, the TF recommends that – by whatever means – all ARs aspiring to quantitative research careers obtain in-depth training in the following areas:

- definitions of, and distinctions between, major quantitative approaches
- concepts and methods related to probability, hypothesis testing and alternatives to hypothesis testing, error rate and power analyses, null hypothesis testing and significance levels;
- relevant design and data collection techniques, such as survey and sampling design
- analytic methods such as descriptive statistics, statistical techniques for determining reliability and validity, correlation and regression, testing of means through ANOVA and related techniques, multivariate statistics, and advanced statistical techniques appropriate to his/her research (e.g. HLM, SEM)

The Table following the two vignettes summarized Denise’s technical training, and the TF recommends this for general consideration.

(11). Qualitative Methods. Just as ARs pursuing careers as quantitative researchers must possess high-level statistical skills, those who hope to pursue research in one of the qualitative

traditions must develop high-level qualitative research skills and practices. However, the TF recognizes that there are several complexities that need to be negotiated here: First, there is a variety of “families” of qualitative inquiry, and each of these has its preferred data, its own way of data collection, and oft-times its own data-analytic techniques. Second, nationally there are two major approaches to teaching introductory qualitative inquiry (a helpful national survey was conducted some years ago, see Webb and Glesne, 1992, and the account below draws upon their analysis of these); we met both of these approaches in the graduate histories of Denise and Catherine.

In one approach, which can be called “skills-based”, students are required to undertake a mini-project that involves gaining access to a field-setting, making observations, taking field notes and then analyzing the data and writing a research report. The second approach, that may be labeled “learning about qualitative inquiry”, and which does not involve field-based research, has two variants:

(a) “Design courses” that focus on “some underlying principles of qualitative research and help students plan qualitative studies. Students may read and critique qualitative studies” (Webb and Glesne, 1992, p.787). The TF is aware that there are a number of texts that cover these principles, for example Miles and Huberman (1994), and Creswell (1998 and later editions) – this latter source discusses overall design, focusing the study, data collection, data analysis and representation, verification and standards of quality, and writing a narrative report.

(b) “Theory-based” courses that “typically examine the theoretical underpinnings of qualitative inquiry. Students in these courses typically read books and articles by philosophers such as Bernstein (1976), Habermas (1971, 1981), James (1890/1950), Natanson (1962, 1973), D.C. Phillips (1987), Ricoeur (1976, 1981), and Winch (1971)...” (Webb and Glesne, 1992, pp.787-8).

In actual practice, many qualitative courses provide both the opportunity to learn about qualitative research and to learn the skills of qualitative inquiry. However, the TF recommends that ARs, with the exception of those who already have an equivalent background or who are launching into a sequence in anthropology or qualitative sociology or the like, avail themselves

of courses or experiences related to both major types -- “skills based”; and “learning about,” although time-constraints might make this difficult. Topics with which all ARs should have in-depth training, in the judgment of the TF, include:

- definitions of, and distinctions between, major qualitative approaches
- interview types, interview design strategies, and interviewing procedures;
- observation protocol design, observation strategies, observation timelines;
- analytic methods such as taxonomic analysis, domain analysis, thematic analysis, constant comparison (“grounded theory”), narrative analysis, discourse analysis, artifact analysis, gestural/proxemic/kinesic analysis.

In closing this section, the TF feels compelled to note that developing the range of skills necessary for the conduct of both quantitative and qualitative research requires multiple courses, workshops, and experiences. In some cases, just-in-time methodological modules might be used, where a method or technique is taught in conjunction with its use in a research project where real data are involved. However, the TF recognizes that the costs and management of such modules and workshops may be prohibitive for individual institutions. One possibility for addressing this resource dilemma is for schools of education to collaborate in offering these alternative learning opportunities.

(12). Knowledge of substantive theories. As we have mentioned earlier, discussions of research training almost automatically gravitate to discussions of research methodology, and the TF wishes to stress the importance of ARs having a strong grounding in the substantive theories that are relevant to the broad domains (the “problem spaces”) in which they will be working. Theories provide direction, and also a conceptual apparatus without which the research process could not proceed fruitfully. Sadly, however, no research-training institution has unlimited resources – and many are only “one deep” in some fields - so there are constraints on the number and variety of theoretically-oriented courses that can be offered. Some RTG institutions have been able to supplement their offerings by way of collaborative relationships with departments across campus, or in a few cases with other nearby universities with which “cross enrolment” is allowed.

## 10.2 Recommendations concerning initiation of ARs into Universe 2

There is a limit to what the AR can learn about Universe 2 – about the complex and often discordant socio-political and educational contexts in which the institutions and processes and phenomena that are the focus of most educational research are situated – from course work and reading. The basic recommendation of the TF here is that ARs should have some direct familiarity or experience of the contexts that are related to the topic of the research that they are (or will be) studying. This can be attained in a variety of ways.

First, we recommend that relevant educational experience prior to entering a graduate research-training program be a criterion that is given due weight during the admission process. Hard data probably does not exist, but it is our impression that prior exposure to social science research methods, or indicators that the candidate has the facility to master these methods in the graduate program coursework, is often given more weight. While we acknowledge, of course, that entering students should be of sufficient caliber to successfully navigate methodological training, we point out that contextual sensitivity and sophistication cannot be fully imparted in coursework, so program faculty need to be aware of the tradeoffs when evaluating the prior academic versus the professional qualifications of applicants.

Second, we point to the important role of advisors and course instructors here. The advisor should have as a priority the identification of internships and research assistantships that would move the AR out into “the field” or that would partner the AR with a practitioner (some projects, such as those in the area of so-called “design research”, are conducted in conjunction with practitioners); this is especially necessary when the AR lacks prior contextual experience. Furthermore, both the advisor and course instructor should encourage the AR to shape course projects and assignments when the structure of the course allows this so that he or she can move out into schools, or observe school board meetings, and the like.

Third, program faculty could usefully organize a lecture or colloquium series involving “outside” speakers who could alert the ARs to some of the intricacies of Universe 2 that impact the phenomena that are being researched.

### **10.3 Recommendations concerning initiation of ARs into Universe 3**

In contrast to Universe 2, Universe 3 is more amenable to being opened-up by means of coursework, directed readings and the like. Indeed, it is the impression of the TF that our respective RTG institutions fare quite well here, and we do not feel a need to make specific recommendations. Examination of the course-offerings at one of our institutions, for example, reveals that about 40% of the two hundred courses regularly offered are what might be called “content” courses that acquaint students with the literature – including research-findings – about one or other area of education; we also note that these content-oriented courses often present opportunities for integration between Universes 1 and 3. The remaining 60% of courses fall into three categories: (i) those directed at teachers-in-training, and which focus upon teaching techniques, classroom management, assessment of learning, and the like; (ii) qualitative and quantitative research techniques and analysis; (iii) advanced research seminars and colloquia.

### **10.4 Recommendations concerning initiation of ARs into Universe 4**

In earlier sections of this report we highlighted the various ways in which Universe 4 is vitally important for fostering both the intellectual and professional growth of the aspiring researcher, and the vignettes presented some examples from the experiences of Denise and Catherine. Reflection reveals that too often the relevant formative experiences seemed to occur by happenstance – chance attendance at a crucial session at AERA, or a chance meeting with graduate students from other institutions who have the same research interests, or the happy coincidence that a brown-bag session on the intricacies of publishing occurs just as the AR is working on the first potentially publishable essay. The following recommendations are designed to remedy this situation.

First, the TF recommends that the discussion of matters pertaining to professional growth be institutionalized; topics such as writing for publication and for grants, the publishing process and selection of a journal to which to make a submission, grant writing, the preparation of a CV, applying and interviewing for a job, the tenure process, how universities function, and the like, should be covered in an institutionally-

supported and regularly offered brown-bag series, or should be part of the first or later year seminar.

Second, we also recommend that the role of advisors be expanded in at least four respects. (i) The sessions at the AERA and other relevant conferences should be treated as resources in the same way that course offerings at the home institution are treated, and advisors should discuss with the ARs which sessions would be valuable to attend, rather than leaving this to chance. (ii) We also recommend that students be strongly encouraged by their advisors to attend several such major conferences during their graduate career. (iii) We recommend, further, that institutional support be given to pre-conference practice sessions at which students who are on the program receive formative feedback from advisors, other faculty members, and from other students who act as discussants. (iv) From time-to-time advisors should provide opportunities for ARs to discuss important and/or controversial papers in current journals; as well as the direct benefits, this might also foster the habit of keeping abreast of current literature and the controversies therein.

### **10.5 Recommendations concerning assessment of research training programs**

We now turn to address a part of the charge to the TF that we have not mentioned in the foregoing discussions – the difficult but extremely important issue of how the effectiveness of research training programs can be assessed.

The difficulty arises in large part because the focus of these training programs – the ability to produce high quality research – is technically a *capacity*; furthermore, it is a capacity to perform a very complex task, and this has necessitated a complex regimen of training during which the ARs have had to acquire a deep and wide knowledge base, a set of appropriate skills, and a stock of relevant experience. And there is another associated issue: students do not enter graduate programs with the same background of knowledge, skills, and experience, so the developmental paths that they follow take different amounts of time – some may take many years before they mature into good researchers. But when the AR finally has developed successfully this complex ability,

there remains an important and unfortunate fact of life – especially with respect to a complex performance, an individual who possesses the relevant capacity or ability may not display it on every expected occasion. Thus a well-regarded stand-up comedian, for example, no doubt has the capacity to tell funny stories – a complex matter that only appears to be simple; but this is compatible with his or her occasionally falling flat. If, of course, this individual was more often un-funny than rib-tickling, we would come to doubt that he or she actually did possess this capacity.

An important consequence follows from this for assessment: “One shot” assessment methods are unreliable ways to determine if individuals have successfully developed the desired capacity or ability. It may also be the case that this could present a problem if the effectiveness of a research-training program is being determined by looking at the quality of a single piece of research done by each person in a sample of recent graduates – for some who have the ability to do fine work might not have displayed it in this particular instance (perhaps they were ill, or working under unreasonable constraints, or just having an “off day” or “off month”). Even worse for the validity of the assessment, some of the individuals who do not possess this capacity might on this one occasion have just happened to have done a good piece of work – even the most unfunny of individuals can occasionally tell a good joke. If the sample of students is large, of course, these countervailing cases might balance out (but they might not); a more secure basis for assessment would be to examine the developmental trajectory of a number of individuals over time, for instance by examining their portfolios, and we return to this below.

Given this problem, there is no single, practicable assessment tool apart from portfolio analysis that the TF can recommend; however, we do believe that a reasonably reliable picture of the effectiveness of a training program also can emerge if the method of triangulation is adopted – that is, if a number of different evaluation instruments and bodies of evidence are marshaled, the conjunction of which yield a fairly consistent picture. And we discovered that a variety of such instruments and bodies of data in fact have been used across the RTG institutions. The following is a discussion of these,

together with a few ideas for evaluation mechanisms that have rarely been put into practice but which could be useful. We begin with portfolios.

- (a) Assessment of portfolios moved into the field of education several decades ago, and it rapidly became a common practice in teacher education programs. Its original home, of course, was the Arts – the complex ability to paint or sculpt, like the ability to do quality research, cannot be assessed validly in a “one shot” manner, for even great artists occasionally produce awful works (in fact, they rarely show all of their productions, but discard many of them) and bad artists can occasionally produce a work of quality. An artist’s portfolio, however – which contains works produced over an extended period – can be the basis not only of a fairly reliable assessment of his or her ability, but also can indicate the developmental trajectory. Similarly with the portfolio of an AR; the assessment of how successfully the ability or capacity for the complex activity of producing quality research has been fostered is certainly much more valid (than looking at a single item, a single “painting”) when based on a portfolio containing such items as class papers, transcript of course of study, research reports, conference presentations, qualifying exam papers, mock or genuine funding applications, responses to criticism, a self-report or reflection on emerging skills and interests and perceived weaknesses that remain, dissertation proposal, evaluations as a research and teaching assistant, and more. One TF member, from a School where faculty members come together in groups to inspect student portfolios, reported that “Our portfolio reviews tell us a lot about our program. In addition, this group assessment of individual students leads to more programmatic work in the sense that we start to see places where we agree and disagree on standards and benchmarks of AR development, and where our courses and other experiences do and don’t attend to important knowledge and skill development.”
- (b) Traditional course assessment. Virtually all institutions have some formal mechanism for evaluating both the content and the adequacy of the teaching in courses; the usual format is a questionnaire that is filled-out anonymously by

students at the end of the quarter/semester. While these instruments do not assess the degree of successful learning that has taken place (some critics have argued that this should be the focus of course evaluation), they do give information about how students judge the effectiveness of the instruction, the interest and relevance of the content – and how much they like the professor.

- (c) Evaluation of advising and/or research mentorship. Only a few institutions have experimented with questionnaires, paralleling those in (b) above, that are focused on student assessment of advising or the quality of the mentored research experience. The problem in both of these cases is, of course, the fact that confidentiality cannot be guaranteed, for even if the questionnaire is anonymous the number of advisees or research assistants a particular faculty member has is relatively small so that respondents often can be identified. Students, however, are so dependent upon the goodwill of faculty in writing letters of recommendation and the like that they usually cannot risk giving frank negative evaluations. Some programs attempt to obviate this problem by collecting anonymous evaluations and aggregating them for a whole cohort of students; it is sometimes found, however, that if the news even in aggregate is bad, individual faculty members can attribute the poor advising to their colleagues but never to themselves!
- (d) Exit interviews or surveys, and/or survey of alumni. Exit interviews are becoming more common; successful graduates looking back at their programs can often give valuable insights, and may even be willing to comment upon quality of advising. Surveys of former graduates after they are several years into their professional careers can be even more informative, for then they have a more informed basis upon which to base comments and suggestions – the strengths and weaknesses of their research training is more apparent to them, and they often are more inclined to give their frank judgments.
- (e) Transcript analyses. Appointment from time-to-time of a small faculty committee representing different research orientations, to analyze transcripts of students who have recently completed their programs, can reveal general patterns of

strength or weakness and of lacunae in training -- it can reveal patterns of neglect or of avoidance of courses in certain areas, and the like.

- (f) Funding proposal exercise. At least one institution makes funds available for dissertation research expenses (the funds have come from the Spencer RTG grant and from internal university resources), but in order to be awarded a grant ARs have to submit a funding proposal that – although of limited length—is of the same kind that established researchers submit to outside agencies. A small committee evaluates these proposals, many of which are at first rejected; detailed feedback is provided to the AR and to the advisor, and a re-application is invited. From time-to-time the funding committee has reported its impressions of the quality of research training at faculty retreats and the like.
- (g) Survey of a sample of dissertation proposals. While there is something to be said in support of the evaluation of completed dissertations, there are significant practical obstacles. Evaluating even just a subset of all dissertations completed in a given year is a formidable, lengthy if not burdensome task; faculty can be prone to “protect their turf” and do not welcome discussion among their peers of dissertations on which they have “signed off”; and of course faculty in most Schools of Education do not have the expertise to evaluate in detail research that is conducted in a theoretical or methodological framework that differs markedly from their own. These difficulties are somewhat diminished when the focus shifts to the evaluation of dissertation proposals. Experience at a small number of RTG institutions has been that the results of this type of exercise have provoked intense discussion at faculty meetings and retreats.
- (h) “Outside” member of dissertation defense committee. Some institutions have instituted the practice of appointing a “disinterested” faculty member -- one who comes from “outside” the dissertation committee itself, often from a different program and sometimes from a different School in the university. The rationale is that this person might prevent the AR and the dissertation committee from becoming “inbred” and insulated within (as Jerome Bruner put it) “their own parish of authorities”. This is not in essence a device for program evaluation, but it can nevertheless sometimes result in higher quality research products.

- (i) Gathering statistical indicators: Each of the following can be informative, but not definitive: Numbers of ARs presenting at conferences; numbers who apply for, and who receive, Spencer pre-doctoral dissertation grants; numbers who, in the years after graduation, apply for and receive Spencer post-doctoral fellowships (the number of alumnae who were unsuccessful applicants is, of course, extremely difficult to obtain); numbers of current ARs who have publications in refereed journals.
- (j) “Visiting” committees. A number of universities now require that, on a regular basis (ten years perhaps), and usually in conjunction with a “self study”, all of their major academic units (Schools or Divisions, and sometimes departments) establish a committee of relevant outside experts. This so-called “visiting committee” evaluates all aspects of the academic work of the unit – the overall curriculum; offerings in the various program areas; quality of teaching and advising; strengths and deficiencies in research training (if this is one of the aims of the unit); diversity of the student body; financial and social support of students; and more. Syllabi and dissertations are often examined, and extensive interviews with faculty and students (and sometimes with alumnae) are conducted. Members of the TF who have served on such committees, or whose own institution has been evaluated in this way, testify to its usefulness.

#### **10.6 Summary of major recommendations about the training of aspiring empirical researchers**

For convenience we provide the following listing of our major recommendations; we stress this list is rather brief, and we do not see it as obviating the need to consult the more detailed exposition that we have given in the preceding text. Furthermore, throughout the Report there are many suggestions, examples, and the like that we hope will point to ways that particular institutions might be able to make small but telling improvements to their programs.

### Advising

1. Advisors should proactively organize experiences that integrate the development of aspiring researchers across each of the 4 universes: research paradigms; social context; substantive knowledge; and professional practice.
2. To facilitate the giving of developmental guidance it is recommended that programs use an “intake survey” along the lines of that given in Appendix 4.
3. Program faculty should reach agreement about core courses, the sequence in which most students should take the methodology courses, and the like, and these guidelines should be made available to the ARs.
4. Advisors should provide detailed, probing, honest (even painful!) feedback to students in ways that constructively promote student development.
5. Deans and program chairs should actively monitor the quality of advising, including the development of strategies to evaluate and improve the quality of advising; and from time-to-time there should be faculty-wide discussion of advising issues and practices.
6. Advisors and programs should institutionalize collaborative writing groups that promote the honest and probing feedback that can develop student thinking (and enculturate students into the universe of professional practice).

### Development in Universe 1: Research Paradigms

Aspiring researchers should develop a basic understanding in the twelve areas discussed in Section 10.1; these are summarized very cursorily below:

1. Epistemological assumptions and issues underlying different research "frames" used in education research. This includes an understanding of guiding principles and controversies concerning the nature of scientific or empirical rigor.
2. Notions of causation, warrant, validity, and threats to validity within different research frames.
3. The strengths and weaknesses of various methodologies for producing causal inferences, and other legitimate foci for research besides establishing treatment effects.
4. The various roles of values and ethics in research generally, and education research specifically. This includes the value-laden nature of research as it relates to researchers'

concepts of the purpose and value of education, and the ethical treatment of individuals acting as subjects in empirical studies.

5. Statistical literacy, meaning a basic understanding of the ideas of and assumptions underlying descriptive v. inferential statistics, correlation, statistical significance and hypothesis testing, regression, reliability, and validity.
6. Qualitative traditions of data collection and analysis, including ethnography, case study, etc. Also including qualitative approaches to dealing with issues of validity of interpretation.

#### Development in Universe 2: Social Context of Education

1. ARs should develop a basic understanding of concepts of "Education" and the competing ideals and visions of education and their historical influence.
2. ARs should have familiarity and some direct experience of the contexts that are related to the research topics they intend to study.
3. Direct experience in a practical educational context should be one criterion in graduate admissions.
4. Program faculty could usefully organize a lecture or colloquium series involving "outside" speakers who could alert the ARs to some of the intricacies of Universe 2 that impact the phenomena that are being researched.

#### Development in Universe 3: Substantive Disciplinary Knowledge

It is the impression of the TF that our respective RTG institutions fare quite well here, and we do not feel a need to make specific recommendations – except that the plethora of courses and resources available in most institutions make careful guidance by the advisor extremely important.

#### Development in Universe 4: Professional Practice

1. Discussion of professional matters such as writing for publication, the peer review and publication process, grant writing, public presentation, the tenure process, and other aspects of professional growth should be institutionalized through brown bags, first year seminars, or some other widely accessible venue.

2. Advisor's roles should be expanded in at least four ways:
  - a. AERA sessions (and those of other conferences) should be treated as educational resources, and advisors should discuss with ARs sessions that may be valuable to attend (rather than leaving it to happenstance).
  - b. Advisors should strongly encourage their students to attend and present their work at several major conferences during the graduate careers.
  - c. Institutional support should be given for pre-conference practice sessions where students can get formative feedback on presentations from advisors, other faculty, and other students.
  - d. Advisors should meet from time to time with small groups of ARs to discuss controversial and/or important articles in current journals.

#### Assessment of research training program effectiveness

1. Aspiring researchers should keep a portfolio that contains samples of their work, both in courses and on research projects, conference presentations, exam essays, dissertation draft, and the like; program faculty should jointly inspect these portfolios (perhaps yearly) and use this process as a stimulus to discuss strengths and weaknesses in their program, areas in which their expectations differ, and so on.
2. Programs should also use a combination of other evaluation devices to produce, by "triangulation", a picture of the effectiveness of their research training. A listing of these was given in Section 10.5 above.

## APPENDICES

### Appendix 1: Biographical sketches of Task Force members

**Robert Floden** is University Distinguished Professor of Teacher Education, Measurement and Quantitative Methods, and Educational Psychology at **Michigan State University**. Floden received an AB with honors in philosophy from Princeton University and an MS in statistics and PhD in philosophy of education from Stanford University. He has studied teacher education and other influences on teaching and learning, including work on the cultures of teaching, on teacher development, on the character and effects of teacher education, and on how policy is linked to classroom practice. His current research examines teacher preparation and teachers' mathematical knowledge for teaching. Floden has been president of the Philosophy of Education Society, a member of the NRC Committee on Education Research, and an Alexander von Humbolt Fellow at the University of Tuebingen. He received the Margaret B. Lindsey Award for Distinguished Research in Teacher Education from the American Association of Colleges for Teacher Education. Floden's work has been published in the *Handbook of Research on Teaching*, the *Handbook of Research on Teacher Education*, the *Handbook of Research on Mathematics Teaching and Learning*, and in many journals and books.

**Hunter Gehlbach** is an educational psychologist at **Harvard**. He has a general interest in using principles from social psychology to improve teaching and learning in classrooms. His specific interest in social perspective taking focuses on how teachers and students discern the thoughts and feelings of others and figure out how others perceive situations. In addition to this substantive focus, he has a methodological interest in the design and use of questionnaires in research. Before moving to Harvard, Hunter was a doctoral student at Stanford, directed an MA program in Stanford's School of Education that introduced students to social science research methods in education, and was a post-doctoral fellow at the University of Connecticut working on their "Teachers for a New Era" grant.

**Carol D. Lee** is Professor of Learning Sciences and African American Studies at **Northwestern University**. She is President of the American Educational Research Association for 2009-2010 and is a Past-President of the National Conference of Research on Language and Literacy. She is

a member of the National Academy of Education, a fellow in the National Conference of Research on Language and Literacy, and a former fellow at the Center for Advanced Study in the Behavioral Sciences. She received a Ph.D. in Curriculum and Instruction from the University of Chicago. Her research focuses on cultural and ecological contexts of learning, with a specific focus on literacy. Over a nearly 40 year period, she has founded four Pre-K – 12 schools in Chicago. Her most recent volumes include *Culture, Literacy, and Learning: Taking Bloom in the Midst of the Whirlwind*, and *Disciplinary Literacy*. She served on the 2004 NRC Committee on *Increasing High School Students' Engagement and Motivation to Learn*.

**Judith Warren Little** is the Carol Liu Professor of Education Policy at the Graduate School of Education, **University of California, Berkeley**. She is a sociologist whose research and teaching interests focus on the organizational and policy contexts of teaching and teacher learning. She concentrates especially on the nature of teachers' professional community and its relationship to teacher development and school reform, investigating how teachers' interactions with one another in workplace settings and in formal professional development contexts supply resources for teacher learning and the improvement of practice. For the past 20 years, she has taught qualitative research design, methods and data analysis to graduate students in the fields of education, social welfare, public health, and public policy. Professor Little is a member of the National Academy of Education.

**Rebecca A. Maynard**. Dr. Maynard is University Trustee Chair Professor of Education and Social Policy at the **University of Pennsylvania**. Presently, she teaches graduate level courses in research methods, economics, and education policy. She directs the University of Pennsylvania's Institute for Education Sciences (IES) Predoctoral Training Program in Interdisciplinary Methods for Education Research and maintains an active research agenda focused on educational improvement, youth risk reduction, and research synthesis. She has a long history designing and directing large-scale, experimental design studies of innovative policies, programs and practices focused on improving the health, education, and welfare of individuals. She earned her Ph.D. in Economics from the University of Wisconsin.

**Mary Haywood Metz** is Professor Emerita of Educational Policy Studies at the **University of Wisconsin in Madison**. She is a sociologist who uses qualitative methods to study school organization with attention to classrooms, schools, and school districts. She has written about authority in the classroom and school and about the effects of magnet schools on school organization and district politics. More recently, she has written about the effects of community social class on teachers' working lives. Currently, she is writing about uses of institutional theory in understanding school organization, and about ways in which the No Child Left Behind Act has accelerated a larger shift in societal thinking about the purposes and practices of education in the United States. She was deeply involved in the Doctoral Research Program funded by the Spencer Foundation in the School of Education at Wisconsin, and was for several years its Director.

**Elizabeth Birr Moje** is Arthur F. Thurnau Professor of Literacy, Language, and Culture in Educational Studies at the **University of Michigan, Ann Arbor**, where she teaches courses in youth literacy, cultural theory, and ethnography, and mixed methods research. Moje also serves as a Faculty Associate in the University's Institute for Social Research (ISR) and Faculty Affiliate in Latino/a Studies in the College of Literature, Science, & the Arts. Her current NIH-funded study examines the intersection between the literacies and texts youth are asked to learn in the school disciplines and the literacies and texts they engage outside of school. Moje also studies how youth construct cultures and enact identities via their literacy practices outside of school.

**D.C. Phillips (Task Force Chair)** has a Ph.D. (philosophy of education and philosophy of science) from the University of Melbourne. Currently he is Professor Emeritus of Education and, by courtesy of Philosophy, at **Stanford**. He served as Associate Dean and Interim Dean of the School of Education, where he also succeeded Lee J. Cronbach as director of the Stanford Evaluation Training Program. Denis is a member of the U.S. National Academy of Education, a fellow of the International Academy of Education, a past-president of the Philosophy of Education Society, and was a fellow at the Center for Advanced Study in the Behavioral Sciences. For more than two decades his focus has been philosophy of social science/educational research, and he has a career-long interest in the history of 19<sup>th</sup> and 20<sup>th</sup> century thought; he is

author, co-author or editor of eleven books, and has published more than 120 journal essays and book chapters. He was an author of the NRC report, *Scientific Research in Education* (2002).

**William A. Sandoval** is Associate Professor and Head of the Division of Psychological Studies in Education at the Graduate School of Education and Information Studies at the **University of California, Los Angeles**. His interests include the development of scientific epistemologies and their effects on learning and teaching, technological supports for science learning, and design-based research methods in education. His publications span the fields of the learning sciences, science education, and psychology. Dr. Sandoval served on the National Research Council study committee that produced America's Lab Report (2006). He is co-editor of the design research strand of the *Journal of the Learning Sciences*, sits on the editorial boards of *Science Education* and *Cognition & Instruction*, and regularly reviews for a number of education journals. He received his B.S. in Computer Science from the University of New Mexico in 1986; and his Ph.D. in Learning Sciences from Northwestern University in 1998.

**Stephen Silverman** is professor of education at **Teachers College, Columbia University** and also served as TC's Spencer RTG coordinator. His research focuses on student learning and attitude in physical education and the methods for doing research in field settings. He is the author of a number of books, including multiple editions of three research methods books (*Proposals that Work: A Guide for Planning Dissertations and Grant Proposals*; *Reading and Understanding Research*; and *Research Methods in Physical Activity*). Steve has served as editor of both the *Journal of Teaching in Physical Education* and the *Research Quarterly for Exercise and Sport* and was the first physical education scholar lecturer for the AERA SIG Research in Learning and Instruction in Physical Education. For over 25 years he has taught classes in research methods, measurement, and statistics.

**Crain Soudien** is a Professor in, and formerly the Director of, the School of Education at the **University of Cape Town**, and teaches in the fields of Sociology and History of Education. He has written extensively in the areas of race, culture, educational policy, comparative education, educational change, public history and popular culture. He is currently the convenor of doctoral studies at the School of Education, University of Cape Town and the research co-ordinator of

the South African Doctoral Consortium. He is also the co-editor of two books on District Six, Cape Town and another on comparative education and the author of *Youth Identity in Contemporary South Africa: Race, Culture and Schooling* and the co-author of *Inclusion and Exclusion in South African and Indian Schools*. Educated at the University of Cape Town, South Africa, he also holds a PhD from the State University of New York at Buffalo. He is involved in a number of local, national and international social and cultural organizations and has recently been elected as the President of the World Council of Comparative Education Societies.

**Rebecca M. Katz (Task Force research assistant)** is an advanced doctoral student at **Stanford University** completing a joint Ph.D. in Interdisciplinary Studies in Humanities and Philosophy of Education as well as a MA in philosophy. Her dissertation focuses on the moral and educational philosophy of John Dewey. Her wider research interests encompass liberal education, moral perfectionism, ethics, and literature (with a dash of German philosophy, namely Nietzsche and Hegel, for kicks). At Stanford she has co-designed and co-taught a course on “Higher Education and Society”, as well as taught numerous philosophy of education, ethics, and literature course sections. Prior to coming to Stanford she worked at Yale University for three years in academic administration and curriculum development. She has a MA from Teachers College, Columbia University and a BA in philosophy and English from Colgate University.

## Appendix 2: Annotated Table of Contents of background briefing package at outset of TF work

1. Shulman, Golde, Bueschel, and Garabedian, RECLAIMING EDUCATION'S DOCTORATES. *Educational Researcher*, April 2006.  
Good background on the problems facing "the doctorate", confusion of purposes, challenges when compared with doctoral training in other areas
2. Kaestle, THE AWFUL REPUTATION OF ED. RESEARCH. *Educational Researcher*, Jan-Feb. 1993.  
Almost a classic piece, based on in-depth interviews; title indicates the topic pursued
3. Tooley and Darby, OFSTED REPORT: EDUCATIONAL RESEARCH – A CRITIQUE. U.K.: *Office for Standards in Education (OFSTED)*, 1998.  
A controversial but fascinating evaluation of published research in the UK; documents many cases where conclusions outstrip evidence, work is unduly ideological, etc..
4. McClintock, EDUCATIONAL RESEARCH. *Teachers College Record*, March 2007.  
A sample from the tradition that finds little or nothing of value in empirical educational research. Should our students have a well-thought out response to such critiques?
5. Flyvbjerg, MAKING SOCIAL SCIENCE MATTER. *Cambridge University Press*, 2001.  
Short extract from a book of growing influence by a Danish social scientist; social science is off-track, but a change in emphasis can save it. Again, should our students be able to assess such lines of argument?
6. SAGE Handbook for Research in Education. *SAGE*, 2005.
  - a. Phillips, MUDDYING THE WATERS: THE MANY PURPOSES OF ED. INQUIRY
  - b. Floden, WHAT KNOWLEDGE USERS WANT
  - c. Lapsley, SECTION THREE: INTRODUCTORY ESSAY: Challenges in formulating and framing meaningful problems (extract).  
The aim of the handbook is to enrich inquiry and to foster engagement with meaningful problems. These chapters complexify the often simple accounts that are given in their respective domains – the titles are informative.
7. Koertge, SCIENTIFIC VALUES AND CIVIC VIRTUES *Oxford University Press*, 2005. (Chapter on "What science can offer democracy")  
A philosopher of science discusses scientific values, and their relation to "civic virtues"; throws some light on epistemological issues in a readable way
8. Kitcher, SCIENCE, TRUTH, AND DEMOCRACY. *Oxford University Press*, 2001. (Extracts)  
A noted philosopher of science discusses scientific and social relevance, and the democratic shaping of scientific research
9. Richardson, STEWARDS OF A FIELD. *Chapter 13 in Golde, Walker and associates*,

*ENVISIONING THE FUTURE OF DOCTORAL EDUCATION. Jossey-Bass, 2006.*

A prominent researcher on teaching reflects on issues in research training.

10. Schoenfeld, THE CORE, THE CANON, AND THE DEVELOPMENT OF RESEARCH SKILLS. *Chapter 9 in Lagemann and Shulman, ISSUES IN EDUCATION RESEARCH. Jossey-Bass, 1999.*

There are so many domains within education, it is hard to see any cross-cutting canon; this raises the specter of unhealthy compartmentalization, and students must be helped in adopting an appreciation for multiple perspectives

11. Heath, DISCIPLINE AND DISCIPLINES IN ED. RESEARCH. *Chapter 10 in Lagemann and Shulman, op. cit..*

A prominent anthropologist critiques superficial understanding of her field, and gives a feisty, biting assessment of the “punishing influence of psychology”. Relevant to our discussion of training in the two traditions

12. Eisenhart and DeHaan, DOCTORAL PREPARATION OF SCIENTIFICALLY BASED ED. RESEARCHERS. *Educational Researcher, May 2005.*

Two members of the writing team of the provocative NRC report reflect on the need for cross-disciplinary training of researchers

13. *Educational Researcher, June/July 2001: Theme Issue: Research for Doctoral Students in Education.*

- a. Young, BORDER CROSSINGS AND OTHER JOURNEYS....
- b. Metz, INTELLECTUAL BORDER CROSSING IN GRADUATE EDUCATION....
- c. Pallas, PREPARING ED. DOCTORAL STUDENTS FOR EPISTEMOLOGICAL DIVERSITY

A symposium organized by Spencer Foundation senior program officer Lauren Young; the titles give an accurate indication of the issues discussed.

14. Wisconsin Spencer Doctoral Research Program (DRP), website excerpts.

<http://www.wcer.wisc.edu/drpl/about.php>

An interesting example of what one of the Spencer RTG institutions has been doing. Does your institution have comparable programs/activities?

16. Joint Committee on Standards for Ed. Evaluation, THE PROGRAM EVALUATION STANDARDS. *SAGE, 2<sup>nd</sup>. Edition, 1994.*

Our final report needs to be relatively specific in its methodological, epistemological and other recommendations. What form should our report, or sections of it, take? The form these evaluation standards adopt, in an effort to avoid complete platitudes, is one interesting model.

















## **Appendix 4: “Intake Survey/Interview”, or guide to analysis of student competencies during the program**

The following is a guide that might be used to gauge where an Aspiring Researcher has already had opportunities for learning and where additional opportunities may be needed. *But we stress that it can be used at the point of intake, or at other stages in the program where the progress of the AR is being assessed.* The guide is organized around the four universes of knowledge and skills. This guide indicates specific areas of competency, with suggestions about documents that might be examined for evidence about competency and topics that might be raised in a conversation with the Aspiring Researcher. The particular documents to examine or questions to ask will vary according to the point which the Aspiring Researcher has reached in the program, but could be adapted for use at many different points, including during review of program applications, as part of an initial meeting between student and advisor, as a component of annual review of progress, or at program completion (where the purpose would be for recommending areas of continued learning). The program might also use the guide as a framework for program evaluation (e.g., exit surveys), for matching students and advisors, and for establishing and operating student research groups.

For each universe, questions are listed that might be answered by examining written materials and topics, and illustrative probes are given for conversations between a program faculty member (often the advisor) and the Aspiring Researcher.

### Universe 1

Questions that might be answered from written materials:

- What discipline(s) have comprised the focal areas of study for the AR
- To what extent was the AR exposed to research methods in that area?
- Has the AR studied of philosophy of science?
- Has the AR studied research methods?
- Does the AR have experience in conduct of research, as research assistant or research participant?

Topics for conversation, aimed at learning about competencies in Universe 1:

- Discussion of questions that particularly interest the AR, followed by probes about what research on the questions the AR knows about, what the AR knows about the methods used. Probing might be used to see what the AR thinks are strengths and weaknesses of studies, with the reasoning behind these judgments.
- Discussion of a social science field the AR has studied, probing for ideas about drawing generalizable conclusions, judging the strength of a chain of evidence, thinking about issues of objectivity and validity.

### Universe 2

Questions that might be answered from written materials:

- Has the AR studied political and social issues?
- Does the AR have practical experience in political and social action (e.g., working for an advocacy group or for a policy maker)?

- Does the AR have experience working in the education system (practitioner, administrator, policy maker)?

Topics for conversation, aimed at learning about competencies in Universe 2:

- Discussion of current education issues, asking about which are important and why.
- Discussion about how administrators or legislators establish policies, asking how the AR sees the roles of research, broad political issues, the views of local constituents and lobbyists.
- Discussion of why practitioners (teachers and administrators) use particular approaches and materials. Probes could address the social and organizational context; the roles of colleagues; fads, research, policies, personal preferences; the relevance of students' age and social backgrounds; and the influence of the backgrounds, training, and continuing professional development of the staff.
- Discussion about the ways in which ideas from research reports are interpreted by journalists, teachers, administrators, and legislative staff. Probes could address the strengths and weaknesses of academic writing styles, the roles of practitioner journals, and the match between practical needs and scholarly calendars.

### Universe 3

Questions that might be answered from written materials:

- Has the AR done concentrated study in social science or humanities discipline?
- Has the AR done concentrated study in relevant subject matter (e.g., mathematics for student in mathematics education)?
- Has the AR studied classroom processes, school organization, processes of learning, or policy implementation?

Topics for conversation, aimed at learning about competencies in Universe 3

- Discussion of recent experiences in a school or classroom (if any). Probes could explore what took the AR there, what was memorable about the events observed, whether anything seemed connected to what the AR has been studying?
- Discussion about what aspects of schooling the AR is most interested in studying. Aspects and related topics could include cognition and learning, testing and measurement, classroom processes, intercultural interaction, school organization, policy implementation, and policy development. Probes could explore the sources, and amounts, of the AR's knowledge of his or her chosen aspect(s) as well as those that are closely related.
- Discussion about the connections the AR sees between research interests and particular school subjects? Probes could explore what the AR knows about issues currently salient in particular subject matter areas (e.g., literacy).

### Universe 4

Questions that might be answered from written materials:

- Is the AR a member of one or more professional associations?
- Has the AR made presentations at conferences?
- Has the AR produced publications or reports? For what audiences?
- Does the AR regularly read professional journals? Which ones?
- Has the AR participated in the review of articles as part of the publication process?

- Has the AR had professional leadership experiences?

Topics for conversation, aimed at learning about competencies in Universe 4:

- Discussion with the AR about how he or she goes about identifying courses to take, workshops to attend, and the range of other learning opportunities to pursue.
- Discussion with AR about who to contact for professional guidance, starting with the faculty guidance committees, but later moving on to identifying both peers with related interests and possible career mentors or collaborators.
- Discussion with the AR about extending professional contacts beyond the local program and institution. Probing could address how the AR thinks about initiating new contacts, about the substance of interactions, about responding to contacts initiated by others.
- Discussion of readings done beyond those assigned in classes or workshops. Probes could address what ideas have seemed particularly important, how readings have contributed to the AR's own research plans.
- Discussion about the range of publication outlets in the field, including differences in topic, audience, reputation, and publication guidelines. Probes could address what the AR understands about publication processes, about how to adjust style for different audiences, and about the AR's current plans for submitting work for publication.
- Discussion of range of opportunities for speaking to non-academics. Probes could address what experiences the AR has had, what is understood about style of presentation, importance of preliminary discussions and follow-up. Probes could also address differences among groups, including students, parents, teachers, administrators, community leaders, and policy makers.
- Discussion of work with professional organizations, including service on committees within or outside of the program's institution. Probes could address the role of such service in professional life, how it figures into decisions about hiring and promotion.

## REFERENCES

- Blalock, H.M., and Blalock, A. (Eds.). (1968). *Methodology in Social Research*. New York: McGraw-Hill.
- Bruner, J. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University Press.
- Bunge, M. (1998). *Social Science under Debate*. Toronto: University of Toronto Press.
- Campbell, D., and Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In R. Lesh and A. Kelly (Eds.), *Handbook of Research Methodologies for Science and Mathematics Education*. Hillsdale, NJ: Lawrence Erlbaum.
- Cook, T., and Payne, M. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller and R. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research*. Washington, DC: Brookings.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Thousand Oaks, CA: Sage.
- Cronbach, L., and Associates. (1980). *Toward Reform of Program Evaluation*. San Francisco: Jossey-Bass.
- Dewey, J. (1938). *Logic: The Theory of Inquiry*. New York: Holt.
- Dewey, J. (1958). *Democracy and Education*. New York: Macmillan.
- Educational Researcher*. (2002). Theme issue on scientific research in education. 31(8), pp. 3-29.
- Educational Theory*. (2005). The education science question; A symposium. 55(3), pp. 235-322.
- Eisenhart, M., and Howe, K. (1992). Validity in educational research. In M. LeCompte, W. Millroy, and J. Preissle (Eds.), *The Handbook of Qualitative Research In Education*. San Diego, CA: Academic Press.
- Flyvbjerg, B. (2001). *Making Social Science Matter*. Cambridge, UK: Cambridge University Press.
- Gamoran, A. (2007). Reflections on exemplary dissertations in education research. Report to the Spencer Foundation, June 2007.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.

- Howe, K. (2003). *Closing Methodological Divides: Toward Democratic Educational Research*. Dordrecht, The Netherlands: Kluwer.
- Huck, S. (2000). *Reading Statistics and Research*. 3<sup>rd</sup>. Edition. New York: Longman.
- Jackson, P., and Kieslar, S. (1977). Fundamental research and education. *Educational Researcher*, 6(6), pp. 13-18.
- Kaplan, A. (1964). *The Conduct of Inquiry: Methodology for Behavioral Science*. Scranton, PA: Chandler.
- Kerlinger, F. (1977). The influence of research on education practice. *Educational Researcher*, 6(6), pp. 5-12.
- Kincaid, H. (1996). *Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research*. Cambridge, UK: Cambridge University Press.
- Lather, P. (2004). This IS your father's paradigm: Government intrusion and the case of qualitative research in education. *Qualitative Inquiry*, 10(1), pp. 15-34.
- LeCompte, M., Millroy, W., and Preissle, J. (Eds.). (1992). *The Handbook of Qualitative Research in Education*. San Diego, CA: Academic Press.
- Light, R., Singer, J., and Willett, J. (1990). *By Design: Planning Research on Higher Education*. Cambridge, MA: Harvard University Press.
- Lindblom, C., and Cohen, D. (1979). *Useable Knowledge: Social Science and Social Problem Solving*. New Haven, CT: Yale University Press.
- Locke, L.F., Silverman, S.J., and Spirduso, W.W. (2010). *Reading and Understanding Research*. 3<sup>rd</sup>. Edition. Thousand Oaks, CA: Sage.
- Maxwell, J. (2002). Understanding and validity in qualitative research. In A. M. Huberman and M. Miles (Eds.), *The Qualitative Researcher's Companion*. Thousand Oaks, CA: Sage.
- McDermott, R. (1993). The acquisition of a child by a learning disability. In S. Chaiklin and J. Lave, (Eds.), *Understanding Practice*. New York: Cambridge University Press.
- Michaels, D. (2005). Doubt is their product. *Scientific American*. 292(6), pp. 96-101.
- Miles, M., and Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. 2<sup>nd</sup>. Edition. Thousand Oaks, CA: Sage.
- Mosteller, F., and Boruch, R. (Eds.). (2002). *Evidence Matters*. Washington, DC: Brookings.
- National Research Council (NRC). (2002). *Scientific Research in Education*. Washington, DC: National Academies Press.

Parlett, M., and Hamilton, D. (1977). Evaluation as illumination. In D.Hamilton et al., (Eds.). *Beyond the Numbers Game*. Berkeley, CA: McCutchan.

Phillips, D.C. (2000). *The Expanded Social Scientist's Bestiary*. Lanham, MD: Rowman & Littlefield.

Phillips, D.C. (2005). Muddying the waters: The many purposes of educational inquiry. In C. Conrad and R. Serlin (Eds.). *SAGE Handbook for Research in Education: Engaging Ideas and Enriching Inquiry*. Thousand Oaks, CA: Sage.

Phillips, D.C. (2009). A Quixotic quest? Philosophical issues in assessing the quality and rigor of educational research. In P. Walters and A. Lareau (Eds.). *Educational Research on Trial: The Search for Rigor and the Promotion of Randomized Trials*. New York: Routledge.

*Qualitative Inquiry*. (2004). Symposium on scientific educational research. 10(1), pp.5-129

Siegel, H. (2006). Epistemological diversity and education research: Much ado about nothing much? *Educational Researcher*, 35(2), pp. 3-12.

Stokes, D. (1997). *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, DC: Brookings.

*Teachers College Record*. (2005). Symposium on the implications of the scientific research in education report for qualitative inquiry. 107(1), pp.1-58.

Thomson, G. (1963). Research in theory and practice. *The First Tizard Lecture, 1963*.  
<https://homepages.westminster.org.uk/sciences/firsttiz.htm>

Tooley, J., and Darby, D. (1998). *Educational Research: A Critique*. Report presented to OFSTED (Office for Standards in Education, UK).

Toulmin, S. (2003). *The Uses of Argument*. Updated Edition. Cambridge, UK: Cambridge University Press.

Webb, R., and Glesne, C. (1992). Teaching qualitative research. In M. LeCompte, W. Millroy and J. Preissle (Eds.). *The Handbook of Qualitative Research in Education*. San Diego, CA: Academic Press.