Introduction

This manual reports on the history, normative data, psychometric properties, and scoring and interpretation for the Barkley Deficits in Executive Functioning Scale—Children and Adolescents (BDEFS-CA) and its clinical use for children 6–17 years of age. The Appendix to this manual contains the long form of the BDEFS-CA (70 items) for parents as well as the short-form (20-item) version. There is also a brief interview based on the short form for use in unusual circumstances in which a parent is unable to complete a rating scale. Executive Functioning (EF) Profiles are also provided for converting the raw scores from the long form and short form to percentiles for the child's age group (6–11 years or 12–17 years) and sex.

The BDEFS-CA is the culmination of more than 17 years of research and development on identifying the most useful items for the assessment of deficits in EF in daily life activities. Initially development began with forms for use with adults. Now, in this manual, a downward extension of the adult BDEFS has been created and normed for use with parents in evaluating their children ages 6–17 years. The various versions of the BDEFS scales are *theoretically based*, having been initially constructed from models of EF in the neuropsychological literature, particularly mine (Barkley, 1997a, 1997c, 2012b) and those of others (see Chapter 2). But the BDEFS-CA is also *empirically developed* being based on statistical analyses that were used to identify the most reliable and valid underlying dimensions of EF deficits in daily life. Initially, these dimensions were identified in various samples of adults with attention-deficit/hyperactivity disorder (ADHD) along with clinical and community control groups (Barkley & Murphy, 2010, 2011) as well as with children with ADHD (hyperactivity) followed to adulthood (along with a general community control group of children) (Barkley & Fischer, 2011). The BDEFS was then further studied and validated in a large sample of adults, ages 18–81 years of age, that was representative of the United

States (relative to 2000 U.S. Census information). In this manual, the BDEFS is further developed and validated to serve as a parent rating scale for the evaluation of EF deficits in daily functioning in children 6–17 years of age. Not only are the various BDEFS empirically and theoretically based, reliable, and valid, but they are also exceptionally convenient to use. This is due largely to the willingness of the publisher to grant limited permission to the individual purchaser to photocopy the scales and score sheets for use in his/her clinical or research practice. This saves the purchaser a great deal of time, inconvenience, and expense typically associated with the continual reordering of those scales and scoring sheets, as is commonplace in the scale-publishing industry.

As with the adult version of the BDEFS, this manual begins with a brief introduction to the concept of EF in neuropsychology, along with some discussion of the use of this construct in children. I then provide details of the developmental history of the initial BDEFS and its progression to this parent-report form. The remaining chapters provide information on the normative sample, factor analysis, and other information concerning the scale's reliability, validity, and clinical utility, along with instructions for scoring and interpretation of the scale. The information available appears to be sufficient to recommend the use of the scale in a variety of circumstances. These may include clinical practice settings with parents of children for whom concerns exist about possible EF deficits; in research settings in which the evaluation of EF deficits in children is a specific aim, in treatment trials as a measure of response to child- and parent-focused interventions in which a child's EF is of interest, or in school settings in which screening children for possible deficits in EF as reported by their parents is an important objective. I hope you will find the BDEFS-CA rating scales to be useful in the evaluation of EF deficits in children in their daily life activities as reported by their parents.

CHAPTER 1

Executive Functioning in Children

The BDEFS-CA is a rating scale designed to evaluate the major components of executive functioning in daily life activities of children ages 6–17 as reported by their parents. The scale is intended for use by professionals to evaluate any child within this broad age range for whom there is concern about deficits in EF, such as children with neurological, developmental, or psychiatric disorders or those having psychological difficulties with which deficits in EF may be thought to be associated. The scale is both theoretically and empirically based, having been founded on common conceptualizations of EF, particularly my theory of EF (Barkley, 1997a, 1997c, 2012b), and on statistical analyses revealing the basic underlying dimensions of EF deficits of the items used to create the scale. This chapter presents a brief overview of the history of EF, efforts at defining the construct, and difficulties that have arisen in the use of psychometric tests to evaluate EF and its components.

A Brief History of EF

According to theorists (Dimond, 1980), the term "executive functioning" appears to have originated with Karl Pribram (1973) as a description of the overarching functions of the prefrontal cortex (PFC), though some (Wolf & Wasserstein, 2001) attribute the initial use of the term to Butterfield and Belmont (1977). Pribram (1973) stated that "the frontal cortex is critically involved in implementing executive programmes where these are necessary to maintain brain organization in the face of insufficient redundancy in input processing and in the outcomes of behavior" (p. 301). The study of those "programmes," however, goes back more than 100 years before the invention of the term EF, originating in the medical and scientific interest in the functions of the PFC (Bekhterev, 1905–1907; Harlow, 1848, 1868; Luria, 1966). The history of EF, therefore, shows its inherent conflation with the functions of the PFC (and vice versa) such that the attempt to understand EF at the neuropsychological level of analysis is routinely confused with or reduced to the neurological level; EF is simply what the prefrontal lobes do (Denckla, 1996; Stuss & Benson, 1986). Research, however, has shown that EF is not exclusively a function of the PFC given that the PFC has various networks of connections to other cortical and subcortical zones, as well as the basal ganglia, amygdala and limbic system, and cerebellum (Fuster, 1989, 1997; Luria, 1966; Stuss & Benson, 1986). And the PFC may well engage in certain neuropsychological functions that would typically not be considered to fall under the umbrella of EF, such as simple or automatic sensory–motor activities, speech, and olfactory identification.

Among the first and certainly most famous clinical descriptions involving a PFC injury and its associated neuropsychological deficits is that of Phineas Gage, the railroad foreman who suffered a penetrating head wound that destroyed a large portion of his PFC. This injury led to drastic alterations in his behavior, personality, and social conduct (Harlow, 1848, 1868). Like Gage, other patients with PFC damage described in later textbooks of neuropsychology (Bianchi, 1895) demonstrated a lack of initiative or drive, a curtailing of their circle of interests, profound disturbances of goal-directed behavior, and a loss of abstract or categorical behavior. They also suffered emotional changes, such as a proneness to irritation, emotional instability, and indifference toward their surroundings, often superimposed on depression. Impulsive actions, trivial jokes, and even euphoria were noted to be symptoms of lesions that involved the more basal aspects of the PFC.

Subsequently, in his *Fundamentals of Brain Function,* Bekhterev (1907, as cited in Luria, 1966) observed that damaging the frontal lobes resulted in a disintegration of goal-directed behavior, which Bekhterev viewed as the principal function of the PFC. Later, Bianchi (1922) took a similar view of the purpose of the PFC, as would Luria, noting that "one of the essential results of destruction of the frontal lobes in animals is a *disturbance of the preliminary . . . syntheses underlying the regulation of complex forms of motor operations and the evaluation of the effect of their own actions, without which goal-directed, selective behavior is impossible* (1966, p. 224; original emphasis). In time, this frontal lobe syndrome noted by Luria would evolve into an executive disorder (Fuster, 1997) or a "dysexecutive syndrome" (Baddeley, 1986; Wilson, Alderman, Burgess, Emslie, & Evans, 1996). It would also lead to the modern view that EF is largely concerned with goal-directed action (Welsh & Pennington, 1988).

A Brief Survey of Definitions of EF

Modern definitions of EF relied heavily on these earlier views of the functions of the PFC, declaring that EF consists of a collection of neuropsychological processes involved in the maintenance of goal-directed action, including problem solving (Gioia, Isquith, Guy, & Kenworthy, 2000; Pennington, 1997; Welsh & Pennington, 1988). But the functions of the PFC are numerous. Likewise, so are the number of EFs: Up to 33 such functions were identified in one survey of experts in the field (Eslinger, 1996). The most common functions attributed to the PFC appear to be: anticipation; goal selection; preplanning (establishing ends and means); the integration and cross-temporal execution of goal-directed behavior; the inhibition of more automatic actions and reactions to extraneous stimuli (distractibility); the production of delayed reactions; the monitoring and evaluation of one's goal-directed actions relative to the external environment and the goal, especially in novel circumstances; and the overall intentionality or purposive quality of behavior (Denckla, 1996; Fuster, 1997; Lezak, 1995, 2004; Luria, 1966; Stuss & Benson, 1986). To this list some might also add the functions of drive, motivation, will, and self-awareness (Lezak, 1995; Stuss & Benson, 1986). Welsh and Pennington (1988) reduced this list to just two functions (working memory and inhibition) that interact to resolve conflicts or competing action alternatives in goal-directed activity (Pennington, 1997). These and many other authors acknowledge, however, that "EF" is an umbrella term referring to a collection of related but somewhat distinct abilities, such as planning, set maintenance, impulse control, working memory, and attentional control. Even this list is thought to be provisional and quite general.

Why are these functions considered to be executive in nature? Apparently the only requirement for being an EF is that a brain function somehow must be associated with goal-directed action and problem solving and also with the functioning of the PFC. The fact that such functions are essential when faced with novel situations that require new solutions is only partly helpful. Problematic in these various views of EF is that they assert that various neurocognitive processes are essential for goal-directed action yet (1) often do not specify what those processes actually are and (2) do not indicate just what basis or rationale qualifies them to be considered executive processes while others, equally as important for goal-directed action, are not so considered. What makes a neurocognitive process "executive"? It cannot just be that it contributes to goal-directed action, because conscious sensory–motor actions can be considered to make such a contribution, whereas walking or hand movements alone, for instance, do not seem to fit the term "executive" as it is used in the literature.

The definition of EF has also been approached from an information-processing perspective in which the computer metaphor of brain functioning has predominated. Typical of that viewpoint of EF was the work of Borkowski and Burke (1996) and other authors whose work they summarized in this field. Borkowski and Burke described EF as a set of three components that are directed at problem solving: task analysis (essentially defining the problem), strategy selection and revision, and strategy monitoring. Those authors also cited a different information-processing model developed by Butterfield and Albertson (1995) that views EF as one of three components of cognition: cognition, metacognition, and EF. A similar information-processing view of EF explains it as follows: "The umbrella concept of 'executive control' encompasses those cognitive functions involved in the selection, scheduling and coordination of the computational processes responsible for perception, memory and action (Norman & Shallice, 1986; Shallice, 1994)" (Ciairano, Visu-Petra, & Settanni, 2007, p. 335).

A quite different view of EF was proposed by Hayes, Gifford, and Ruckstuhl (1996) using a more behavior-analytic model and particularly the concept of rule-governed behavior (see Hayes, 1989). Their analysis of the terms often believed to make up EF, as well as many of the tests used to assess EF, led them to conclude that

EF is a special subset of rule-governed (verbally regulated) behavior. Rule-governed behavior is behavior that is being initiated and guided by verbalizations, whether self-directed or provided by others. Hayes and colleagues (1996) argued that EF tasks place people in situations in which previously learned sources of behavioral regulation come into conflict with rules laid down by the task and the examiner. Those task-specific rules are competing with behavior that is otherwise automatic and well practiced. Thus the typical automatic flow of behavior must be interrupted and delayed long enough for the person to discover a new rule or select among previously learned rules that may apply in this situation. Yet interrupting a well-practiced behavior itself often requires that a rule be selected and followed that is initiating the delay in responding. And so in EF tasks an individual often has to implement a rule to inhibit his/her usual ongoing responding, even if he/she is only asked a question about the task. He/she must then either select from among a set of relevant previously learned rules or generate a new one. The latter is a verbal means by which we problem-solve by using second-order rules to discover first-order rules. This view of EF appeared to gain little favor among traditional neuropsychologists, who seemed more comfortable using a cognitive neuropsychological or information-processing view of EF than one derived from behavior analysis. Yet features of it would be incorporated into my own model of EF (Barkley, 1997a, 1997c) through Vygotsky's model of the internalization of self-directed speech.

Barkley's Hybrid Theory of EF

In 1997, I sketched the broad outlines of a theory of EF and applied it specifically to ADHD, a relatively chronic developmental disorder of inattention, impulsivity, and hyperactivity known to be associated with various deficits in EF (Frazier, Demareem, & Youngstrom, 2004; Hervey, Epstein, & Curry, 2004; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). This "hybrid" theory was an amalgamation of many of the ideas to be found in other theories or definitions of EF but began with its foundation based on the ideas of Bronowski (1977) concerning the unique properties of human language and cognition. These ideas were combined with those of Fuster (1997) on the functions of the PFC, especially his emphasis on the PFC being of central importance for the cross-temporal organization of behavior toward the future. This theory was subsequently further developed (Barkley, 1997c) and eventually published as a book (1997a). Unique to this theory of EF is that it specifies that:

1. EF is self-regulation toward a goal specifically and the future generally.
2. Self-regulation is a set of self-directed actions, often private but not necessarily so, that the individual uses to change his/her subsequent behavior so as to alter the likelihood of some future outcome (to attain some goal).
3. There appear to be at least five of these self-directed (executive) activities being used to choose goals and to select, enact, and sustain actions toward them.

Each executive (self-directed) activity is argued to have arisen in the same manner as the self-direction and internalization of speech described by Vygotsky (1962).

Five such self-regulatory actions were elaborated in this initial theory, each of which is a conscious, voluntary, effortful action:

1. *Self-inhibition.* This represents (1) the capacity to consciously or voluntarily suppress or otherwise disrupt or prevent the execution of a prepotent or dominant response to an event (the response that has been previously associated with reinforcement or has the highest likelihood of being performed under ordinary circumstances); (2) the capacity to consciously or voluntarily interrupt an ongoing sequence of behavior toward a goal if it is proving to be ineffective; and (3) the capacity to protect the self-directed actions that will subsequently occur and the goal-directed actions they are guiding from interference by external and internal goal-irrelevant events.

2. *Self-directed sensory–motor action.* This is an alternative means of defining nonverbal working memory and refers to the use of self-directed visual imagery, along with the private rehearsal of visuomotor actions it permits. Although this consists largely of visual imagery, or the "mind's eye," it also includes the other senses, such as private rehearing, retasting, resmelling, and refeeling (kinesthetic–proprioceptive experience). In totality, this component of EF provides for the conscious, willful reexperiencing of past events—a metaphorical Cartesian theater of the mind. It is the initial basis of ideas or ideational thinking. Yet it also includes the capacity to practice or reperform sensory–motor actions privately to oneself, providing a form of mental behavioral simulation.

3. *Self-directed private speech.* This is the alternative explanation for the verbal working memory system to the one offered by information-processing and traditional cognitive models of EF, such as Baddeley's (1986) phonological loop in his model of working memory (Baddeley & Hitch, 1994). It is largely based on Vygotsky's (1962) theory of the internalization of speech to form the mind's voice. Individuals talk to themselves to permit private rehearsal of utterances, as well as engaging in self-directed instructions, self-questioning, and other means by which language can be used for self-regulation and problem solving.

4. *Self-directed emotion/motivation.* This component of EF is believed to arise from the combination of the first three, in which the individual learns to use inhibition, private imagery (and private sensing more generally), and private speech to initially inhibit strong emotions and then to down-regulate or otherwise moderate them. The individual then consciously employs the other components to replace the initial strong emotion with alternate emotional responses more consistent with social demands and the individual's goals and longer-term welfare. Because emotions are motivational states, this component of EF also provides for the capacity for self-motivation—the drive states needed to initiate and sustain action toward the future.

5. *Self-directed play (reconstitution).* This component of EF is an alternative to the planning, problem solving, and inventive or generativity/fluency components noted in other views of EF. It is hypothesized to be founded on the development and internalization of play and is seen as essentially a two-step activity. The first is analysis, or taking apart features of the environment and one's own prior behavior

toward it. This is followed by synthesis, or the recombination of components of the environment and behavioral structures into novel combinations. These novel combinations can then be tested against a criterion, such as a problem or goal, for their likely effectiveness in overcoming obstacles to goal attainment (i.e., solving the problem). Although this component begins as observable manual play in young children, it progresses, like the other EF components, to being turned on the self and internalized as a form of private mental play. This permits the manipulation of mentally represented information both about the environment and about prior behavioral structures so as to yield new combinations that can serve as options for goal-directed problem solving. It is possible that this component can be subdivided into both nonverbal and verbal modules that provide for fluency and generativity in each of these forms of behavior. Nonverbal, verbal, and action fluencies are believed to arise from this component of EF.

Unlike many prior efforts to define or conceptualize EF, this view provides a specific operational definition of it; EF consists of acts of self-regulation across time toward future goals. The theory specifies five components or types of self-directed actions that humans use for purposive, goal-directed behavior and that become self-directed: (1) inhibition, (2) sensory–motor action, (3) speech, (4) emotion, and (5) play. The theory states that these functions are executive in nature because they are self-directed actions that alter future consequences. Actions that are not self-directed, including any that are used to attain the goal, are not considered executive in nature (e.g., walking, seeing, hearing, balance, spatial orientation, speech to others). The self-directed actions that are guiding those other activities *are* executive in nature.

More recently, I have greatly elaborated on this initial theory of EF (Barkley, 2012b). First, I have now separated out self-directed attention that gives rise to self-awareness and self-monitoring as a sixth EF component. Previously it was housed under the self-directed sensing component. The two components may still be a single function. Yet, given the importance of self-awareness to evaluating both current and future (desired, possible) states that serves to initiate the other EF components, it seemed to deserve separate standing as an EF in its own right. Second, and more important, I borrowed the concept from evolutionary biology of an extended phenotype (Dawkins, 1982) in order to better understand and explicate the profound impact that EF produces in human social living and cultural life. EF is a person's principal means of survival. Its effects can be observed to radiate outward across spatial, temporal, and social distances from the individual and his/her genotype beyond merely his/her immediate appearance and behavior. These extended phenotypic effects-at-a-distance have a considerable impact on a person's adaptive functioning, reproductive success, and survival. In this expanded model, the six EFs described earlier form the Instrumental–Self-Directed level of EF that is most proximal to PFC development and functioning. It represents the usually private, internalized, or cognitive form of EF by young adulthood. But this level is not sufficient to account for how humans employ EF in daily life activities toward accomplishing their goals in major domains of functioning across days, weeks, and months of time. To do that, one must describe the social functions of EF and create a hierarchy of levels of EF that will emerge over the three decades it appears to take the EF/PFC system to mature.

Like models of driving (Ranney, 1994), EF in this more expanded view (Barkley, 2012b) comprises a hierarchy of increasingly organized sets of goal-directed actions that consist of nested sets of shorter-term subgoal/action complexes. These sets and the lesser goals they achieve can be nested together and strung into sequences to accomplish increasingly larger, more abstract, and longer-term goals. Those goals increasingly involve others, through acts of self-defense and self-reliance to reciprocity, on to cooperation, and eventually to mutualism. Individuals also increasingly utilize cultural devices to achieve their goals as they progress up the hierarchy of the EF extended phenotype. These levels, in order, comprise the Methodical–Self-Reliant, Tactical–Reciprocal, Strategic–Cooperative, and Principled–Mutualistic ones that are perched atop the initial Instrumental–Self-Directed one. Seven developmental capacities are argued to arise from the six Instrumental EF components that increase with maturation and permit the movement upward into higher levels of the EF extended phenotype. These are: (1) increasing behavioral complexity, (2) increasing self-restraint (inhibition), (3) contemplation across increasing spatial distances, (4) contemplation over increasing temporal fore-periods (the future), (5) increasing creation and sustainment of self-motivation, (6) increasing use of and reliance on others to achieve goals, and (7) an increasing use of cultural information and devices for goal selection and attainment. Injuries to or maldevelopment of the EF/PFC system can be seen to result in a collapsing downward of this hierarchical arrangement and a contraction in the seven capacities arising from the development of that system.

All of the Instrumental–Self-Directed EF components used for engaging in goal-directed actions are initiated by, directed toward, and intended for the benefit of oneself and one's self-interests. Thus EF is, by definition, *self-regulation for the purposes of one's longer-term or later self-interests.* As noted previously, self-regulation is any action directed at oneself that changes one's subsequent behavior so as to alter the likelihood of a future outcome (a goal) for oneself (Barkley, 1997a, 1997c; Skinner, 1953). The chief executive is therefore *the self*—the conscious sense of ourselves over time (past, present, and future) and also of our needs, wants, and desires (self-interests) as we perceive them to be *over time.*

This expanded EF phenotype respects the fact that the goals toward which EF and its components are being used are largely *social goals and occur in the context of a social group.* Those social goals are being pursued often with others, sometimes against others, and nearly always in the larger context of others. That is the environmental niche in which EF likely evolved, and the problems posed by it are the ones it most likely evolved to solve (Barkley, 2001, 2012b; Dimond, 1980). From this model, I define EF as being *those self-directed actions (self-regulation) used to choose goals and to select, enact, and sustain actions across time toward those goals typically in the context of others and often using social and cultural means* (Barkley, 2012a). EF is a *meta-construct* in the sense that such self-regulation requires multiple self-directed actions (or interacting mental modules or neuropsychological capacities) each of which contributes to EF (the meta-construct). Disturbances or difficulties in the effective functioning of any one type of self-directed activity or mental module disrupts EF. But it does so in a different way from disturbances in other modules. In other words, EF, defined as self-regulation, is not one thing but involves multiple types of self-directed activities.

Executive functions (self-directed activities) are not independent but interactive and probably hierarchically organized in development. The capacity to hold information in mind (working memory), for example, using visual imagery, requires that the individual inhibit responding to momentary external events and resist interference by competing sources of goal-irrelevant information (inhibition and resistance to distraction). The holding of such information in mind can be further enhanced by self-directed speech, or the phonological loop in Baddeley's (1986) model of working memory. Both of these activities (speech and imagery) may then be subjected to mental manipulation (self-directed play, or the fluency, generativity, or reconstitutive EF) to develop multiple possible options for responding to the situation. The emotional restraint and self-motivation that may be required to drive the goal-directed behavior toward the intended goal may be facilitated or even initiated by self-imagery and self-speech. Because a problem can be defined as a situation for which one does not have a readily available response, it poses by definition a novel or nonroutine situation for the individual. Therefore, EF and its self-directed activities will appear to be most needed and most activated when *novelty* has been encountered.

Developmental Considerations

As just summarized, in my view EF consists of those self-directed actions that permit self-regulation. These self-directed actions are founded on preexecutive mental and behavioral activities that are directed at controlling or reacting to the immediate environment and others within it. But during development some types of human action become self-directed to achieve self-change. Those self-directed activities, or EF capacities, first develop as actions that are largely publicly observable. Over maturation, the public aspects of these actions become inhibited or privatized such that the person can engage in this activity without it activating the peripheral musculoskeletal system; the action can be engaged in within the brain just as if it were going to be publicly expressed, but in this case it is not released into the spinal cord for actual execution. A form of private mental simulation of actions results from this developmental internalization of human self-directed actions. That effort at self-management is being done to alter the likelihood of a later consequence (a goal).

The extended phenotype model of EF, like other views of EF (Anderson, 2002), argues that these various self-directed activities do not arise all at once but in a sequence. Most likely, the first EF component to arise in early child development is self-awareness (self-directed attention), followed closely by inhibition and interference control (self-restraint) and nonverbal working memory (sensory–motor actions toward the self). These serve to initiate self-awareness over time—a conscious sense of time, past, and the future (hindsight, foresight). Within a few years, these initial EF components are followed by the development of other self-directed actions, such as verbal working memory (self-speech), emotional/motivational self-control (self-directed emotional states), and eventually planning and problem solving (self-directed play). By adulthood, much self-directed activity is private, being unobservable or very difficult to observe and therefore is cognitive or mental in form (self-awareness, visual imagery, private audition and speech, private emotion regulation, and mental play, for instance). Other self-directed activities may be more public and outwardly observable (i.e., observable self-speech, writing notes to oneself, posting cues or other stimuli in one's work or living space to increase the likelihood of a later behavior, meditating to reduce a strong negative emotional state, doodling various options for solving a problem on paper, etc.), but they are executive nonetheless.

Obviously, the extent to which EF may be deficient or impaired in a child can be gauged only by making reference to same-age peers and to our understanding of the developmental trajectories of the EF components. An EF component cannot be viewed as delayed or deficient in a child of a particular age if that component has not yet arisen, as in infants or toddlers. The assessment of EF deficits in children, whether by test, rating scale, or direct observational recordings, requires that normative information be available for that measure for a wide range of ages throughout childhood development. Such information is critical in permitting a particular child's score to be compared with those of others of the same age in making the determination that a delay or deficit in that EF component is present. Should such developmental trajectories of the EF components differ as a function of sex, then comparisons with same-sex norms for that age group would also be indispensable.

Typically, developmental models of EF have been based on studies employing factor analyses of psychometric test batteries (Anderson, 2002). To the extent that this approach to measuring EF may be seriously limited (see Chapter 2), such models of EF development will also be limited. Studies using factor analysis usually conclude that EF is both unitary (a single large factor often emerges that accounts for a majority of variance, or evidence is found that the factors are at least significantly related to each other) and diverse, involving multiple components (typically represented by smaller, more specific factors that account for far less variance) (Miyake et al., 2000; for reviews, see also Anderson, 2002; Barkley, 1997a; Best, Miller, & Jones, 2009). Less often acknowledged is that some of these factors likely arise due simply to shared method variance, or the use of multiple scores from a single test that give rise to a factor that simply represents that type of test. In many instances, these specific tests of EF are impure, being confounded by other nonexecutive psychological functions such as language, IQ, and visual–spatial and mechanical abilities, among others (Anderson, 2002; Best et al., 2009). On average, four or more factors are commonly observed across these studies, typically labeled as reflecting inhibition, attentional control, working memory, cognitive flexibility, and planning/problem solving. Often the nonexecutive factor of response speed is also evident in many of these studies. Of these, several are believed to be foundational and even inseparable, these being inhibition and working memory (Bell, Wolfe, & Adkins, 2007; Pennington, 1997). In contrast, rating scales of EF behavior in daily life typically identify one or two large initial factors related to (1) behavioral inhibition (cognitive, motor, and verbal) and emotion regulation and (2) metacognition (working memory, planning, self-monitoring, etc.) in children (Gioia et al., 2000; Thorell et al., 2010). In adults, up to five factors may be evident in such ratings (Barkley, 2011a; Barkley & Murphy, 2011), reflecting time management, self-organization and problem solving, self-restraint (inhibition), self-motivation, and self-regulation of emotion.

Because these EF components may develop at different rates, their trajectories over maturation are not identical or necessarily linear (Anderson, 2002). Nor do these components start their development at the same time or reach their apex of maturation simultaneously. Inhibition (and resistance to distraction) is often viewed as being the first EF component to arise. It seems to be more important to problem solving in early (preschool) childhood, when neuropsychological tests are used to evaluate EF development. This attribution of inhibition as the first of the EF components to arise may be the result of the fact that most of these developmental studies do not have specific measures of self-awareness, which, as argued earlier, ought to be considered as the first EF component or at least as developing simultaneously with inhibition. Other EF components besides inhibition are apparently less differentiable from inhibition in the preschool age group (Hughes, Ensor, Wilson, & Graham, 2010). Available evidence suggests a rapid rise in inhibitory ability during these preschool years, with continuing but less rapid improvement during the school-age years. By early adolescence, this capacity to suppress prepotent responses and demonstrate interference control (resistance to distraction) approaches adult levels of performance (Best et al., 2009).

Beginning in the late preschool to early school years, first working memory (WM) and then set shifting begin their developmental ascent and can be distinguished from the inhibitory EF component (Huizinga, Dolan, & van der Molen, 2006). WM appears to be related to set shifting ability despite these two constructs being distinguishable EF components during childhood and adolescence (Huizinga et al., 2006). Both components continue to mature in a linear fashion until mid- to late adolescence (Huizinga et al., 2006) and may be relied on more often in problem solving by school-age children or adolescents (Best et al., 2009) than was evident in the preschool age group. Once the capacity to inhibit responding to distracting events and to inhibit prepotent responses more generally has developed sufficiently, it permits the child to develop and utilize working memory and mental flexibility for more complex problem solving. Later in the school-age years of development, planning, which is thought by some to be the pinnacle EF, may emerge as a more important EF for both problem solving and goal-directed behavior (Anderson, 2002; Best et al., 2009). It may not achieve its apex or mastery until late adolescence. Across all of these EF components, the age of mastery is directly related to the difficulty of the task being used to assess that component of EF (Schleepen & Jonkman, 2010). For instance, even preschool children seem capable of mastering WM tasks in which just a few pieces of information must be held in mind. But when more units of information are added, the age of mastery may shift into later childhood or even late adolescence (Huizinga et al., 2006).

The development of these components of EF as evaluated by psychometric tasks is related to the development of both social and emotional self-regulation—a relationship that appears to increase with age (Best et al., 2009). This supports the previous contention that EF in humans is highly important for social functioning and not just for the performance of purely cold cognitive tasks, such as WM tests or academic performance. However, the low ecological validity of EF tests (see the next section) may limit what are in actuality even stronger relationships between EF and social functioning. Such low validity would argue for the inclusion of behavioral ratings of EF in future studies of EF development (Anderson, 2002). For instance, Vazsonyi and Huang (2010) found that ratings of self-control, a construct highly related if not identical to EF, that were collected at 4 years of age were highly predictive of children's trajectory of deviant (antisocial) behavior over the next 6 years. It may be that rating scales of EF will prove more valid and important in understanding the development of these relationships between EF and sociocultural functioning than have test batteries that are largely devoid of social factors and that ascertain behavior across relatively brief periods of time, often in highly structured and unnatural clinical settings.

Assessing EF: The Problems with EF Tests

A number of tests have been declared to be measures of EF in the neuropsychological literature (see the meta-analyses of Frazier et al., 2004, and Hervey et al., 2004, for a lengthy list of those EF tests used just in studies of ADHD; see Lezak, 2004, for a more comprehensive review). A major premise in many studies and in the clinical evaluation of EF is that these EF tests are the gold standard for measuring EF. As I have described elsewhere (Barkley, 2011a, 2012b), this largely unquestioned assumption can be criticized for many reasons:

- *Most tests used to assess EF were not originally designed to measure EF* (see Lezak, 2004). Many EF tests were originally developed to assess other psychological functions, such as attention, memory (both verbal and visuospatial), sequencing, abstract reasoning, and language. Others were intended to directly assess response inhibition, planning, and problem solving without regard to these constructs being involved in EF. Many do not assess frontal lobe (EF) functions exclusively or differentially (Dodrill, 1997). But when EF is viewed as a collection of various brain functions that include these constructs, or once it could be demonstrated that deficits on the test were apparently associated with injuries to the PFC, these tests were appropriated for use as EF tests without regard to whether or not those tests actually were sampling the *conceptual* domain of EF as a neuropsychological construct.

- *No consensus definition of EF exists that can be used as the standard for determining the construct validity of EF tests.* As shown earlier, EF is a term that is both general and provisional. It is also quite ambiguous. There is a lack of agreement among researchers on the precise meaning of the term (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006; Denckla, 1996; Willcutt et al., 2005). Instead, reviews of the research seem to focus more on listing the constructs thought to be subsumed under the term and the tests believed to evaluate those constructs, such as response inhibition, resistance to distraction, working memory, planning and problem solving, and set shifting, among others (Boonstra, Oosterlaan, Sergeant, & Buitelaar, 2005; Frazier et al., 2004; Hervey et al., 2004; Willcutt et al., 2005). When no operational definition of EF exists, anything goes. This status ensures that ambiguity will abound with regard to what tests may or may not be considered to reflect the conceptual domain of EF.

• *The structured clinical setting is ill suited to assessing self-regulation across time and its use in novel settings.* EF, as noted previously, is most in demand in ambiguous or novel situations, especially social ones, that are often lacking in external guidance or structure as to how to behave effectively in such settings. The individual must engage in acts of self-regulation and problem solving so as to successfully negotiate the situation, solve the problem, or attain the goal (either self-chosen or assigned by others). These factors rarely exist in one-to-one clinical settings in which a child is provided with specific tasks under the guidance and supervision of an adult (Anderson, 2002; Gioia et al., 2000). In contrast, rating scales can be used to capture a far wider range of child behavior across multiple contexts and months of time. Parents clearly possess a substantial body of knowledge about their children's self-regulation and EF that can be captured by rating scales and used to great advantage in the clinical evaluation of EF deficits in their children and in treatment planning for such deficits.

• *Traditional tests of EF cannot evaluate the cross-temporal nature of EF as used in daily life because of their small ascertainment windows for sampling behavior in the clinic (typically 5–30 minutes per test).* EF tests use exceptionally short temporal durations relative to the hours and even days over which children and adolescents sustain their goal-directed activities and weeks to months over which adults are likely to do so. Such a short ascertainment window makes it difficult if not impossible for EF tests alone to capture the lengthy cross-temporal structures of normal human action. By itself, such a problem guarantees that tests will be related only modestly, if at all, to other means of evaluating EF in naturalistic settings, such as by observations over days or by ratings over weeks and months. Ratings of EF in daily life, on the other hand, do sample behavior across considerably longer periods of time (weeks to months). Ratings or observations, therefore, may be better indicators of the cross-temporal organization of behavior and problem solving toward goals in naturalistic settings than might EF tests.

• *Some of the most important features of EF are not captured by EF tests.* Tests of EF do not evaluate many of the capacities believed to be central to the construct of EF (Dimond, 1980; Eslinger, 1996; Lezak, 1995, 2004). They do not capture constructs such as volition and human will, intentionality or the purposive quality of behavior, self-awareness (of self, context, and others), and even aspects of planning (foresight, objectivity, choice and comparative judgment, hierarchical structuring) and plan execution (self-motivation, self-monitoring, prolonged resistance to interference by goal-irrelevant events, etc.) (Lezak, 1995).

• *EF tests do not directly evaluate self-regulation.*

• *EF tests do not capture the principally social functions of EF, such as reciprocity and social exchange, competition, cooperation, and mutualism more generally* (Barkley, 2012b). EF can be usefully viewed as an extended phenotype, as I have recently shown (Barkley, 2012b). This means that EF produces significant effects at considerable spatial and temporal distances and across substantial social networks that influence the survival and reproductive success of people. Among these extended phenotypic effects, the most important for human survival are social activities such as self-interested reciprocity, competition, cooperation, and mutualism (the mutual or reciprocal concern for the longer-term welfare of others). The extended phenotype model of EF views it as a hierarchy comprising these levels of important human social activities that contracts or even collapses as a consequence of PFC damage or maldevelopment. As presently constructed, EF tests do not capture this essential social purpose of the EF extended phenotype.

- *EF tasks are contaminated by multiple nonexecutive cognitive processes.* EF tasks are rarely pure measures of EF and often require various nonexecutive psychological abilities in their performance (Anderson, 2002; Castellanos et al., 2006). For instance, many EF tests are often found to be significantly influenced by overall general cognitive ability or level of intelligence (Mahone, Hagelthorn, et al., 2002; Riccio, Hall, Morgan, Hynd, & Gonzalez, 1994). This makes their results difficult to interpret as reflecting unadulterated measures of EF. Statistically removing IQ from relationships between EF tests and observations and ratings of EF in natural settings may reduce any significant relationships to nonsignificant status (Mahone, Hagelthorn, et al., 2002; Mangeot, Armstrong, Colvin, Yeates, & Taylor, 2002). And it may also account for the fact that some of the strongest relationships noted to date have been between EF tests and academic achievement scores (Biederman, Petty, Fried, Black, et al., 2008; Gropper & Tannock, 2009; Thorell, 2007) or self-ratings of academic performance (Ready, Stierman, & Paulsen, 2001). Given that both academic achievement and self-rated academic performance are significantly related to IQ, not to mention shared method (testing) variance when academic tests are used, this finding is not surprising. In contrast, EF rating scales are typically not associated with IQ (Alderman, Burgess, Knight, & Henman, 2003; Barkley & Murphy, 2011), and so the issue of contamination by general cognitive ability is far less problematic for ratings than for EF tests.

- *Functional deficits in patients experiencing frontal lobe injuries or those with ADHD presumed to have a frontal lobe disorder are often not detected by EF tests.* Ratings of EF in daily life activities or direct observations of EF performance in natural settings are often found to be superior to EF tests in detecting impairment (Alderman et al., 2003; Barkley & Fischer, 2011; Barkley & Murphy, 2011; Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Kertesz, Nadkarni, Davidson, & Thomas, 2000; Mitchell & Miller, 2008; Wood & Liossi, 2006). Given the strong linkage of EF to PFC functioning in the history of EF (discussed earlier), it is unlikely that those patients having disorders of the PFC would not likewise show impairment in EF. This situation has led some to argue, however, that ADHD is probably not a disorder of EF given that the majority of cases show no deficits on EF tests (Boonstra et al., 2005; Jonsdottir, Bouma, Sergeant, & Scherder, 2006; Willcutt et al., 2005), a conclusion much harder to justify logically in patients with demonstrated PFC damage.

- *EF tests have very low or no ecological validity.* In fact, they correlate poorly if at all with ratings of EF in daily life activities in natural settings. This is true in adults (Alderman et al., 2003; Bogod, Mateer, & MacDonald, 2003; Burgess et al., 1998; Chaytor, Schmitter-Edgecombe, & Burr, 2006; Hummer et al., 2011; Ready et al., 2001; Wood & Liossi, 2006) and children with frontal lobe lesions, in people with traumatic brain injuries (TBI), or in those having other neurological or developmental disorders (Anderson, Anderson, Northam, Jacobs, & Mikiewicz, 2002; Mangeot et al., 2002; Vriezen & Pigott, 2002; Zandt, Prior, & Kyrios, 2009). This is also the case both in adults with ADHD and in children with ADHD followed to adulthood (Barkley & Fischer, 2011; Barkley & Murphy, 2010). The results of those studies usually reveal that any single EF test shares 0–10% of its variance with EF ratings. The relationships are frequently not statistically significant. Even the best combination of EF tests shares approximately 12–20% of the variance with EF ratings or observations as reflected in these studies. Yet these two types of measurement are supposed to be measuring the same construct—EF. If IQ is statistically removed from the results, the few significant relationships found in these studies between EF tests and EF ratings may even become nonsignificant (Mangeot et al., 2002). Something is terribly amiss here if different approaches to measuring the same construct are found to be so poorly related to each other.

Additional evidence for the low ecological validity of EF tests comes from studies in which the performance of frontal-lobe-injured patients in tasks in daily life have been directly observed and correlated with EF test batteries. These studies, too, find little or no relationship between impairment in such performances and EF test results (Alderman et al., 2003; Mitchell & Miller, 2008). Here again, EF tests may account for just 9–15% of the variance or even less in ratings of adaptive impairment, primarily in work activities (Ready et al., 2001; Stavro, Ettenhofer, & Nigg, 2007). In contrast, research has noted moderate relationships between EF ratings and measures of daily adaptive functioning in children with various disorders, including TBI (Gilotty, Kenworthy, Sirian, Black, & Wagner, 2002; Mangeot et al., 2002) and in adults with ADHD (Biederman, Petty, Fried, Doyle, et al., 2008), those with frontal lobe disorders (Alderman et al., 2003), and college undergraduates (Ready et al., 2001). And EF ratings substantially outpredict EF tests in the variance shared with measures of impairments in various major life activities such as occupational functioning, educational history, driving, money management, and criminal conduct (Barkley & Fischer, 2011; Barkley & Murphy, 2010, 2011). The totality of findings to date concerning the relationship of EF tests to EF ratings and of each to impairment in daily life indicates that EF tests are largely not sampling the same constructs as are EF ratings or direct evaluations of EF in daily life (Alderman et al., 2003; Shallice & Burgess, 1991). The available evidence also provides a basis for refusing to accept EF tests as the primary or sole source in establishing the nature of EF deficits in various disorders. If assessing how well people do in using EF in their daily life activities is important in clinically evaluating EF, then rating scales assessing EF are superior to EF tests in doing so. Yet the very plethora of studies using tests exclusively to evaluate EF with various patient and general population samples indicates that these serious problems with EF tests have gone largely unknown, unappreciated, or ignored.

This significant failure of EF tests to relate well to EF ratings, daily life activities, or impairment in major domains of life could well indicate that such tests are not assessing EF. This seems doubtful given that many of these tests have been shown to index activities in various regions of the PFC that largely underlie EF. And it is surely unlikely to be the case that EF ratings are not actually evaluating EF. After all, their item content has been drafted directly from definitions of EF and from lists of putative EF constructs in the literature, as well as from observations and clinical descriptions of patients with PFC lesions believed to manifest the "dysexecutive syndrome" (Burgess et al., 1998; Gioia et al., 2000; Kertesz et al., 2000). Moreover, as noted previously, these ratings are substantially related to impairment in various daily life activities and various domains of adaptive functioning, such as work, education, driving, social relationships, self-sufficiency, and so forth, in which EF would surely be operative.

The solution to this paradox of why EF tests and EF ratings are so poorly related lies elsewhere, most likely in the fact that EF is more hierarchically organized than models built entirely on EF tests indicate (Barkley, 2012b). EF tests likely assess the most rudimentary, moment-to-moment, instrumental, and cold cognitive level of EF. But they are very poor at capturing the higher adaptive, tactical, strategic, and principled levels of EF as they are deployed in daily adaptive functioning, human interactions, and socially cooperative and reciprocal activities that play out over much longer spans of time (days, weeks, months, and years)(Barkley, 2012b). It is here, at these higher, more complex, and longer-term levels, that a rating scale of EF can be useful. The reason is that it uses a far longer ascertainment window for capturing summary judgments of behavior over time and is able to capture EF symptoms in their important, largely social, contexts via the reports of the respondent and from others who know that person well.

Also, if the purpose of evaluating EF in patients is to render some judgment as to their likelihood of experiencing impairments in major domains of life activities, then EF ratings are far superior to EF tests. On the other hand, if the purpose of the evaluation is to assess the most proximal neuropsychological EF activities related to moment-to-moment brain activity, as may be important to do in functional brain imaging studies, then EF tests might be preferable. Yet even that point is arguable in view of recent studies linking neuroimaging results to traits that were assessed via rating scales (Buckholtz et al., 2010). Moreover, such tests can be criticized for stripping out important social and cross-temporal elements of EF, which may reduce the validity of the task in sampling any particular EF or its adaptive (evolutionary) purposes. Because of these difficulties, I undertook the development of a rating scale of EF as reflected in daily adaptive behavior. The development of this scale is described in the next chapter.