Observational Methods*

David R. Cross, Ph.D.†
Texas Christian University‡

September 13, 2007

Introduction

Observation is the one essential ingredient of science. One of the hallmarks of the scientific revolution during the 15th, 16th, and 17th centuries was the emphasis placed on empirical observation by such men as Kepler, Galileo, Hook, and Vesalius. Arguments based on empirical observation replaced arguments based on theological dogma and/or Greek philosophy. Without observation, all science, including psychology, would be an empty shell. Although the phenomena being observed will be different for each science, the skills of observation are a key ingredient of scientific training in any discipline. Indeed, it is my observation that the greatest scientists have very often been the keenest observers of nature.

Observation is also an essential ingredient of parenting. Good parents are keen observers of their children's behavior and actions. In this sense good parents are like good scientists, and for parents and scientists alike observation is the key to insight and understanding. As you will discover elsewhere in this course, good parents are sensitive to their children's thoughts and feelings, responsive to their actions, and closely monitor their behavior and progress towards desired goals. All of these characteristics of effective parenting depend fundamentally on observation. Although there is more to parenting than keen observation—knowledge, skills, and resources are also important—keen observational skills are a cornerstone of good parenting.

An explicit goal of this course is to improve your observational skills (see the eCollege Syllabus). These skills will be important to you both as a parent and as a professional. You will have several opportunities to learn and practice various observational methods, and this handout is designed as an introduction to the various methods for scientific observation.

*Prepared for *Child Psychology* (PSYC 30323)
†www.davidcross.us
‡www.tcu.edu

Observational Methods

There are four parts to this brief overview of observational methods:

1. *Use of Observational Techniques* includes a discussion of the place of observation in the overall scientific enterprise, and a set of basic question to ask before embarking on observational adventure.

2. *Types of Observations* provides an overview of the different types of observation used in child development research; this section also indicates how each of the observational activities fits into the proposed typology.

3. *Estimating Interobserver Reliability* addresses an essential feature of scientific observation, which is the reliability of the observations; this section includes a simple formula for computing inter-rater reliability for many of the activities you will be doing in this course.

4. *Sources of Observational Bias* includes an overview of the basic sources of bias that can creep into the observational process.

Use of Observational Techniques

A common misconception among psychologists is that observational methods are suitable for field studies, but not for laboratory research. This couldn't be further from the truth, for it is becoming increasingly common—at least in developmental psychology—for observational methods to be used in laboratory as well as field settings. There are numerous well-known studies where observational techniques were used in laboratory settings, and this is not the place to survey this empirical literature. Nevertheless, it might be useful to briefly discuss one well-known example.

Albert Bandura conducted a series of studies, the so-called "Bobo-doll" studies, during the 1960s that are classics of the developmental literature (Bandura et al., 1961, 1963a,b). In these studies Bandura sought to determine whether viewing an aggressive model on television would increase children's aggressiveness. In a typical experiment, children were randomly assigned to one of two groups: A control group watched a program with little or no violence, whereas an experimental group watched a program with relatively high levels of violence. After watching the film, the children in both groups were allowed to play freely in a laboratory room with lots of toys, including a "Bobo" doll. (In the violent film, models were shown hitting the "Bobo" doll.) As might be expected, the children in the experimental condition (violent content) group were more aggressive during the free play session: They hit the Bobo doll far more often than their counterparts in the control condition.

Two points are worth making for present purposes. First, observational techniques were used to measure the dependent variable (aggression). Second, this was a laboratory study with a true experimental design (i.e., there was random assignment of participants to experimental conditions). Clearly, observational

techniques can be, and have been, used in laboratory situations. The converse is also true: Measurement techniques other than observation (e.g., tests, tasks) can be, and have been, used in field settings. Thus, there is no one-to-one correspondence between measurement techniques and research settings. (The myth that observational methods are only used in field settings has been surprisingly persistent in psychology textbooks, which is why I dwell on it here.)

The main point so far is that observational techniques can be suitably (and fruitfully) employed in almost any kind of psychological study (see Bakeman & Gottman, 1987; Pelligrini, 1996; Zaslow et al., 2006). But there is another important point to be made, which is also relevant to this course. This is the point made above about observational techniques and skills not only being important for researchers, but for clinicians, teachers, and parents also. It is for this reason that observational methods are emphasized in *Child Psychology*. Anyone planning to use observational methods must begin by answering four general questions. The answers to these questions may vary dramatically depending upon the observational method to be used, and more importantly, on the questions to be answered.

What will be observed? This apparently simple question may have a complex answer. You will wish to observe certain behaviors of the child (and/or the child's parent), but exactly which ones and precisely how to define them may not be straightforward. For example, if one is interested in observing attachment behaviors in infants, the list of specific behaviors could be quite long. These attachment-specific behaviors could include approach, visual pursuit, clinging, smiling, and crying, to name just a few. It is also necessary to define the behaviors of interest in manner that makes them easy for others to identify. The behavior "vocalizations directed to the mother" may require detailed explanation when observers actually try separate these behaviors from other infant vocalizations occurring in the mother's presence. In addition to observing the behaviors of the child, the investigator may wish to record other events as well. The behavior of others in the child's environment as well as the entire range of physical and social events occurring during the period of observation, are also frequently data of interest. Clearly, many decisions of this type must be made before observations can be made.

When will the observations be made? Related to this question is the issue of how long or how many observations will be made. This answer again depends upon the nature of the question or research topic, but there is a very important issue here that is common to observational studies. Because the observations will occur for only a limited period of time (even if the period is very long), it is crucial that the obtained data be *representative* of the target construct or behaviors. For example, a single one-hour observation of the child in the home may provide a very distorted sample of the child's typical behavior, especially if it happens to occur during an unusual circumstance, such as during illness or holiday. Likewise, observations made during a particular part of the day may not accurately reflect the total story, as with children who are attentive during morning hours but get restless and disruptive later in the day. And if there is any suspicion that the observational procedures are initially reactive

(see *Sources of Observational Bias* below), it may be necessary to continue the procedure until such effects stabilize or disappear.

Which observational method will be used? There are many possible answers to this question, and there is really just one. The investigator should use the method that most efficiently and accurately provides the information of interest. The various methods will not be discussed at this point, because they will be examined below. But it is important to note that the researcher should be aware of the different types of methods available, including their strengths and weaknesses, so the most appropriate method may be selected.

How will the accuracy of the observations be verified? In many ways, this is the most important question of all, because if the researcher cannot be confident that the observations represent what actually occurred, the data are of no scientific or practical value. The major problem that can arise here is that the observer may not accurately record the behaviors or events being observed. This could happen for a variety of reasons, such as insufficient observational skills, prior expectations of what will occur, poorly defined behaviors, and so on. The usual safeguard for dealing with these possibilities is the use of several observers, each recording the behaviors independently. The observations are then compared, and similarity or *reliability* is determined mathematically. If they are sufficiently similar, the researcher typically assumes that they are also reasonably accurate. Methods for computing inter-observer reliability will be discussed below.

The four general questions just discussed are only the minimum number that must be considered before observations are made. Many other specific questions could be discussed as well, including the nature and training of the observers, the use of videotape or audiotape equipment to provide a permanent record of the observed events, the openness of the observational procedures versus the use of hidden or unobtrusive observers, and so on. In each case, the answers to these questions should be considered in advance, keeping in mind the unique characteristics or requirements of the research problem. It is always a good idea for novices in a particular research area to consult with experienced researchers, or at least to consult published investigations on related topics.

Types of Observations

There are many different types of procedures that have been used to observe children, but most of them can be grouped into one of three groups: *Informal Observations*, *Indirect Observations*, and *Formal Observations*.

Informal Observations

Informal Observations are the least rigorous of the three types that are considered here. Although not rigorous, informal observational procedures have played an important role in the development of child psychology. These procedures are considered informal because there is typically no systematic plan for timing the

observations, nor is there any attempt to verify the reliability of the observations. As a result, the observations have limited scientific value (although, of course, informal observations can be extremely valuable to a parent, teacher, or clinician). But informal observations *can* be a valuable source of new insights and hypotheses about child development, as we shall see.

The earliest of the informal observations are the now-famous *baby biographies*, some of which were produced during the eighteenth century. Among the most famous of these reports is Charles Darwin's description of the behavior and development of his son "Doddy," published in 1877 (Darwin, 1877). Darwin recorded many detailed characteristics of the infant, including initial reflex reactions, early evidence of fear or anger, and even more sophisticated behavior sequences presumed to display reasoning or moral development. More recently, under the supervision of psychologist Joseph Church, three mothers kept detailed diaries of their newborns during the first year of life. Church gave the mothers some broad guidelines for their observations, but he generally suggested that they record "... whatever they found amusing, surprising, or puzzling." These biographies were then revised and published in 1966 as a book entitled *Three Babies* (Church, 1966).

Perhaps the most famous informal observations in the annals of developmental psychology are the observations made by Jean Piaget on his three children when they were infants and toddlers. These observations are the empirical foundation for Piaget's developmental theory of sensorimotor intelligence. Incredibly, these informal observations of Piaget's own three children were the main source of information for his theorizing about the development of thought during the first two years of life. Although developmental psychologists are revising the details of Piaget's theory, it remains true as a broad outline of early cognitive development. Piaget's ideas about the early growth of intelligence were published in 1952 (in French) as *The Origins of Intelligence in Children* (Piaget, 1952).

Indirect Observations

Sometimes it is not possible to directly observe children's behavior, or perhaps it is inefficient to do so. In these cases researchers may resort to *indirect observations*, where the information of interest is obtained from sources other than the child.

The most common method of indirect observation involves interviewing adults who know the target children well enough to provide meaningful information. Parents and teachers are the most common sources, but ministers, coaches, and counselors, among others, could be sources also. (Peer nomination procedures are another type of indirect observation.) The actual interview procedures can be quite controlled and structured, or the adult may be permitted to elaborate on particular situations in a less-structured manner. A well-known example of a highly structured method of indirect observation is Thomas Achenbach's *Child Behavior Checklist* (Achenbach, 1991, 1992).

Another method of indirect observation involves interviewing an adult or an

adolescent about their own childhood experiences. Although these self-reports, or *retrospective reports*, are a valuable approach to gathering developmental information, there are obvious limitations. First, we remember virtually nothing about our own infancy, so retrospective reports are useless as sources of information about this important stage of development.[1] Second, indirect observation itself requires a certain amount of skill and knowledge, and children are not always reliable sources of information about their own experiences. An example of this is the difficulties associated with the use of children as "eyewitnesses" to their own physical or sexual abuse. The upshot of this limitation is that it is usually not practical to interview people about their own development until adolescence or adulthood, but by then they may have forgotten much of their own developmental history! Despite problems, self-reports remain an attractive method for gathering information about human development, for in some sense no one knows about any one of us as we do ourselves. One good example of a self-report method is Mary Main's *Adult Attachment Interview*, which we will be discussing in this course (see Hesse, 1999).

There is a trade-off with indirect observations. On the one hand, the source—parent, teacher, the "children" themselves—are highly knowledgeable sources information. In fact, these people are "gold mines" of information. On the other hand, the information obtained from these sources can be unreliable, incomplete, and/or biased. To continue with the "gold mine" analogy, the "ore" obtained, while plentiful, can be of low grade, and must be carefully processed to obtain "pure gold". Thus, when using indirect observations, great care must be taken to ensure the accuracy of the information obtained. The *Child Behavior Checklist* "works" because the items are simple and objective, requiring little subjective interpretation. The *Adult Attachment Interview* "works" because of the great care and effort taken during the interview and with the scoring, and because of Mary Main's extraordinary insights about subconscious attachment processes.

Formal Observations

The third group of observational methods are those that include systematic observing and recording procedures. Such methods not only consider the time and duration of the observation period, but also the precise behaviors to be observed and the procedures used to record them. In addition, reliability estimations are typically provided by multiple, independent observers. Three principle methods of formal observation have been used by psychologists (see Vasta, 1979, ch. 10, for further examples): *Rating Scales, Frequency of Occurrence, Time-Sampling.*

When using *rating scales*, observers are required to score the subject on a number of dimensions, or continuous scales, for which the end-points (and perhaps the midpoints) are carefully defined. Such rating scales typically use

[1] Freud referred to this phenomenon as "infantile amnesia", but his explanation—that we repress the sexual urges of our infancy as a defense mechanism—has been discredited; Piaget's explanation in terms of differences between sensorimotor and representational intelligence is closer to the mark (see the *Infant Play* activity).

five or seven values, and are sometimes called *Likert* scales after a prominent psychometrician who first described the method (see Likert, 1932). A good example of the use of rating scales is with the *strange situation*, developed by Mary Ainsworth and her colleagues (Ainsworth et al., 1978). (You will have an opportunity to score strange situation tapes as part of the *Attachment Activity*.) Central to scoring the strange situation tapes are the *interactive rating scales: proximity seeking, contact maintaining, avoidance,* and *resistance*. Each scale is defined by seven anchors, and the anchors themselves are richly described. I know of no better rating scales than those used in the strange situation.

Another method of formal observation involves counting the *frequency of occurrence* of the behaviors of interest during the observation period. The crucial feature of this procedure is that these behaviors, called *target behaviors*, must be defined with enough detail so that observers working independently can reliably identify their occurrence or nonoccurrence. The target behaviors in Bandura's "Bobo doll" studies (see above) were aggressive behaviors such as striking or kicking the Bobo doll. A variant of the *frequency of occurrence* approach was used by Belsky and Most when investigating their infant play scale (Belsky & Most, 1981). You will have an opportunity to score infant play tapes using this scale as part of the *Infant Play Activity*.

A third group of formal observational methods are called *interval-sampling* or *time-sampling* methods. The idea behind this approach is that it is much easier for observers to agree on the occurrence or nonoccurrence of a behavior during a short period than during a longer one. Observation intervals, therefore, are divided into very small units, often ten seconds or less. At the end of each interval, observers record whether the target behaviors have occurred and then immediately begin the next interval. Twenty minutes of such observation could produce over one hundred intervals on which independent observers could compare their recordings. Again, target behaviors typically are well-defined to increased the likelihood of accurate observations. An additional benefit of time-sampling approaches is that they preserve the temporal structure of the observations, a benefit that can pay off if one wishes to look at sequences of behaviors. Time-sampling methods are the bedrock of research on early parent-infant interactions, as seen in the work of several contemporary developmental psychologists (see Papousek, 2007; Trevarthen, 2005). You will have an opportunity to score the social play of toddlers using a time-sampling scheme, in the *Social Play* activity.

Estimating Inter-Observer Reliability

When collecting data for reliability purposes, it is necessary to have two observers independently observe the same subject at the same time. These multiple observations do not have to cover the entire observation period, however. For example, if a child is being observed each day for sixty minutes by the principal observer, the second observer (reliability assessor) commonly observes that child for only ten to fifteen minutes of that time. In some research, reliability observations are not even conducted during every observation session (or tape),

although it is usually considered desirable for the reliability to cover the entire time span of the research.

The data from the reliability assessor are then compared with that same portion of the principal observer's data. Two decisions must be made at this point. The first concerns the *unit* of reliability assessment. For example, it is usually the case that several behaviors are coded from an observation episode (e.g., smile, vocalize, look away, etc.). The methodological question is, should reliability be assessed for each behavior individually, or for the set of behaviors? The basic rule of thumb is this: Compute reliability at the level you are going to report your results. So, if you are going to present frequencies (or some other statistic) for the individual behaviors, you should compute reliability at the level also.[2]

The second decision concerns the *time frame* of the reliability estimate. For example, the degree of inter-observer reliability could be calculated for each session, for a group of sessions, or simply for all the sessions combined. Again, the general rule is that reliability computations should not be combined over time frames that are to be discussed or reported separately. For example, in the Social Play activity, one of the research questions is whether there is significant change in the children's social play across the two weeks of the summer play camp. In this case reliability should be assessed for each week separately, and not for the entire two weeks. If it were desirable to analyze the data by day, then the reliability should be assessed within that same time frame.

Although the computation of inter-observer reliabilities has gotten more complex than it used to be twenty or thirty years ago,[3] the procedures we will use in class are quite simple. Although simple, they capture the essential elements of inter-observer reliability. The basic idea is this:

$$Interobserver\ Reliability = \frac{Agreements}{Agreements + Disagreements} \times 100.$$

This formula yields percent agreement, which is satisfactory for our purposes. Eighty percent agreement is usually considered the minimum for adequate inter-observer reliability, and values over ninety percent are considered excellent.[4]

[2] Actually, my recommendation is that you try to achieve adequate inter-observer reliability at a *finer* level of analysis than you will be reporting. For example, in attachment research, reliability is usually assessed at the level of the subclassifications (A_1, A_2, ...) for the strange situation, even though the data will be reported at the level of the major classifications (*A, B, C, D*). By striving for adequate reliability at a finer level than will be reported, you stand a much better chance of having adequate inter-observer at the level you actually want to report.

[3] For example, the procedures described in Vasta (1979) are out of date.

[4] As an aside, note that modern procedures are based on Cohen's *kappa* (Cohen, 1968), which is similar to percent agreement, but with an additional twist; the additional twist is that *kappa* corrects for *chance agreement*, which can make an important difference. Imagine the following two scenarios: In both scenarios two observers agree at the 60% level (not so good, but good enough for this example). However, in one scenario, there are only two behavioral codes, so the observers might be expected to agree 50% of the time, just by chance. In the other scenario there are ten codes, so the observers might be expected to only 10% of the time by chance, so that 60% agreement looks pretty good. Cohen's *kappa* statistic takes these differences into account, which is one reason why it is the standard approach for researchers today.

Please note that in this class you will never be graded on your level of reliability (percent agreement) with you scoring partner. In this class you simply do not receive enough training on any one scoring system to achieve proficiency, and so it would be unrealistic for us to expect highly reliable scorers. For our purposes, it is enough for you to realize that it takes a lot of work for two observers to independently score children's behavior and get to the point where each observer records the same thing.

Sources of Observational Bias

The term *observational bias* refers to the possibility that observational data—even data arising from two observers who are highly reliable—might be inaccurate or consistently distorted. Below I discuss several of the most important sources of observational bias, with possible remedies.

The first source of bias involves the subjects of the observations and is termed *reactivity*. This means simply that the behavior of the subjects sometimes changes when they know they are being observed. Two general solutions have been proposed to deal with this problem. Sometimes the reactive effects of being observed decrease after a period of time (this is an example of *habituation*). Therefore, one solution involves simply continuing the observations over a long enough period of time for the behaviors to stabilize—presumably at "normal" levels. Unfortunately, it is often difficult to know when and if this has occurred, making this solution less than certain. A second alternative is to employ unobtrusive or covert means of observation. Hidden cameras or one-way viewing mirrors are frequently used in observational research so that subjects are not aware (or are less aware) of the data collection. This approach is more likely to reduce or eliminate reactivity, but is not always a practical solution in many settings.

A second source of bias involves the observers themselves. We have already discussed the value of objective procedures and reliability assessment as means of reducing observer distortion. But even well-defined and well-quantified procedures can result in biased observations under some circumstances. For example, some kinds of observational data seem to be particularly sensitive to the *expectations* of the observers. Scott Miller, in his book *Developmental Research Methods* (Miller, 1998), gives the following example of observer bias:

> A study by Kent, O'Leary, Diament, and Dietz (1974) provides an example. The observers in this study viewed the videotapes that supposedly showed the baseline and treatment phases of a program intended to reduce disruptive behavior in a classroom setting. Half of the observers were told that a decrease in disruptive behavior from baseline to treatment was predicted; half were told that no change was predicted. In fact, all the observers viewed the same videotape, in which no change occurred for any of the behavioral categories. When later asked to make a global rating of the effectiveness of the program, however, 9 of the 10 observers who had been led to

> expect a decrease in disruptive behavior reported that a decrease had in fact occurred. In contrast, 7 of the 10 observers led to expect no change reported no change. It is interesting to not that there were *no* differences between the two groups in the actual behavior recordings made while watching the tape. Only on the overall global rating did the expectancy manipulation have an effect. (pp. 88–89 of the 1st edition, published 1987)

The fact that the expectancy manipulation only affected the global ratings illustrates how objective observational procedures *may* protect the data from observational bias. A problem related to observer expectancies is the problem of the feedback given to observers by the research team. If the observers are led to believe that the experimenter is either pleased or disappointed with their findings, there is some tendency for observers to produce data that are consistent with the *experimenter's* expectancies. Both of these problems can be solved if the observers are kept "blind" (unaware of the experimental conditions and the experimenter's expectations). Again, this is not always practical.

Another problem involving the observers is *knowledge of reliability assessment*. It has been shown that if an observer is aware that a second observer is assessing reliability, there is a greater tendency for their observations to be similar (and hence reliable) than if reliability is assessed unobtrusively. This effect probably results from the observers being more attentive or alert in their observations when they know they are being checked. The best way to solve this problem is to keep observers unaware of when reliability assessments are occurring. This is relatively easy to do when scoring is done off of videotapes, but can be quite difficult if observations are done "live."

Another source of bias is related to the observational procedures themselves. One such problem involves *code complexity*. If observers are required to use a scoring system involving many different types of behaviors or events, lower reliability estimates may occur. This can be partially eliminated by better training procedures, but it also indicates that are limits to human observer capabilities. Another possible bias involves your selection of procedures for *reliability estimation*. As was suggested in the previous section, there are many ways to calculate inter-observer reliability, with some methods producing artificially high estimates of agreement (e.g., the simple procedure we use in *Child Psychology* does not correct for chance, so will in some cases yield artificially high—biased—levels of agreement). In selecting a reliability procedure, therefore, the investigator should take precautions against using a method that will imply greater observer accuracy than actually occurred.

In this section I have indicated various approaches to observing children's behavior, and have warned against several possible pitfalls when making such observations. It may seem to you at this point that the dangers outweigh the benefits, so why would anyone want to employ observational methods at all? The answer lies in the following quote taken from Scott Miller's text:[5]

[5]Miller, in turn, quotes Yarrow and Waxler (1979), Observing interaction: A confrontation

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M. (1992). *Manual for the child behavior checklist/2-3 and 1992 profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Ainsworth, M. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Oxford, England: Lawrence Erlbaum Associates.

Bakeman, R. & Gottman, J. M. (1987). Applying observational methods: A systematic view. In J. D. Osofsky (Ed.), *Handbook of Infant Development* (2nd ed.). (pp. 818–854). Oxford, England: John Wiley & Sons.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology*, *63*(3), 575–582.

Bandura, A., Ross, D., & Ross, S. A. (1963a). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology*, *66*(1), 3–11.

Bandura, A., Ross, D., & Ross, S. A. (1963b). Vicarious reinforcement and imitative learning. *The Journal of Abnormal and Social Psychology*, *67*(6), 601–607.

Belsky, J. & Most, R. K. (1981). From exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology*, *17*, 630–639.

Church, J. (1966). *Three babies; biographies of cognitive development*. New York: Random House.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220.

Darwin, C. (1877). A biographical sketch of an infant. *Mind*, *2*, 285–294.

Hesse, E. (1999). The adult attachment interview: Historical and current perspectives. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment* (pp. 395–433). New York: Guilford Press.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 55.

Miller, S. A. (1998). *Developmental research methods* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

---

with methodology. In R. B. Cairns (ed.), *The analysis of social interactions: Methods, issues, illustrations* (pp. 37–65).

Papousek, M. (2007). Communication in early infancy: An arena of intersubjective learning. *Infant Behavior and Development*, *30*(2), 258–266.

Pelligrini, A. D. (1996). *Observing children in their natural worlds: A methodological primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Trevarthen, C. (2005). First things first: Infants make good use of the sympathetic rhythm of imitation, without reason or language. *Journal of Child Psychotherapy*, *31*(1), 91–113.

Vasta, R. (1979). *Studying children: An introduction to research methods*. San Francisco, CA: Freeman.

Zaslow, M. J., Weinfield, N. S., Gallagher, M., Hair, E. C., Ogawa, J. R., Egeland, B., Tabors, P. O., & De Temple, J. M. (2006). Longitudinal prediction of child outcomes from differing measures in a low-income sample. *Developmental Psychology*, *42*(1), 27–37.