

# Applied Research

As we saw in Chapter 7, one general way to distinguish quantitative and qualitative research is in terms of the goals that motivate the research. In a somewhat different sense, a difference in goals also defines the distinction to which we turn now.

Chapter 1 began with a discussion of why we do research in developmental psychology. As I argued there, we do research for two general reasons. We do research for reasons of *basic science*, to advance our understanding of human development; and we do research for purposes of *application*, to better the lives of children and other vulnerable groups. These two goals are by no means incompatible, nor do the basic methodological principles that govern good research change when we move from a basic to an applied focus. Nevertheless, applied research can present special challenges that add to the challenges and complexities of research in general. Because most of the book deals with the basic research, the purpose of the present chapter is to add some points with respect to research in applied developmental psychology.

What is meant by “applied research”? Here are several definitions. Fisher and Lerner (1994, p. 4) define applied developmental psychology as “that aspect of psychology which bears upon enhancing developmental processes and preventing developmental handicaps.” Lerner, Jacobs, and Wertlieb (2003, p. 4) offer a similar definition: “the programmatic synthesis of research and

applications to describe, explain, intervene, and provide preventative and enhancing uses of knowledge about human development.” Finally, Lerner (2010, p. 38) writes that “applied developmental science involves the integration of basic issues of description and explanation with issues about how developmental science may be applied . . . to improve the human condition.” In short, applied research attempts to make life better for the objects of its study.

A number of specific goals and related forms of research fall within the scope of applied developmental psychology. Table 8.1 summarizes these goals and gives examples of the kinds of research or writing that relate to each; the table is derived in part from Fisher and Lerner (1994) and in part from other discussions of the applied approach (e.g., Zigler & Finn-Stevenson, 1999).

Clearly, all of the goals identified in Table 8.1 are important. In what follows I limit the discussion to the first three entries in the table: assessment of at-risk populations, research directed to socially important issues, and interventions designed either to prevent or to correct problems. Fuller discussions of both these and other kinds of applied work can be found in Durkin and Schaffer (2016); Fisher, Busch-Rossnagel, Jopp, and Brown (2013); Lerner et al. (2003), Maholmes and Lomonaco (2010), and Sigel and Renninger (2006). Good sources with respect to the translation of research into public policy

Table 8.1  
*Forms of Applied Research in Developmental Psychology*

Category	Goal	Examples
Assessment	To identify developmentally important characteristics in an at-risk target population	Use of a neonatal screening test to measure developmental status in newborns subjected to prenatal stress. Use of the Stanford-Binet IQ test to identify possible cognitive delays in preschoolers from impoverished homes.
Intervention	To alter the environment in ways that may prevent, correct, or reduce problems of development	Project Head Start or similar programs intended to boost the academic readiness and cognitive and social competence of preschoolers from disadvantaged backgrounds. Memory training programs for older adults experiencing memory difficulties.
Research on socially important issues	To provide evidence relevant to the resolution of pragmatically important questions	Research directed to the accuracy of children's testimony in cases of suspected abuse. Research directed to the determinants of teenage substance abuse.
Contribution to public policy	To use knowledge gained from research to inform policy-making decisions	Use of findings from developmental research to establish standards of quality for outside-the-home child care. Incorporation of findings from developmental research in court decisions regarding custody and visitation in cases of divorce.
Dissemination of psychological knowledge	To make available the results of research to those who might benefit from the knowledge	Production of books or pamphlets of advice for teenage parents. Consultation by psychologists in situations (e.g., school board decisions, court cases, design of nursing homes) in which their knowledge is relevant.

include Aber, Bishop-Josef, Jones, McLearn, and Phillips (2007); Bogenschneider and Corbett (2010); and Huston (2008). Good sources with respect to dissemination of findings from applied research include McKoy (2015), Shonkoff and Bates (2011), and Welch-Ross and Fasig (2007). Note also that there are several journals and newsletters devoted specifically to applied issues. Table 8.2 lists the most relevant such sources.

## Assessment

Issues of assessment can arise at any point in the life span. Our example concerns a very early

phase of life: the newborn period. The question is how to identify babies who are at risk—at risk at the least for developmental problems and at the extreme for failure to survive.

What kinds of experiences might place a newborn at risk? There are, unfortunately, many. Being born prematurely constitutes a risk, as does being malnourished when in the womb. Maternal illness during pregnancy can pose a threat. So can the use of obstetrical medication during delivery. And so can the mother's ingestion of nonfood substances while pregnant, such as alcohol and drugs.

In the United States and many other countries, babies undergo their first assessment within

Table 8.2

*Sources for Publications in Applied Developmental Psychology*

---

*Analyses of Social Issues and Public Policy*  
*Applied Developmental Science**The Future of Children**Journal of Applied Developmental Psychology**Journal of Applied Gerontology**Prevention Science**Psychological Science Agenda**Psychological Science in the Public Interest*

---

*SRCD Social Policy Reports*

---

a minute or so of birth. The Apgar test is given in the delivery room to assess the immediate status of the newborn (Apgar, 1953). Five dimensions are measured, each on a 2-point scale: heart rate, respiration, muscle tone, color or complexion, and reflex irritability. A total score of 6 or below is considered problematic; scores of 3 or less indicate that immediate intervention is necessary to ensure survival.

Although the Apgar performs a vital screening function in the minutes following birth, it provides no information about the more long-term functioning of the newborn. For that, we need a measure that can be administered across the early weeks of life and that assesses newborn functioning more generally. By far the leading such instrument is the **Neonatal Behavioral Assessment Scale** (NBAS), first developed over 40 years ago and now in its fourth edition (Brazelton & Nugent, 2011).

The NBAS was not the first attempt to assess newborn status. Earlier tests, however, relied heavily on measures of reflexive behaviors as an index of the baby's neurological integrity. An absence of certain reflexes, for example, can be a sign of brain damage, as can the failure of a reflex to drop out at the expected time in development. The NBAS retains a focus on reflexive behavior;

a total of 20 reflexes are elicited at various points in the exam. But it adds to them a behavioral component that goes well beyond simple reflexes—28 behavioral items that tap individual differences in varied forms of newborn adaptive competence.

Table 8.3 presents a subset of the behavioral items. It can be seen that the scale includes various measures of the *habituation* response described briefly in Chapter 6 and discussed more fully in Chapter 12. It also includes several measures of the infant's reaction to the social stimulation presented by the examiner, such as orientation to her face or voice and the ability to be soothed by the adult when distressed. The emphasis on the baby not simply in isolation but in interaction with others is, in fact, one of the distinctive characteristics of the NBAS. Another characteristic is the emphasis on eliciting the baby's *optimal* level of performance—on administering the exam in a flexible enough manner to reveal the very best that the newborn can do.

The NBAS illustrates many features that are true of assessment instruments more generally. It encompasses a range of content; because our target is newborn competence in general, it is important that the assessment not be limited to one or a few types of behavior. There is an emphasis in administering the test on *standardization*—on both giving and scoring the test in the same way for all infants. The reason for standardization is obvious. We can interpret any newborn's performance, as well as draw comparisons among babies, only if we can assume that all babies have been tested in the same way. On the other hand, there is also an emphasis, as I noted in the preceding paragraph, on *flexibility*—on adjusting the preset procedures when necessary to elicit the baby's optimal response. Most assessment instruments are designed to allow for some on-the-spot adjustment, within the bounds of standardization, as the tester deems necessary; they vary, however, in where they fall on the standardized-to-flexible continuum. The NBAS is definitely toward the flexible end. The need for flexibility, in turn, reflects a basic fact about infant assessment, and that is that the baby is in charge,

Table 8.3

*Examples of Items From the Neonatal Behavioral Assessment Scale*

Item	Description
Response to decrement to light	While infant is asleep, shine light in eyes and observe response; after response disappears, wait 5 seconds and re-present; continue for either 10 trials or until habituation occurs.
Response to decrement to rattle	While infant is asleep, shake rattle near ear and observe response; continue for either 10 trials or until habituation occurs.
Inanimate visual orientation	Slowly move a red ball across the infant's field of vision; record ability to track both horizontally and vertically.
Animate visual orientation	Have examiner slowly move his or her face across the infant's field of vision; record ability to track both horizontally and vertically.
Animate auditory orientation	While out of the infant's line of sight, have examiner speak softly into baby's ear; record ability to localize on each side.
Defensive movement	With infant in supine position, hold cloth over eyes; observe for about 15 seconds, noting rooting, head turning, neck stretching, directed or nondirected swipes.

*Note.* Adapted from *Neonatal Behavioral Assessment Scale* (4th ed.) by T. B. Brazelton and J. K. Nugent, 2011. London: Mac Keith Press. Copyright 2011 by Mac Keith Press. Adapted with permission.

and babies, as discussed more fully in Chapter 12, are not always cooperative research participants. The challenges of infant testing are multiplied with the NBAS, given the young and immature nature of the target group, and they may be multiplied even further if the particular newborn being tested is experiencing problems. And this leads to a final point about the NBAS, which is that the exam can be administered only by a highly trained tester. All assessment instruments require some in-advance training and some on-the-spot clinical judgment; the NBAS, however, ranks especially high on both dimensions.

Does the NBAS provide a valid assessment of newborn status? Chapter 4 discussed the general issue of test validity. Validation of the NBAS relies to some extent on each of the kinds of evidence reviewed there; the primary validation, however, is in the form of *criterion validity*—the ability of the measure to predict to some external indication of developmental status. Criterion validity,

in turn, divides into concurrent validity (relation to contemporary status) and predictive validity (relation to future status). One index of concurrent validity is the test's ability to identify immediate problems in babies subjected to adverse circumstances during prenatal development or at birth. Thus, the NBAS verifies that each of the conditions sketched earlier (e.g., prematurity, drug exposure) does in fact place the infant at greater than average risk for developmental problems. It is important to note, however, that problems are not inevitable in such cases; conversely, problems may arise—and be identifiable with the NBAS—even in the absence of any known history of risk.

Establishment of predictive validity requires longitudinal study, in which we determine whether scores on the NBAS predict to measures of development later in infancy or in childhood. A large research literature demonstrates that early status can be predictive of a variety of

aspects of later development, generally within the span of infancy but to some extent beyond as well (Lester & Sparrow, 2010; Nugent, 2013; Nugent, Petrauskas, & Brazelton, 2009). The predictive power is, to be sure, far from perfect, but then this is hardly surprising: No one expects a measure given in the first days of life to predict everything about later development. What is striking is that it is predictive at all.

Several further points about assessment follow from this validation evidence for the NBAS. When we use an assessment instrument, we are at least interested in determining the current status of the individual being assessed. When adults are the target for the assessment, our interest may not extend beyond this information about contemporary functioning. With assessments of children, however, the goal is often to predict not just current but *future* functioning. How, for example, do certain problems during the newborn period relate to social development at age 1? As would be expected, the accuracy of prediction decreases as the time span between assessment and outcome measure increases. Furthermore, even if we confine ourselves to one time period, there is no assessment device, at least within the psychological realm, that provides a complete, error-free measure of the construct of interest. Assessments are always imperfect to varying degrees, and the test user must be aware of the limitations when making diagnoses or prognoses that impact people's lives.

A final point is that the work with the NBAS illustrates the two general benefits that accurate assessments, appropriately used, can have. One is to the individual assessed. An assessment can provide information about the individual—often, but by no means always, concerning difficulties in development—that can help us to structure the future environment in ways that will reduce problems and optimize the developmental course. The second possible benefit is to the target population more generally. If, for example, a measure like the NBAS reveals that some prenatal practice place infants at risk, then educational efforts and policy changes can be implemented to reduce the risk for future generations.

## Socially Important Issues

What topics fall under the heading of socially important issues? Table 8.4 provides a sampling of topics that are frequent foci for work in applied developmental psychology. The list is a partial one; for other examples simply look through the tables of contents of the publications listed in Table 8.2.

### Children as Witnesses

The first example I consider is both one of the most important and one of the most intensely studied of the topics listed in the table: children as witnesses in suspected cases of child sexual abuse. Research directed to this issue falls under the heading of **forensic developmental psychology**: the study of children's behavior and understanding in legal contexts (Holliday & Marche, 2013; Malloy, Lamb, & Katz, 2011).

Some statistics can help to convey the importance of the issue. It has been estimated that at least 100,000 children testify in court cases in the United States every year (Court Appointed Special Advocates for Children, 2012). This figure does not include the much larger number of instances in which children provide evidence outside of court. The cases in which children testify span a range of topics, but the most frequent category among criminal trials is child sexual abuse. In most instances of alleged abuse, the child witness is also the target of the abuse. In many, the child is the only witness.

Can the testimony of a young child be trusted? Should such testimony be admissible in court? As the figures just cited indicate, questions like these can be critically important to answer.

They can also be very difficult to answer. We have not yet considered how researchers typically study children's memory—that is one of the topics discussed in Chapter 13. I will tell you now, however, something you can no doubt guess, and that is that the typical study of children's memory bears little resemblance to the memory-for-abuse situation. Two general differences exist.

Table 8.4  
*Examples of Topics of Study in Applied Developmental Psychology*

Adolescent pregnancy	Literacy
Aggression and violence Bullying	Marital disruption and divorce
Children's eyewitness reports	Mass media and computers
Depression	Obesity
Developmental psychopathology	Parenting and parent education
Domestic violence and maltreatment	Pediatric psychology
Early child care	Poverty
Early childhood education	Preterm birth
Education reform and schooling	Prevention science
Elder care	Successful children and families
Immigrant families	War trauma

The first difference concerns the content for the memory. Memory studies in the basic-science literature focus on children's memory for benign content, either arbitrary material such as list of words or personal experiences of a pleasant sort such as a family trip. In contrast, the memory at issue in cases of abuse is memory for personal experience of a very negative, often highly traumatic sort. Most basic memory studies ask about memory for one-time experiences or events; abuse may sometimes be a one-time occurrence, but it also may occur on multiple occasions across an extended period of time. In much basic memory research the child is a passive recipient of the memory materials; instances of abuse may involve the child as a participant and not merely as an onlooker or bystander. Finally, in basic memory research there is no reason for children not to report everything they remember; this, however, is not necessarily true in cases of abuse. Complex social and emotional factors may affect children's reports, such as the child's guilt about being a participant or reluctance to implicate a parent or friend. Thus, what

children say may not be identical to what they remember.

The second general difference concerns the method of questioning. In basic memory research the questioning is generally designed to be optimal—that is, to elicit the most accurate recall of which the child is capable. To this end, researchers use clear and straightforward questions, provide appropriate encouragement and social support, and in general do the things that research has shown help children to perform at their best. Accurate recall is certainly the goal in most real-life forensic interviews, but the questioning may be far from optimal. To begin with, children may be questioned by a range of people—parents, teachers, doctors, lawyers, police officers—none of whom is necessarily expert in the best way to question children. The sheer number of questioners introduces a potentially daunting factor not found in most laboratory research, as does the number of separate occasions on which questioning may occur—estimates are 12 occasions on the average, and the figure is much higher in some instances. In addition, the questioning in

real-life cases is almost always delayed to some extent, and in some instances it may be quite delayed; children may be questioned several years after the alleged incidents took place. Finally, the questioning that children encounter in legal cases is often far from straightforward; rather it may include various suggestions and prompts and reinforcers that lead the child in the direction of particular answers. Such shaping of the child's responses is especially likely if questioners believe that they already know the truth or if, as advocates within the court system, they have an interest in a particular outcome. The following examples from court transcripts give some idea of just how leading the questioning can be.

Prosecutor: Did you have to lay on top of Bridget?

Bobby: Yes.

Prosecutor: And when you were laying on top of Bridget, where was your private?

Bobby: I forgot.

Prosecutor: Do you remember telling Miss Judy that you had to put your private next to her private? Did you have to do that, Bobby?

Bobby: No Sir.

Prosecutor: What did you say?

Bobby: No Sir.

Prosecutor: Did you say No or Yes?

Bobby: Yes Sir.

Prosecutor: Did she touch you with a spoon?

Child: No.

Prosecutor: No? Okay. Did you like it when she touched you with the spoon?

Child: No.

Prosecutor: No? Why not?

Child: I don't know.

Prosecutor: You don't know?

Child: No.

Prosecutor: What did you say to Kelly when she touched you?

Child: I don't like that. (Ceci & Bruck, 1998, p. 741)

How might research speak to all these complexities? One challenge, clearly, is to discover or devise situations we can ask about that bear some similarity to the abuse situation. Two general approaches have been tried. One possibility is to create experimental settings that reproduce some elements of the real-life situations of interest. One set of investigators, for example, had the child play a Simon Says game with the experimenter, in the course of which both the child and the experimenter touched parts of each other's bodies (White, Leichtman, & Ceci, 1997). In another study the child interacted with an adult dressed in a clown suit; in the course of their play parts of the suit were put on and taken off both participants, and the adult took photos of the child (Rudy & Goodman, 1991).

Although studies such as these capture many aspects of real-life cases (e.g., removing clothing, taking photos), they do have an obvious limitation, which is that they omit a central element of the abuse situation—namely, the trauma of being abused. Obviously, researchers cannot intentionally subject children to traumatic experiences. They can, however, capitalize on instances in which such experiences have already occurred, and the second type of study falls in this category. Among the examples of painful or embarrassing experiences that children have been questioned about are medical emergencies (Peterson, 2015), receiving an injection (Chae et al., 2014), receiving a genital and anal examination (Eisen, Goodman, Qin, Davis, & Crayton, 2007), undergoing urinary catheterization (Quas et al., 1999), undergoing a painful medical procedure (E. Chen, Zeltzer, Craske, & Katz, 2000), and experiencing a hurricane (Fivush, Sales,

Goldberg, Bahrick, & Parker, 2004). Although such experiences can hardly equal the trauma of abuse, they do capture some of its characteristics.

As noted, the second general challenge in this research is to simulate the types of questioning that suspected victims of abuse must undergo. This, too, has been the focus of a number of studies. A child may be questioned several times across a period of weeks, for example, or questioned by different people on different occasions. Children may be told to “keep a secret” about what happened to them during the experimental session (Bottoms, Goodman, Schwartz-Kenney, Sachsenmaier, & Thomas, 1990) or a police officer rather than a research assistant may do the questioning (Melinder & Alexander, 2010). Finally, probably the most thoroughly explored variable is the nature of the questioning, with a basic contrast between straightforward questions (e.g., “Tell me what happened.”) and suggestive or leading questions (e.g., “He kissed you, didn’t he?” “How many times did he kiss you?”). As we saw, suggestive questioning is a part of many real-life cases; it is therefore important to determine how susceptible children are to an adult’s suggestions.

Although methods rather than findings are our concern here, I will note a few general conclusions that emerge from this very active research literature (for further discussion, see Bottoms, Najdowski, & Goodman, 2009; Bruck, Ceci, & Principe, 2006; Saywitz & Camparo, 2014). First, research verifies what we would expect: that memory improves with age and that older children typically report more of their experiences than do younger children. Second, as the delay between an event and questioning about it increases, the completeness and accuracy of recall decline, and this is especially true for young children. Third, in at least some cases, young children are more suggestible than are older children or adults—that is, they are more likely to be influenced by leading questions from an adult authority figure. On the other hand—and more positively—in many studies memory differences between younger and older children

or between children and adults are not very great. Age differences, as well as memory inaccuracies in general, are most likely when specific questions are used; conversely, they are minimized by the use of free recall measures that allow children to say what happened in their own words. Finally, the memory problems that children do show are mainly errors of omission rather than of commission—that is, they are more likely to fail to report certain details than they are to introduce false information. This finding suggests that any clearly spontaneous mentions of abuse by children should be taken very seriously.

Having offered these conclusions, I will add that there remains much controversy about exactly what the research shows and what the implications are for children’s legal testimony. Among the topics under active investigation is the issue of how best to question children in order to meet two goals: maximizing the accuracy of testimony and minimizing stress to the child (D. A. Brown & Lamb, 2015; Brubacher, Poole, & Dickinson, 2015; Lamb, Herskowitz, Orbach, & Esplin, 2008). Having to testify can add to the trauma of an already traumatic experience. It is therefore important to devise procedures that protect the child from further harm.

The studies of children as witnesses illustrate several general points about the relation between basic research and applied research. First, the starting point for applied work on some topic is the relevant basic-science literature—in the present case, the literally thousands of studies of children’s memory that emerged from the first century of psychology as a science. Such studies produced numerous findings (e.g., differences between recognition and recall, effects of delay, effects of retrieval cues) that can help us to understand the applied situation. Second, the basic-science research is seldom a sufficient basis for applied efforts; rather, the complexities of real-world problems may require kinds of study that go beyond what has been attempted before. This, of course, is a point I stressed in introducing the research on children as witnesses. Finally, the relation between basic and applied is a two-way



**BOX 8.1 RESEARCH WITH MINORITY POPULATIONS**

The demographics of the United States population have changed markedly across recent decades. In the most recent census, members of minority groups (i.e., any group other than non-Hispanic Whites) made up 38% of the population, more than double the percentage 30 years ago. Projections indicate that by 2045 minorities will no longer be a minority, for they will constitute more than 50% of the population. This is already the case in several states (Arizona, California, Hawaii, New Mexico, and Texas).

The representation of minorities in research falls well short of their representation in the population (Knight, Roosa, & Umana-Taylor, 2009). Recent years, it is true, have seen some increases in both the frequency and the quality of relevant research, in part because of several special issues of developmental journals devoted to research with minorities (Cabrera, Beeghly, & Eisenberg, 2012; McLoyd, 1990; Quintana et al., 2006), as well as the launching of the journal *Cultural Diversity and Ethnic Minority Psychology*. Research on the development of immigrant children has been especially active in recent years (e.g., Crosnoe & Fuligni, 2012; Garcia Coll & Marks, 2012). Nevertheless, large gaps remain, and researchers who attempt to fill these gaps face formidable challenges. It is to these challenges that this box is devoted.

As always, an initial issue is that of sampling. Sampling from any minority group is a two-step process. Researchers must first decide exactly whom they wish to study, for considerable diversity exists within any group. Hispanics, for example, represent many countries of origin, and they vary substantially in income level, proficiency in English, and other indices of adaptation to the majority culture. The second step, of course, is to secure the cooperation of those whom one has decided to target. Recruitment of participants can be difficult in any kind of research, but the difficulties are often compounded in research with minorities. Members of minority groups often do not share the cultural membership of the researchers who approach them, and they may be less likely to believe in the value of research than are more mainstream research participants. Indeed, some may be actively suspicious and defensive with regard to researchers who come into their community, a position for which, sadly, there is a good deal of historical support (Knight et al., 2009).

Numerous researchers have addressed the challenges of recruiting minority samples. Interestingly, some of the most helpful discussions have come from researchers in disciplines other than psychology—anthropology (e.g., Spicer, 2010), for example, and health research (e.g., Yancey, Ortega, & Kumanyika, 2005). A common feature that cuts across the various discussions is the need to treat the participants in such research as partners at every step of the research enterprise and not merely as subjects to whom research is done. The precepts that have emerged from this collective recruitment effort are summarized in eight “values” delineated by two anthropologists (Jacklin & Kinoshameg, 2008): partnership, empowerment, community control, mutual benefit, holism, action, communication, and respect.

A second general issue is measurement. A point made in Chapter 6 in the discussion of cross-cultural research applies as well to much of the research that has been carried out with minority groups within the United States. Often, the starting point for such research has been some measure or set of measures developed primarily for White middle-class populations, measures that are then administered unchanged to some minority sample. Often, the outcome has been poorer

performance by the new sample, a result that has nurtured “deficit” lines of theorizing with regard to the perceived problems of minorities. As we have seen, however, use of the same measure with different groups does not guarantee that it is functionally the same—equally fair, equally comprehensive, equally informative. Good discussions of both measurement equivalence in general and measurement equivalence in research with minorities in particular are provided by Knight and Zerr (2010) and Knight et al. (2009).

A final general issue concerns design. Group status—whether the group be age, sex, race, or ethnicity—is a subject variable, and as such, it is subject to all the problems of interpretation discussed under Subject Variables in Chapter 2. Any two groups that we might compare in such research are likely to differ in numerous ways (e.g., income level, language proficiency, geographical residence, degree of acculturation) in addition to racial or ethnic status. Drawing cause-and-effect conclusions about group differences is therefore very difficult.

One more point is important to make. Members of minority groups face many obstacles in adapting to the majority culture, and applied research can help both to identify and to remove those obstacles. Placing research with minorities under the Applied heading does not mean, however, that such groups are of interest only because they present problems that must be solved (Quintana et al., 2006). One consistent message from those who work with such populations has been the need for more research directed to the normative development of children from different minority groups.

one: As findings emerge from the new paradigms of applied research, they feed back on and expand our basic-science understanding—in the present case, what we know about children’s memory.

## Bullying

Our next topic is a relatively recent focus for research and policy, even though the behavior at issue has probably been around ever since peers first began to interact with one another. **Bullying** is a particular form of aggression that has two features not found in most instances of aggressive behavior. One is that there is an imbalance of power between the participants: The bully has power; the victim lacks it. The other is that bullying is (at least usually) not a onetime occurrence; rather it involves repeated acts of aggression. In most cases bullies bully multiple times, victims are victimized multiple times, and the same pairing of bully and victim occurs multiple times.

As always, a first question for any topic is that of measurement. Before we can determine how often bullying occurs or what effects it has, we must be able to measure it. Although observational measures sometimes appear in the bullying literature, the majority of studies rely on verbal reports. And of these, the great majority rely on self-reports, as opposed to reports from parents, teachers, or peers (Hymel & Swearer, 2015). Thus what we can learn about bullying depends on how successful we are at the difficult task of getting children and adolescents to report their experiences accurately. Table 8.5 shows a sampling of items from one of the many self-report instruments that contribute to this literature.

Specifying exactly how often bullying occurs is a challenge, given that findings vary not only across samples but also across methods of measurement. Two recent reviews (Hymel & Swearer, 2015; Juvonen & Graham, 2014) report ranges, summed across studies, of 4% to 13% among school-aged children and adolescents for engaging in bullying and 9% to 33% for being

Table 8.5  
*Examples of Items From the Illinois Bullying Scale*

For each of the following questions, choose how many times you did this activity or how many times these things happened to you in the LAST 30 DAYS.					
	Never	1 or 2 times	3 or 4 times	5 or 6 times	7 or more times
I upset other students for the fun of it.					
In a group I teased other students.					
I fought students I could easily beat.					
Other students picked on me.					
Other students made fun of me.					
Other students called me names.					
I got hit and pushed by other students.					
I helped harass other students.					
I got in a physical fight.					
I threatened to hurt or hit another student.					

*Note.* Adapted from "Bullying and Victimization During Early Adolescence: Peer Influences and Psychosocial Correlates," by D. Espelage & M. K. Holt, 2001, *Journal of Emotional Abuse*, 2, p. 130. Copyright 2001 by Hayworth Press. Reprinted with permission.

the victim of bullying. Many more children, of course, experience bullying not as participants but as spectators.

Much of the research on bullying has been directed to the characteristics of bullies and victims. What features predispose children to either role, and what are the consequences of being either a bully or a victim? Although there is no single answer that fits every case, some average patterns do emerge (R. C. Cook, Williams, Guerra, Kim, & Sadek, 2010; Kijakovic & Hunt, 2016; McDougall & Vaillancourt, 2015). Bullies, perhaps not surprisingly, tend to be low in empathy and high in need for control; they often hold distorted views about aggression; they are especially susceptible to peer influence; and

they often come from a troubled home background. Across time they are at heightened risk for depression, school disengagement, and criminality in adulthood. The class of victims includes a small subset labeled *provocative victims*; these children are themselves combative and aggressive, and they tend to respond to bullying with aggression, although usually unsuccessfully. The characteristics of most victims, however, fit a different pattern that is again not surprising. Such children tend to be physically weak, they are often socially isolated, they manifest internalizing difficulties (e.g., anxiety and depression), they often have fearful temperaments and a history of over-protective parenting, and they tend to give in to and therefore reinforce aggression. Over time

their internalizing problems worsen, their relations with peers are typically poor, their academic performance suffers, and they are at heightened risk for psychiatric disorders in adulthood.

The research just discussed is necessarily correlational, and as such it cannot establish cause and effect with certainty. Consider the relation between internalizing problems and victimization. Even if we assume a causal relation, the directionality question still remains: Are children bullied because they possess these qualities, or do they develop problems because they are bullied? Longitudinal research can help to pull apart these possibilities, and it suggests that the causality runs in both directions: Such problems are often a precursor to bullying, but a history of being bullied also makes the initial problems even worse (Reijntjes, Kamphuis, Prinzie, & Telch, 2010). Note that this conclusion of a back-and-forth causal relation is an example of the transactional model of development discussed in Chapter 3.

I turn next to a relatively new form of bullying. We saw in Chapter 6 that recent years have seen the emergence of an important new context for development: namely, experience while on the Internet. We saw also that not all such experiences are positive, for time on the Internet brings risks as well as rewards. One of the most important examples of this point is a form of bullying made possible by the Internet, the phenomenon of **cyberbullying**. Cyberbullying is bullying that occurs not in face-to-face interactions but rather via communications in the digital world of the net. Texts, chats, e-mails, instant messages, blogs, Facebook, Twitter, Instagram—all can be and have been used to inflict pain on others (Bauman, Cross, & Walker, 2013; Vollink, Dehue, & McGunkin, 2016).

Cyberbullying shares many features with traditional bullying. Most basically, cyberbullying is an aggressive act, it is often repeated, and it often involves an imbalance of power between perpetrator and recipient. In addition, cyberbullying, like traditional bullying, can clearly lead to negative outcomes in its victims, including in extreme cases suicide. Indeed, some well

publicized instances of suicide as an apparent result of cyberbullying clearly helped spark the current interest in the topic. Finally, traditional bullying and cyberbullying are moderately correlated; that is, children who are bullies or victims in one setting tend (although with exceptions) to be bullies or victims in the other setting as well (Kowalski, Giumetti, Schroeder, & Lattanner, 2014).

There are also some important differences between traditional bullying and cyberbullying. Unlike many instances of traditional bullying, cyberbullying does not inflict physical harm on the victim (although it may include the threat of doing so). The aggression, rather, is verbal—damage produced not by what is done but by what is said. One of the factors that makes cyberbullying so worrisome is the broad potential scope for such aggression. A partial list of ways to hurt others digitally includes flaming, harassing, impersonating, excluding, spreading rumors, and posting embarrassing pictures (Willard, 2007). Other characteristics of cyberbullying also heighten its frequency and potential damage, including its anonymity, its freedom from restrictions of time and space, the absence of adult control, the absence of typical social cues from the victim, and its potentially large audience (Mehari, Farrell, & Le, 2014).

As with bullying in general, the frequency of cyberbullying varies across samples and across methods of study. One recent summary concludes that approximately 20% of teens have been the victims of cyberbullying at some point in life, and approximately 15% have engaged in cyberbullying themselves (Hinduja & Patchin, 2015; but see Olweus, 2012, for a cautionary view). Such surveys, it is interesting to note, reveal another difference between traditional bullying and cyberbullying. Traditional bullying shows a consistent gender difference: Boys engage in it more often than do girls. In contrast, girls are just as likely to be involved in cyberbullying, as either perpetrators or victims, as are boys.

The discussion to this point has not addressed an obvious question: What can be done to reduce

the prevalence of bullying and to minimize its damaging effects? The question of how to address and correct important social problems is the subject of the next section of the chapter, and we will consequently continue with the topic of bullying there.

## Intervention

The goal of intervention research is to identify methods to correct or prevent problems in some developmentally vulnerable population. There have been thousands of such efforts, with populations ranging from newborns to octogenarians and interventions ranging from a few minutes to several years. All, of course, are intended to provide benefit to their recipients. But they are also intended to generate knowledge that can be applied more broadly to similar populations in the future—which is what makes them research and not just intervention. In what follows we consider two examples of this point.

## Bullying

Intervention programs have been a part of bullying research since the pioneering work by Dan Olweus in Norwegian schools in the 1980s (Olweus, 1993). Because it set the mold for many subsequent intervention efforts, I begin with a brief summary of the Olweus Bullying Prevention Program (OBPP).

Four general goals underlie the Olweus approach. Each reflects the fact that the approach is not limited to known bullies and victims; rather, it is a school-wide endeavor that targets all of the parties relevant to the control of bullying. One goal is to heighten awareness of the problem of bullying among these various groups: students, parents, teachers, and other school officials. A second goal is to develop clear rules against bullying, based partly on input from students, and to ensure that everyone is aware of the rules. A third goal is to provide support

and protection for the victims of bullying, both directly from teachers and through consultations with the parents and peers of victims. A final goal is implied by those already stated, and it is to actively involve teachers and parents in every phase of the attempt to control bullying. Beyond this general prescription, the program provides a number of specific suggestions for achieving such involvement—for example, guidelines about how to respond when the rules about bullying are violated, or ways to help victims become more confident and assertive.

As noted, more recent intervention programs have typically adopted at least some elements of the Olweus approach. Some differences in emphasis are also evident, however. Although most interventions are school-wide, a subset of projects take a targeted approach, that is, are directed specifically to known bullies and victims (typically, primarily the former). Another difference is that some programs place more emphasis than does the OBPP on influencing how bystanders respond to bullying (e.g., Kärnä et al., 2011). Two findings from the general bullying literature suggest that such an emphasis should be helpful (as indeed it turns out to be). One is that bystanders intervene in only a minority of bullying episodes, even though most children say that they disapprove of bullying; when bystanders do intervene, however, they are often successful at stopping the aggression (Hawkins, Pepler, & Craig, 2001). The other finding is that receiving support from at least one friend can lessen both the probability of being bullied and the negative effects from such experience (Sainio, Veenstra, Huitsing, & Salmivalli, 2011).

How successful are antibullying programs? The original application of the OBPP was quite successful, resulting in reductions of 50% to 70% in incidents of bullying in the schools selected for study (Olweus, 1993). Most subsequent interventions, including applications of the OBPP in the United States, have obtained lower levels of success. One recent summary of the literature reported values of 20% to 23% for reductions in bullying and 17% to 20% for reductions in

victimization (Bradshaw, 2015). I should note, however, that other reviews have arrived at somewhat different values, in some cases more positive (e.g., Zych, Ortega-Ruiz, & Del Rey, 2015) and in some cases more negative (e.g., Merrill, Gueldner, Ross, & Isava, 2008). Two general conclusions seem safe. A first is that although advances have been made, the problem of bullying is far from solved, and more research is clearly needed. The second is that a host of methodological factors can affect what we conclude about intervention, including the groups being compared, the form or forms of intervention under study, and the outcome measures used to evaluate success. I will return to these points shortly; first, however, I add a second example of intervention research.

## Memory Training

Our second example of intervention research concerns memory problems in old age. One common stereotype of old age—and one common complaint of many older adults—is of problems with memory. As we will see in Chapter 15, there are many qualifications and exceptions to this stereotype; nevertheless, memory difficulties do constitute a genuine, and distressing, reality for some older adults. The issue to address now is whether interventions in the form of memory training programs can undo some of these difficulties.

The general answer to this question is yes, as reviews of the memory training literature indicate (A. L. Gross et al., 2012; Karbach & Verhaeghen, 2014). Here I first present an example of a representative intervention study, after which we will consider some of the general issues raised by such efforts.

Woolverton and colleagues (Woolverton, Scogin, Shackelford, Black, & Duke, 2001) examined the possibility of memory training in a sample of older adults ranging in age from 60 to 88 (mean = 71). The forms of memory targeted were ones that are known to be frequent foci for memory complaints in old age: memory

for names, memory for locations of household objects, and memory for dates and appointments. The training focused on **mnemonic strategies**—that is, techniques that people can use to facilitate their memory performance. As is discussed in both Chapter 13 and Chapter 15, a large research literature indicates that strategies play an important role in both individual and developmental differences in memory; not surprisingly, therefore, strategies are a major component in many memory training programs. In the Woolverton et al. study, participants received 24 one-hour training sessions during which they were taught and then practiced a variety of mnemonic strategies, including use of physical reminders (e.g., notes), categorization of material, and mental imagery.

This training proved beneficial. Participants who had received training improved on a variety of memory measures from the pretest to the posttest; they also outperformed a nontrained control group. Effects of the training were still evident one month later.

## General Issues

The example studies just considered illustrate many of the issues that arise in intervention projects more generally. Table 8.6, which is adapted from Flay et al. (2005), summarizes the points to be discussed.

A first issue concerns identification of participants: To whom will the treatment be applied? As we saw, antibullying programs vary in this regard; although some specifically target bullies and victims, most are school-wide efforts that do not require identifying any particular population. Memory programs for older adults show a similar distinction. In some cases (as in the Woolverton et al., 2001, study), the target group is older adults in general, with no particular selection criteria beyond age and willingness to participate; in others, the focus is on older adults who are known to be experiencing memory difficulties. In many other domains of intervention the target group

Table 8.6  
*Criteria for Evaluating Intervention Programs*

Criterion	Description
Identification of target group	Use of assessment methods sufficient to identify participants in need of intervention and to specify the nature of the sample
Appropriate control group	Inclusion of an untreated group that is equivalent to the target group prior to intervention
Generalization of effects	Demonstration that positive effects of the intervention generalize to outcomes other than those directly trained
Maintenance of effects	Demonstration that positive effects of the intervention persist once the program is no longer in effect
Meaningfulness of effects	Demonstration that positive effects are not merely statistically significant but rather make a genuine difference in the lives of those affected
Clarity and practicality	Creation of an intervention that is both clear enough and practical enough to be successfully applied by others
Fidelity	Extent to which subsequent applications of an intervention retain its critical features

is a distinct subset of the general population—not preschoolers in general, for example, but only those with attentional deficits, or not adolescents in general but only those with a history of substance abuse. It is in such instances that assessment becomes especially important: We need a measure that can accurately identify those in need of help and perhaps also specify areas of need in a way that can guide subsequent intervention.

Once prospective participants have been identified, considerations of experimental design become important. At the core of most intervention projects is the sequence labeled in Chapter 3 as the *One Group Pretest-Posttest Design*: an initial assessment, followed by the intervention, followed by the final assessment. As we saw in Chapter 3, a significant pretest-posttest difference in such research is not certain evidence of the value of the treatment, because this design is subject to various alternative explanations or threats to the validity of its cause-and-effect conclusions. For this reason, many interventions

randomly assign a subset of the participants to an untreated control group. Such was the case in both the Olweus (1993) and Woolverton et al. (2001) studies, although the assignment was at the level of schools rather than individual participants in the Olweus research. The control group provides an estimate of the contribution of factors other than the experimental treatment (e.g., practice effects, regression to the mean), thus allowing a clearer determination of the gains that are specific to the intervention.

As is often the case with matters of design, however, implementation of some methodologically desirable procedure may be complicated by pragmatic difficulties. Participants may balk, for example, at being assigned to a no-treatment group, thus introducing the possibility of selection bias in the formation of groups. In some instances two of the threats to validity listed in Table 2.5 may apply: compensatory rivalry, in which the control group shows enhanced motivation to succeed, or resentful demoralization, in which the opposite effect occurs. Finally,

ethical issues may arise if some needy individuals are denied a potentially beneficial treatment for purposes of experimental control, an issue discussed in Chapter 10 under the heading of *withholding treatment*. Because of such ethical concerns, many intervention projects offer some form of assistance to control participants following the completion of formal data collection. In the Woolverton et al. (2001) study, for example, the control participants received comparable mnemonic training at the conclusion of the study.

Accurate assessment is important not only to identify participants at the time of the pretest but also to determine effects of the intervention at the time of the posttest. What we are interested in, clearly, is documenting improvement following the intervention—improvement above and beyond what might be attributable to factors other than the intervention (thus the need for a control group) and improvement that is not merely statistically significant but also pragmatically important, that is, that makes a meaningful difference in the lives of the participants. Exactly what is reasonable or desirable to expect will vary across different projects. Two general considerations, however, are important in the evaluation of many intervention efforts, including both antibullying and memory training programs.

One concerns *transfer* or *generalization*: Do the benefits of the treatment extend beyond the specific tasks, stimuli, and contexts that were involved in the intervention? In the case of a study like that of Woolverton et al. (2001), for example, is there transfer to forms of memory other than those targeted by the training or to any sort of memory performance outside the laboratory context? As it happens, the Woolverton et al. report is silent on the issue, given that none of these forms of transfer was included in the study. Other projects, however, make clear that some degree of generalization is a common, although not inevitable, finding in memory training research. Similarly, in addition to direct effects on bullying, at least some antibullying interventions have produced positive effects on other

outcomes, for example, increases in self-esteem and peer acceptance.

The second general issue is *maintenance*: Do the benefits of the intervention persist beyond the study itself, or do they fade away once the program is no longer in effect? As we saw, the Woolverton et al. (2001) study demonstrated that effects held up on a delayed posttest, albeit one that was not very delayed. Both the original and subsequent applications of the Olweus program reported positive levels of maintenance once the intervention was terminated, that is, continued changes in school climate and in frequency of bullying (Olweus & Limber, 2010). Maintenance is by no means always found in these literatures, however. Indeed, a basic methodological point is that many intervention projects—both in the memory training and bullying domains and more generally—fail even to include tests of transfer or maintenance. And when such tests *are* included, a general conclusion seems safe: Immediate effects are easier to produce than long-lasting effects, and narrow and specific effects are easier to produce than broad and general ones.

A final issue concerns future applications of the intervention. As noted, many interventions are intended not simply to improve the lives of the participants to whom they are directed but also to furnish information that can be used to help the target population more generally in the future. The goal, in other words, is to create a program that others can use. To achieve this goal, the intervention must be of feasible scope—that is, it cannot be so time-consuming and demanding of resources that others cannot implement it successfully. It must also be clearly conveyed so that others can identify and apply its critical elements (the “Fidelity” entry in Table 8.6). I will add that antibullying programs, especially those of the school-wide sort, are especially challenging in this respect. Some interventions are relatively easy to apply faithfully, given that they involve a limited target population and a handful of researchers. In contrast, a program such as the OBPP requires extended involvement by dozens of individuals in multiple settings, and these individuals change



from one application of the program to another. One difficulty in deriving clear conclusions about such programs is that specific implementations may vary so widely (Ryan & Smith, 2010).

## Summary

This chapter is directed to *applied research*—that is, research whose explicit goal is to prevent or correct problems in development and to optimize developmental outcomes. Three forms of applied research are discussed.

The first topic considered is *assessment*. The purpose of assessment is to identify important characteristics of some population of interest—most often, negative characteristics in a population perceived to be at risk for developmental difficulties. The NBAS is described as an example of an assessment instrument. The ability of the NBAS to identify effects of early risk conditions is one kind of evidence for its validity; another kind comes from its ability to predict future developmental status. The predictive power is far from perfect, however—a point that applies in varying degrees to every psychological assessment instrument.

Another important form of applied study is *research directed to socially important issues*. Two examples are discussed. The first is *forensic developmental psychology*: the study of children in legal contexts, especially as witnesses and especially in cases of suspected abuse. Research on children as witnesses builds on but also extends the relevant basic-science literature—a point that applies in general to applied research. In the present case, two challenges must be met: devising or discovering situations that bear some similarity to instances of abuse and determining the effects of questioning that is similar to that in real-life forensic cases.

The second example of a socially important issue is bullying. The last several decades have seen the emergence of a large research literature directed to the nature and prevalence of bullying, the characteristics that predispose children to

the bully or victim roles, and the consequences of being a bully or a victim. In recent years the research focus has expanded to include cyberbullying: bullying committed not in person but rather through activity on the Internet.

A third important form of applied study is *intervention research*: programs designed to correct or prevent problems. Antibullying programs in school settings and memory training programs for the older adults serve as the examples from which general points about intervention are drawn. One such point concerns the need to identify prospective participants, a task for which an accurate assessment device may be essential. Assessment is also important at the termination of the intervention to document effects; among the questions that are usually of interest are the *transfer* of effects beyond the intervention context itself and the *maintenance* of the effects over time. Finally, a goal of many intervention projects is to develop programs that can be applied more broadly; it is important, therefore, that programs be feasible and their essential elements well understood.

## Key Terms

bullying	mnemonic strategies
cyberbullying	Neonatal Behavioral
forensic	Assessment Scale
developmental	
psychology	

## Exercises

- Clearly, all of the topics listed in Table 8.4 are important issues that are worthy of study. Suppose, however, that you had to select just one of the topics as especially deserving of an increased research effort in the coming years. Indicate which topic you would select and justify your selection. Ground your selection at least in part in a survey of the relevant literature.

2. An important issue in forensic psychology is labeled the *juror perspective*: the extent to which adults tend to believe different kinds of witnesses, including witnesses of different ages. Imagine that you are a juror about to hear the testimony of an alleged victim of sexual abuse. Imagine also that the alleged victim is either 4 years old or 10 years old or 16 years old. Do you think that your tendency to believe the testimony might vary across these ages? If so, which witness would you expect to find most credible and which least credible? Why?

3. Like most research in developmental psychology, most studies of bullying have been carried out within the quantitative psychology

framework. Imagine that you wished instead to take a qualitative psychology approach to the topic. Describe how you might proceed.

4. The text discusses two examples of intervention research: antibullying programs in school settings and memory training studies with older adults. Select one of the topics in Table 8.4 that especially interests you and outline a possible intervention study for that content area (i.e., the sample that would be studied, the measures that would be used, the form that the intervention or interventions would take, etc.). If possible, locate a review of research on the topic in one of the sources cited in the chapter and compare your ideas with what has actually been done.