

Table of Levels of Research and Associated Research Strategies and Data

	LEVEL 1 →	LEVEL 2 →	LEVEL 3
RESEARCH STRATEGIES			
<p><b>1. Experimental.</b> An experimental research strategy usually involves <b>comparison</b> of two or more situations (classes, schools, districts) that are the same in every way possible, but that differ in the <b>few factors</b> (such as curriculum materials, teaching methods, leadership style) whose effects or outcomes you are <b>testing</b>.</p> <p>One group is the experimental group, that receives the intervention. The other is the control group (that receives the usual “treatment”) or an alternative treatment, that receives a different intervention that is being tested).</p> <p>Certain <b>design elements</b> make the data from experiments more <b>valid</b> (accurate), and therefore more believable and</p>	<p>May use <b>experimental designs</b>, but (since this is <b>basic research</b> or a <b>pilot study</b>) the experiment may not have equivalent control and comparison groups and standardized measurement.</p> <p>For instance, a supervisor has a hunch (from earlier observations or from several research articles that she read), that teachers may benefit more from directive supervision, with explicit instruction from the supervisor.</p> <p><b>She does pilot research to see if there is anything to this hunch.</b> She provides highly directive supervision to five new teachers, and <b>nondirective</b> supervision (teachers evaluate themselves and suggest ways to improve their teaching) to five other teachers---to see if there is <b>any consistent different</b> in these teachers’ improvement.</p> <p>If the researcher DOES find the expected consistent differences, she next uses Level 2 research to test more <b>rigorously</b> (i.e., using design</p>	<p>Let’s say that the <b>researcher finds from Level 1 research</b> that teachers who received more directive supervision improved instruction more than teachers who received <b>nondirective</b> supervision.</p> <p>The researcher turns this finding into an <b>hypothesis</b> to test more rigorously.</p> <p>“Teachers who receive more directive supervision will improve their instruction and will be more satisfied with supervision than teachers who receive nondirective supervision.” [Notice that several outcome variables are measured, to get a bigger picture of the possible effects of the</p> <p>The researcher: a. Develops a <b>conceptual definition</b> of “proficient instruction.” Then she transforms this definition into an <b>operational definition</b>---specific features of proficient instruction; e.g., using clear words, using a proper set of examples, correcting errors.</p>	<p>Let’s say that the researcher finds from Level 2 research that the data support the hypothesis.</p> <p>“Sure enough, teachers who received more directive supervision DID improve their instruction and WERE more satisfied than teachers who received nondirective supervision.”</p> <p>Great!</p> <p>But is one study of 20 teachers in two suburban schools enough to be <b>confident</b> that there would be the <b>same results with different samples</b> of teachers (older, male, different grades); in different schools (urban, rural) and with different supervisors? No.</p> <p>Is it ethical for the researcher to say that all schools ought to supervise in a directive way? No.</p> <p>The researcher can’t be sure</p>

<p>useable. These features are:</p> <p>a. <b>Samples are representative</b> of the population for which the findings are relevant.</p> <p>b. <b>Randomization or matching</b> are used to produce equivalent groups.</p> <p>c. Variables to measures are <b>clearly defined</b>.</p> <p>d. A <b>wide range of variables</b> are measured so that a comprehensive picture is obtained (e.g., achievement, the quality of a curriculum, the quality of instruction, and student and teacher satisfaction are all measured).</p> <p>e. The same variables are measured in several different ways (<b>triangulation</b>).</p> <p>f. Measures (e.g., how to count behavior) and instruments (for tests) are <b>validated</b>.</p> <p>g. Measurement (e.g., testing) is checked for tester or observer</p>	<p>elements cited in the left-hand column) her hypothesis (not merely a hunch) on the effects of directive supervision</p>	<p>b. Develops a <b>conceptual definition</b> of “teacher satisfaction with supervision.” Then she transforms this definition into an <b>operational definition</b>---specific things that teachers might do and say, that represent different degrees of satisfaction with supervision.</p> <p>c. Develops <b>instruments</b> (See <b>DATA</b> in the left-hand column.) for <b>rating</b> teachers’ “proficient instruction” and satisfaction with supervision.”</p> <p>d. Tests the <b>reliability</b> of the instruments to see if different observers score the same teachers’ proficiency and satisfaction the same way.</p> <p>e. Develops a <b>protocol</b> that tells exactly how to deliver directive vs. nondirective supervision</p> <p>f. Trains some supervisors in two suburban schools to supervise in a directive way and others to supervise in a nondirective way.</p> <p>g. Randomly assigns <b>20 teachers</b> to the two supervision groups.</p> <p>h. Pre-tests (rates) teachers’ proficiency at instruction.</p>	<p>that the findings in the first study WEREN’T a chance fluke (that is, wouldn’t happen again if the study could be repeated exactly).</p> <p>And the researcher can’t be sure that the findings apply anywhere else.</p> <p>Therefore, the researcher conducts <b>larger-scale evaluation research</b>. For instance, the researcher:</p> <p>a. Does the study again with samples that are similar to the first ones, to see if the results of the first study were a fluke .</p> <p>[Let’s say there are the same results. Probably not a fluke.]</p> <p>b. Does the study across the whole district. There are 40 elementary schools.</p> <p>The schools are <b>randomly assigned</b> (by flip of a coin) to the two supervision groups.</p> <p>A <b>random sample</b> of 10 teachers is selected from each school.</p> <p>Observers are <b>trained</b> to use the instruments for measuring</p>
--	--	--	--

<p><b>reliability.</b></p> <p>h. Both <b>pre-tests</b> and <b>post-tests</b> are given. Even better, data are collected periodically <b>during an intervention</b> to see if there is any trend; e.g., gradual vs. rapid change.</p> <p>i. The experiment is <b>replicated</b> (done again) with <b>similar samples</b> (to see if the findings are consistent and strong), and with <b>different samples</b> (to see if the findings apply elsewhere).</p>		<p>i. Implements the two kinds of supervision.</p> <p>j. Post-tests (rates) teachers’</p> <p>k. Uses <b>structured questionnaires</b> or <b>interviews</b> to find out how satisfied teachers were with their kind of supervision.</p> <p>l. Analyzes the data to see if the data support the hypothesis; that is, do the data say that teachers who received directive supervision improved more (and were more satisfied with supervision) than teachers who received nondirective supervision?</p> <p>For example,</p> <p>What is the average amount of change in skill for each feature of proficient instruction in teachers who received directive vs. nondirective supervision?</p> <p>What percentage of teachers who received directive supervision vs. nondirective supervision rated themselves as “highly satisfied”?</p>	<p>proficient instruction and satisfaction with supervision. The <b>reliability</b> of their scorings is checked.</p> <p>Supervisors in the schools are trained to use the protocols for providing directive vs. nondirective supervision.</p> <p>Teachers receive <b>pretests</b> of their teaching skills; then teachers receive one of the two kinds of supervision; and then teachers receive post-tests of their teaching skill.</p> <p>Teachers are given structured questionnaires or interviews to find out how satisfied teachers were with their kind of supervision.</p> <p>The researcher analyzes the data. For example,</p> <p>What is the average amount of change in skill for each feature of proficient instruction in teachers who received directive vs. nondirective supervision?</p> <p>What percentage of teachers who received directive supervision vs. nondirective</p>
---	--	---	---

			<p>supervision rated themselves as “highly satisfied”?</p> <p>The researcher finds that, again, directive supervision yields significantly more improvement in teaching skill and significantly higher ratings of satisfaction.</p> <p>However, the researcher now <b>disaggregates</b> the data.</p> <p>She divides the data on all teachers into groups whose characteristics may be relevant: male/female; veteran/new teachers; ex-military/non military.</p> <p>Does the teacher’s sex, time as a teacher, or military experience influence the effectiveness of the two forms of supervision?</p> <p>The researcher finds that in general, directive supervision <b>STILL</b> works best---whether male or female, new or veteran, military experience or not.</p> <p>However, male teachers and ex-military (males and females) receiving directive supervision <b>improve more</b></p>
--	--	--	--

			<p>than female teachers and nonmilitary males and females receiving directive supervision.</p> <p>The researcher now has good reason to feel fairly confident of the findings, and might implement directive supervision across the district. The researcher also knows that if there is ever a choice, it is a safer best to use directive supervision with men and with ex-military men and women.</p>
<p><b>2. Survey.</b> Survey means <b>overview</b>. The survey research strategy is used to obtain a larger picture.</p> <p>Various kinds of <b>survey data</b> may be collected, but some data are more difficult to obtain than others. For example, a researcher might use <b>official statistics</b> to survey teacher retention or students achievement across a state.</p> <p>Or, a researcher could survey teachers' proficiency by observing and scoring (on rating</p>	<p>Based on research he has read and observations he has made, a researcher has a hunch that there are different styles of school (principal) leadership, and that these styles are associated with different degrees of teachers' job satisfaction.</p> <p>This hunch is certainly not enough even to <b>suggest</b> that principals use a certain leadership style. Basic research is needed to see if there is anything TO the hunch.</p> <p>The researcher distributes <b>structured questionnaires</b> (a series of questions that all teachers are asked) to six schools (two elementary, middle, and high schools) in a district to see <b>if there</b></p>	<p>Let's say that the <b>researcher finds from Level 1 research</b> that there is an association (correlation) between teachers' evaluations of school leadership and teachers' job satisfaction. For instance, teachers were most satisfied when principals had the following set of traits:</p> <ul style="list-style-type: none"> <li>a. Establishes a clear and feasible school mission that organizes all aspects of the school around proficient instruction.</li> <li>b. Is well-read in current scientific research. Shares this with staff.</li> <li>c. Skillfully evaluates and helps teachers to select effective curriculum materials.</li> </ul>	

scales) the teaching skills of a large sample of teachers in a district. However, this would be a lot of work for one person.

**is any association** (correlation) between teachers' evaluations of school leadership and teachers' job satisfaction).

d. Skillfully models and provides useful suggestions for effective instruction methods when teachers need help.

e. Provides useful resources for assessment, instruction, and remediation---workshops, materials.

f. Listens and finds useful points in differing opinions, but makes and takes responsibility for final decisions about curriculum, instruction, supervision, etc.

g. Visits all teachers in class almost every day, and praises teachers in front of class.

However, the sample was not large, and was not representative of all schools. Therefore, the researcher decides to test whether the findings "hold up" in other settings.

The researcher turns the findings into an hypothesis to test more rigorously.

"Teachers will be most satisfied when principals have a leadership style that involves the following features:....(described above)."

<p><b>3. Field (ethnographic) research.</b>  Field research is done in naturalistic settings (e.g., classrooms). The point is to obtain a big picture of behavior, interactions, activities, social roles, physical environment—similar to what anthropologists do.</p> <p>Data are usually more qualitative, such as informal interviews and field notes. However, the researcher certainly can count (structured observations) what he or she thinks are important events (such as how often a teacher calls on diverse learners). This research can yield findings that a researcher might want to investigate in a more formal way.</p>	<p>A school district wants to find out why diverse learners are doing so poorly. They have low achievement and a high drop out rate. Does anything in early elementary grade classrooms predict this?</p> <p>The researcher observes a sample of 12 classes—two each in grades k-5 in a school. The classes have at least five diverse learners in them. The observer takes field notes on teacher-student interaction. The observer notices: (a) where different students sit; (2) how often the teacher gives different students opportunities to respond; (3) how often the teacher gives praise and other forms of recognition and acceptance; (4) how patient the teacher seems to be, as judged by facial expression and tone of voice; (5) how the teacher responds to off-task or disruptive behavior.</p> <p>The researcher analyzes the field notes and finds that there are modest associations between teacher-student interaction and student achievement. For instance, the majority of diverse learners with higher achievement were in classes where teachers made sure they sat up front; gave lots of opportunities to respond; were patient; gave</p>		

	frequent praise; and responded to off-task and disruptive behavior in a matter-of-fact way.		
<b>DATA</b>			
1. <b>Tests with quantitative data</b> ; e.g., percentage correct; percentile rank. Collected by the researcher or from (for example) a district's official statistics	<p>For example, the researcher uses a standardized test of phonemic awareness (can students distinguish the separate sounds in words) to see if there is any association between the degree of phonemic awareness and students later speed at learning to sound out (decode) words.</p> <p>A researcher obtains <b>official statistics</b> from a district or across a state to see if there is any association between schools using certain math materials and student achievement.</p>		
2. <b>Structured observations</b> (e.g., in class) <b>with quantitative data</b> ; e.g., (a) <b>counting</b> specified behaviors or interactions; or (b) <b>rating scales</b> of skill.	<p>Using published research on features of effective reading instruction, the researcher develops an instrument (tentative, because it has not been tested) that enables the researcher to rate teachers on a four-point scale on a variety of important instructional behaviors.</p> <p>For example, "The teacher corrects:  (1) almost every student error; (2) about two-thirds of errors; (3) about one-third of errors; (4) almost no errors.  Teachers are then rated before and</p>		

	<p>at the end of a week-long refresher course on the identified features of effective instruction.</p> <p>The researcher examines the data to see which (if any) instructional behaviors have improved.</p>		
<p><b>3. Field notes; qualitative narratives of events</b></p>	<p>A researcher observes 20 different peer reading tutoring partnerships four times during a month of school. The tutors naturally “teach” in somewhat different ways. The researcher uses narrative recording (field notes) to describe the interactions. The researcher examines the data (field notes) to see if there are: (1) certain styles of tutoring (e.g., highly directive); (2) if there are style differences depending on students’ age and sex; and (3) if there is any association between style of tutoring and later student achievement.</p>		
<p><b>4. Informal interviews; qualitative information</b></p>	<p>The researcher (above) interviews peer partners to see if there is any association between: (1) style of tutoring; (2) tutors’ perception of effectiveness; (3) tutees’ satisfaction and perception of effectiveness of tutoring.</p>		
<p><b>5. Structured interviews and questionnaires with quantitative questions (e.g., rating scales) and perhaps qualitative questions (e.g., “Tell me about...”)</b></p>	<p>A sample of students and the families of students who dropped out of school are interviewed. Interviewers ask a series of questions, such as:  “How strongly do you agree or disagree with this statement: ‘If I</p>		

	<p>had been taught to read better, I probably would not have dropped out of school.”?”</p> <ul style="list-style-type: none"><li>(a) Strongly agree.</li><li>(b) Agree.</li><li>(c) Disagree.</li><li>(d) Strongly disagree.</li></ul> <p>The researcher analyzes the responses to see if there is any association between the responses and dropping out. For example, is there a lot of agreement on the preventive effects of proficient reading?</p>		
--	--	--	--