Psychological Review
2008, Vol. 115, No. 3, 527–543

Copyright 2008 by the American Psychological Association
0033-295X/08/$12.00 DOI: 10.1037/0033-295X.115.3.527

What Is the Shape of Developmental Change?

Karen E. Adolph
New York University

Scott R. Robinson
University of Iowa

Jesse W. Young
Stony Brook University

Felix Gill-Alvarez
New York University

Developmental trajectories provide the empirical foundation for theories about change processes during development. However, the ability to distinguish among alternative trajectories depends on how frequently observations are sampled. This study used real behavioral data, with real patterns of variability, to examine the effects of sampling at different intervals on characterization of the underlying trajectory. Data were derived from a set of 32 infant motor skills indexed daily during the first 18 months. Larger sampling intervals (2–31 days) were simulated by systematically removing observations from the daily data and interpolating over the gaps. Infrequent sampling caused decreasing sensitivity to fluctuations in the daily data: Variable trajectories erroneously appeared as step functions, and estimates of onset ages were increasingly off target. Sensitivity to variation decreased as an inverse power function of sampling interval, resulting in severe degradation of the trajectory with intervals longer than 7 days. These findings suggest that sampling rates typically used by developmental researchers may be inadequate to accurately depict patterns of variability and the shape of developmental change. Inadequate sampling regimes therefore may seriously compromise theories of development.

Keywords: continuity, longitudinal, microgenetic, motor development, stages

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Developmental Trajectories

Understanding developmental change is a central goal for developmental science. However, despite numerous treatises by prominent developmental theorists in a variety of areas urging researchers to focus on change processes (e.g., Elman, 2003; Flavell, 1971; Siegler, 1996; Thelen & Smith, 1994), developmental psychologists have made surprisingly little progress toward understanding the process of developmental change. Part of the problem is historical. Much of the work in developmental psychology has concentrated on descriptions of children's behavior at various ages or on the earliest manifestations of particular abilities. Decades of reliance on cross-sectional designs, demonstration proofs, and broad-sweeping longitudinal approaches have left researchers with a gallery of before and after snapshots, studio portraits of newborns, and fossilized milestones but little understanding of the process of development itself. What we need are accurate, fine-grained depictions of developmental trajectories for cognitive, language, perceptual, motor, and social skills.

The staggering variety of developmental trajectories has also contributed to the lack of progress in understanding change processes. The shape of developmental change might assume any number of patterns (Figure 1). For instance, a trajectory might show smooth and monotonic improvements with age, proceeding at a steady pace, as in children's use of retrieval strategies in addition (Siegler, 1996), or with accelerating or decelerating rates of change, as in infants' acquisition of new words (McMurray, 2007) and improvements in toddlers' walking skill (Adolph, Vereijken, & Shrout, 2003), respectively. The path of change may show discontinuities such as abrupt, stage-like shifts in performance between periods of relative stability, as in children's stage-like success on many Piagetian tasks (Shultz, 1998), their abrupt shift from ignoring to marking the past tense of verbs (Marcus et al., 1992), and the sudden transition to grasping while reaching (Wimmers, Savelsbergh, Beek, & Hopkins, 1998). Variability may increase during the period of acquisition, with a series of reversals vacillating between less and more mature expressions of the skill, as in children's conservation of volume (van der Maas & Molenaar, 1992). Or a variable acquisition period may entail use of multiple, unsystematic use of strategies between incorrect and correct endpoints, as in children's gesture–speech mismatches (Church & Goldin-Meadow, 1986) and their acquisition of a theory of mind (Flynn, 2006). Discontinuities can take on other shapes, such as episodic changes, where development advances like climbing a staircase, with sudden improvements in children's conceptual understanding separated by long periods in a single stage (Case & Okamoto, 1996) or small fits and starts of physical growth sepa-

Karen E. Adolph and Felix Gill-Alvarez, Department of Psychology, New York University; Scott R. Robinson, Department of Psychology, University of Iowa; Jesse W. Young, Interdepartmental Program in Anthropological Sciences, Stony Brook University.

This research was supported by National Institute of Child Health and Human Development Grants R01-33486 to Karen E. Adolph and R01-33862 to Scott R. Robinson. We thank the parents who participated in the study for their painstaking care in filling out the diaries and the New York University students who so meticulously collected and entered the data and helped us to think about the ideas (Ofronama Biu, Pallika Kunani, Jenny Pethkongathan, Toni Webster). We are grateful to Robert Siegler, Peter Molenaar, Martha Alibali, and James Cutting for their useful suggestions for improving the article.

Correspondence concerning this article should be addressed to Karen E. Adolph, Department of Psychology, New York University, 4 Washington Place, Room 410, New York, NY 10003. E-mail: Karen.Adolph@nyu.edu

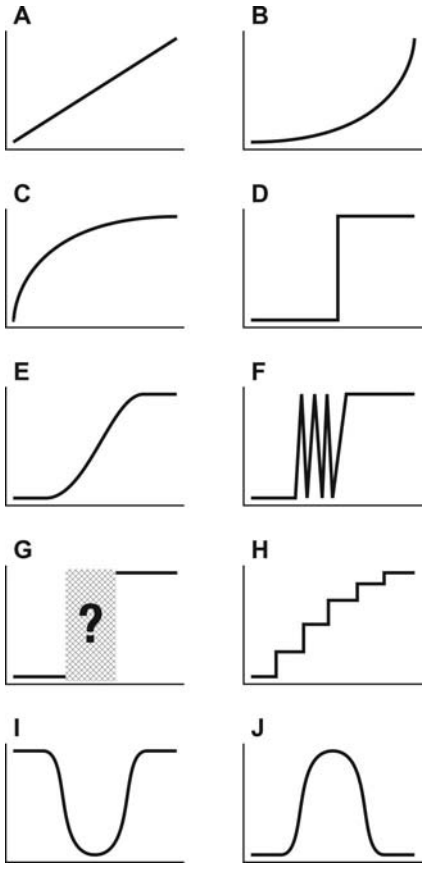

Figure 1. Idealized shapes of developmental change, with age shown on the *x* axis and an index of behavioral expression or level of performance on the *y* axis. A: Linear. B: Accelerating. C: Asymptotic. D: Step-like. E: S-shaped. F: Variable. G: Unsystematic. H: Stair-climbing. I: U-shaped. J: Inverted-U-shaped.

rated by periods of stasis (Lampl, Veldhuis, & Johnson, 1992). Discontinuities can involve reversible patterns of change, as in the U-shaped course of children's success on math equivalence problems (McNeil, 2007), infants' alternating stepping movements (Thelen, 1984), the classic description of overregularizations in past-tense verb forms (Marcus et al., 1992), the inverted-U-shaped trajectory of cognition over the life span (Craik & Bialystok, 2006), and infants' zigzag-shaped error rate in detecting threats to balance as they learn to sit, crawl, cruise, and walk (Adolph, 2005).

Such descriptions of developmental trajectories play an instrumental role in formulating and testing theories of development (Gottlieb, 1976; Siegler, 2006; Smotherman & Robinson, 1995; Wohlwill, 1973). For example, a contentious theoretical debate was spurred by descriptions of a sudden, stage-like increase in children's rate of word learning, the so-called "vocabulary spurt," or "naming explosion" (Bloom, 2004; Ganger & Brent, 2004). According to the classic description, at about 18 months of age, when children have acquired approximately 50 words, they display a sharp transition from an initial stage of slow vocabulary growth to a later stage of faster growth. Several influential theories were advanced to explain the putative shift, invoking major cognitive or linguistic changes that coincided with the spurt (e.g., Gopnik & Meltzoff, 1987; Reznick & Goldfield, 1992). However, recent work shows that for most children, the increase in the rate of word learning is best fit by a quadratic rather than a logistic function (Ganger & Brent, 2004). Without a stage-like spurt in the trajectory, theories positing a sudden, fundamental change in cognitive or linguistic abilities become superfluous.

As illustrated by this example, regardless of whether the theoretical perspective is one of discontinuity or continuity, spurts or quadratics, theoretical accounts of how change occurs are built upon the foundation of an accurate portrayal of the pattern of developmental change (Wohlwill, 1970, 1973). As we demonstrate in this article, an accurate characterization of the developmental trajectory depends on the rate at which observations are sampled.

The Problem of Sampling Rate

More than 75 years ago, Vygotsky (1978) criticized researchers' reliance on sampling methods that merely characterize the stable endpoints in cognitive development. As a remedy, he proposed a "microgenetic method" of sampling at small time intervals to observe development in progress. More recent researchers also have cautioned against overreliance on cross-sectional and long-term longitudinal designs (Wohlwill, 1970, 1973) and have espoused the microgenetic method for capturing the process of developmental change (e.g., Granott & Parziale, 2002; Kuhn, 1995; Siegler, 2006; Thelen & Ulrich, 1991).

However, apart from the general criticism that researchers' typical sampling intervals are too large, the microgenetic method does not quantify the potential consequences of various rates of data collection for detecting and characterizing different patterns of development. Proponents of the microgenetic method have offered general suggestions that researchers should collect observations spanning the entire period of change from one stable state to another and that the frequency of observations should be high relative to the rate of change of the phenomenon (Siegler, 2006). But these proponents have not addressed the problem of how to decide whether a sampling interval is small enough to detect the shape of the underlying trajectory. That is, when does one stable state end and another begin? Is the development step-like, or is there an intervening period of variability, partial or intermittent expression, or disruption of performance? Similarly, critics of developmental methodology have recognized that overly large sampling intervals in longitudinal research can cause important patterns of change to go undetected and have suggested that developmental researchers sample at smaller intervals (Burchinal & Appelbaum, 1991; Collins, 2006; Hertzog & Nesselroade, 2003; McArdle & Epstein, 1987). But how small is small enough?

In fields of inquiry such as physiology, psychobiology, health psychology, and neuroscience, principles are available to guide the selection of an appropriate sampling rate to ensure recovery of the underlying pattern. For instance, the Nyquist–Shannon sampling theorem (Nyquist, 1928/2002; Shannon, 1949/1998) provides an algorithm for calculating the minimum sampling rate to fully characterize complex waveforms. The sampling theorem stipulates that for a waveform composed of one or more frequencies, with a maximum relevant bandwidth (B), the minimum sampling frequency (f_s) necessary to reconstruct the original waveform must be at least twice the bandwidth ($f_s > 2B$). In other words, sampling frequency must be at least twice as frequent as the highest frequency component. For example, recording sounds at 20 kHz, the upper limit

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

for human auditory perception, would require sampling the waveform at a minimum of 40 kHz (which is one reason why mp3 digital sound files have such poor quality for higher frequency sounds). Assumptions about the maximum relevant bandwidth are dictated by the nature of the research question. A study of human color discrimination would not require light wavelengths to be sampled beyond the blue end of the visible spectrum.

Ironically, the same developmental psychologists who scrupulously use principles such as the Nyquist–Shannon theorem to select sampling rates to estimate functions for physiological and psychophysical variables rely on intuition, convenience, and tradition to select sampling intervals to characterize developmental change in said functions. For example, to describe age-related changes in the event-related brain potential associated with face and object recognition, Webb, Long, and Nelson (2005) sampled the electroencephalogram at 100 Hz to ensure that they could characterize specific components of the electroencephalogram response distributed during the first 1,500 ms after presentation of the stimulus. But they relied on arbitrary 2-month intervals to chart the developmental trajectory of the event-related brain potential signals. To describe the development of stereoacuity in infants, Held, Birch, and Gwiazda (1980) estimated the psychophysical functions by ensuring a sufficiently high sampling rate to distribute intervals of visual angle along the inflection of the curve. However, they relied on an arbitrary, 1-month sampling interval to estimate infants' developmental trajectories and onset ages. Similarly, Adolph (1997) described developmental changes in infants' perception of affordances for crawling and walking by sampling at sufficiently small intervals of difficulty to ensure robust estimates of the psychophysical functions, relying on an arbitrary 3-week sampling interval to estimate the developmental trajectories.

A recommended remedy for researchers' sampling dilemma is to design the spacing of observations on the basis of a formal theoretical model about the shape of the underlying developmental function (Boker & Nesselroade, 2002; Burchinal & Appelbaum, 1991). Such a model would dictate the minimum number of data points and their optimal spacing in time (e.g., a linear function requires only two observations at each end of the acquisition period). However, formal rules such as the Nyquist–Shannon theorem are applicable only when the data consist of complex waveforms and the maximum frequency of interest is known in advance. If the temporal scale of developmental changes also were known in advance, then applying a formal rule like the Nyquist–Shannon might be possible (e.g., sample at twice the frequency of the smallest significant change). Unfortunately, most developmental data are not periodic and are not generated by simple mathematical functions, where the relevant scale of temporal change can be obtained by deduction. Thus, developmental researchers must determine patterns of developmental change empirically and discover, rather than deduce, the temporal scale of events that make a difference in the process of change.

The problem is compounded because, as Collins and Graham (2002) pointed out, empirically derived sampling intervals lead to a "chicken and egg" situation: Without prior knowledge about the shape of the underlying trajectory to inform a statistical function, researchers cannot know how frequently to space their observations; and the underlying function that determines the shape of the developmental trajectory cannot be discovered empirically without making a decision about sampling interval. Often, researchers do not even have prior information about the approximate ages that span the period of developmental change.

Implications for the Shape of Change

Few examples of developmental research have systematically assessed the empirical costs and benefits of large and small sampling intervals on descriptions of developmental change. A notable exception is the research of Lampl and colleagues on patterns of physical growth (Johnson, Veldhuis, & Lampl, 1996; Lampl, Johnson, & Frongillo, 2001; Lampl et al., 1992). Traditionally, children's growth is characterized as a continuous function from birth to adulthood, with more rapid growth rates during infancy and adolescence. However, when children's height is measured every day, growth appears to be episodic. Infants' height, for example, can increase 1.65 cm in the course of a single day, separated by long periods of days or weeks during which no growth occurs. Sampling at weekly intervals results in developmental trajectories that preserve the episodic nature of children's growth but reduce the observed number of growth spurts, increase the amplitude of the spurts, and prolong the periods of stasis. Sampling at quarterly or yearly intervals, as in traditional studies of growth, results in the smooth, continuous growth curves on standard growth charts.

Even within a 24-hr period, growth is not continuous. In a tour de force of micromeasurement, Noonan et al. (2004) demonstrated episodic growth on two time scales: brief periods of substantial growth on a scale of minutes and days, flanked by long periods of no growth on an hourly and weekly scale. Leg growth in freely moving lambs was measured with a microtransducer surgically implanted across the tibial growth plate. Bone length was sampled at 167-s intervals over a period of 3 weeks, synchronized with video recordings of the lambs' activity. Periods of bone growth revealed by the microtransducer coincided with periods of recumbency revealed by the video recordings, and periods when bones did not grow coincided with periods of loading the limbs in stance or locomotion. The authors calculated that 90% of bone growth occurs while lying down, even though lambs spend just over 50% of their time in a recumbent position, and little or no growth occurs while standing or walking. Clearly, tradition, intuition, and convenience that informed traditional studies of physical growth have been inadequate for capturing the richness of the actual trajectory.

The case of physical growth shows how increased sampling resolution from years to days to minutes can provide novel insights into the developmental process. The episodic growth pattern from minute to minute indicates that bones lengthen only when compressive forces on the leg are absent. Paradoxically, other research has demonstrated that the presence of physical forces applied to bone promote growth by stimulating the expression of genes that regulate cartilage and bone formation (Muller, 2003). Together, these research findings imply that cellular processes involved in regulating physical growth must be coordinated and synchronized on a previously unsuspected temporal scale.

As exemplified by the research on physical growth, overly large sampling intervals cause interval data to appear smooth and continuous, regardless of whether the underlying trajectory is episodic or U-shaped. Similarly, overly large sampling intervals distort the shape of change for binary data (skills that are indexed as absent or present). Figure 2 shows the potential impact of sampling at monthly intervals on characterizing patterns of development using actual data from

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

daily observations of two infants' progress in balancing upright. Figure 2A shows a step function, where the infant exhibited a single transition from not standing to standing, from 1 day to the next. Figure 2B shows a variable developmental function, where standing was expressed intermittently (21 times) over a protracted transition period of several weeks. For skills with variable trajectories and reversals, interpolating over the existing data points—which is what all developmental researchers do when measurements are collected at weekly, monthly, and yearly intervals—can distort the shape of the developmental trajectory. Infrequent observations cause binary data to appear as a step function, with a single abrupt transition, regardless of whether the underlying trajectory is variable, with a series of reversals. As illustrated by the gray curves in the figure, the variable data in Figure 2B appear to follow the same developmental path as the stage-like data in Figure 2A.

Implications for the Timing of Change

Overly large sampling intervals are also likely to produce errors in estimating onset ages: the earliest age at which children consistently and reliably express a behavior, skill, or physiological milestone. Identification of onset ages plays a prominent role in normative and clinical studies of human development, screening for developmental delay, and experimental manipulations of development in animals. The onset ages of cognitive and motor milestones are commonly used to document developmental delays in clinical populations, such as the delay in autistic and deaf children's acquisition of theory of mind (Peterson & Siegal, 1999). Researchers use age at onset to compare the development of different skills, such as language comprehension and production (Clark & Hecht, 1983), or to compare the development of the same skill expressed in different contexts, such as the age of attaining conservation of quantities in different cultures (Dasen, 1984) or the age of reaching for objects in the light and in the dark (Clifton, Muir, Ashmead, & Clarkson, 1993). Researchers use onset ages to assess effects of prior experiences on the development of a target skill, such as interactions with siblings on acquiring a theory of mind (Perner, Ruffman, & Leekam, 1994), experience with pottery making on the onset of conservation (Price-Williams, Gordon, & Ramirez, 1969), or the effect of sleeping prone versus supine on the subsequent development of crawling (Majnemer & Barr, 2005). Measures of experience in human infants typically are calculated as the number of days between onset and test dates, for assessing effects of crawling experience, for example, on improvements in perceptual, cognitive, and social tasks (Campos et al., 2000).

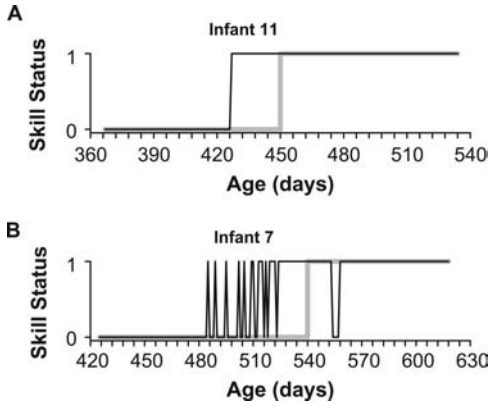

Figure 2. Examples of developmental trajectories derived from daily data (black curves) for standing (balancing upright for ≥3 s without holding a support) in 2 infants. A: Trajectory that exhibits abrupt step function from absent to present from 1 day to the next. Simulated monthly sampling (gray curve) results in an error in identifying the skill onset age but does not distort the shape of the trajectory. B: Variable trajectory, where skill vacillated 21 times between absent and present over the course of several weeks. Simulated monthly sampling (gray curve) misrepresents both the shape of the variable trajectory and the estimated onset age.

It is easy to imagine how sampling at longer intervals results in reduced accuracy in estimating the onset age of skills that exhibit abrupt, step-like transitions (e.g., monthly sampling risks 1-month delays in estimates of onset ages; see Figure 2A). But it is less intuitive how the choice of sampling interval affects the accuracy of estimating onset ages in skills with variable developmental trajectories. As shown in Figure 2B, infrequent sampling is likely to miss the period of variability and thereby provide a later estimate of the onset age. Occasionally, the observations fall on a day when the skill is present but not yet stable and, thus, distort the estimate of onset by providing a prematurely early estimate.

As we have argued in the foregoing account, the rate at which behavior is sampled is likely to have a significant impact on our ability to discern the shape and timing of developmental change. Sampling at inappropriately large intervals can yield an erroneous picture of the underlying developmental trajectory, which in turn may provide misleading inferences for developmental theory. But the real cost is not just in misrepresenting the pattern of change. It is the loss of the ability to distinguish among alternative trajectories, such as the ones depicted in Figure 1. An important principle of empirical science is that theories and hypotheses must be falsifiable. It should be possible in principle to obtain some set of measurements that would not accord with the theory (Popper, 1959). Inferences about particular developmental trajectories are not falsifiable unless the data could have revealed alternative patterns of change. Confidence in the shape of a developmental trajectory depends on whether the data were sampled at appropriate intervals to permit the possible detection of alternative paths. Because there are no generally accepted rules or theorems to guide selection of a sampling interval in a particular developmental context, appropriate sampling intervals must be determined empirically.

Current Study

In the present study, we aimed to meet the challenge of the microgenetic method by establishing empirically whether the different sampling rates typically used by developmental psychologists in microgenetic and longitudinal research—days, weeks, and months—are sufficient to accurately characterize the pattern of developmental change. Our aims were fourfold. First, we sought to demonstrate that real data with real patterns of variability could yield dramatically different trajectories when sampled at rates commonly used in developmental research. Second, we aimed to quantify how quickly researchers lose the picture of developmental change when sampling at increasingly large intervals. It is a mathematical certainty that coarser sampling is less sensitive to fluctuations in the data, but it is not clear at what rate researchers incur the cost of misrepresenting the underlying trajectory. Third, we assessed the consequence of different sampling intervals for

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

estimating onset ages: the earliest manifestation of stable expression of skills and abilities. Fourth, we tested whether the effects of sampling interval generalize across children, the first 18 months of life, and a range of different skills.

Specifically, in this study, we measured the impact of collecting developmental data at intervals of varying length on loss of sensitivity to detect the underlying trajectory. To ensure that natural patterns of variability would be included in the data, we compiled a real data set of daily changes in 32 infant motor skills (sitting, crawling, standing, walking, etc.) obtained from parent checklist diaries, rather than an artificial data set of experimenter-generated data. We focused on motor skills because motor performance is overt and amenable to objective, reliable measurement; new motor skills are highly salient to parents; and motor development has a long history of longitudinal and microgenetic research. However, in principle, the data set could have been constructed from any skills appearing at any point in the lifespan, indexed in terms of competence, rather than performance, and obtained in the laboratory or during home visits, rather than by parents' reports.

Following in the long tradition of language studies (e.g., Darwin, 1877/1974; Dromi, 1987), parents served as informants by noting the presence or absence of each skill at the end of the day in a checklist diary. Although readers' first inclination may be skepticism regarding parental reports, home observations integrated over the course of the day may be the best way to determine if a skill is in a child's repertoire, because parents are with their children in many different situations, including contexts that are likely to elicit and support the emergence of new skills (Bodnarchuk & Eaton, 2004). For language skills, laboratory tests and experimenter home visits grossly underestimate children's early abilities, necessitating parental reports to avoid false negatives (Bates, 1993). For motor skills, parent checklist diaries of basic motor skills are concordant with experimenter home visits (Bodnarchuk & Eaton, 2004).

As is customary in the literature, we treated the appearance and disappearance of motor skills as binary, categorical data (present or absent). Like researchers in other developmental domains that treat skills categorically (e.g., object permanence, conservation, and theory of mind), we established operational definitions for the performance of each skill. For several skills, we included multiple criteria for successful performance (e.g., walking <3 m and walking >3 m) to determine whether more stringent criteria would affect the trajectory.

From the daily assessments, we constructed developmental trajectories for each skill at the finest available grain of temporal resolution. Then we systematically removed observations to simulate the effects of sampling at intervals ranging from daily to monthly and reconstructed the developmental trajectories on the basis of the reduced number of observations. We compared key features of the resulting trajectories with the original data to determine the loss of sensitivity for detecting various patterns of developmental change that result from different sampling schedules. In addition, we formulated a method on the basis of a neurally inspired activation function for objectively estimating onset ages for each skill. We compared the estimated onset ages derived from the original daily observations with those derived from the simulations of larger sampling intervals to determine the magnitude of error that could be attributed to sampling frequency.

Method

Checklist Diary

We compiled a database of daily diary data from 11 families with infants (5 boys, 6 girls). Nine infants were Caucasian and 2 were Asian. All parents were middle class and highly educated. Eight infants had parents who were doctoral students or professors in psychology or anthropology, including the daughter of Karen E. Adolph, and thus, most respondents were experienced in methods of behavioral data collection. Parents began keeping diary records before their infants could perform any of the target skills and ended participation several weeks after their infants could walk independently. One family stopped participation abruptly when the infant was 9 months old because of a medical emergency. For the other 10 infants, length of participation ranged from 10.94 to 17.00 months ($M = 12.59$ months). One additional family ceased participation after only 3 months because the parents found it to be too grueling; data from this infant were not included in the database.

Parents were trained to make daily entries into a three-page, paper-and-pencil, checklist diary containing 32 gross motor skills involving balance and locomotion, all of which could be performed in a minimally structured environment (i.e., with a floor and furniture). Instruction manuals accompanied parents' diaries with detailed descriptions of the criteria for each skill (see the Appendix) and a reminder of how to fill out the diary. The diaries were similar to those used by Bodnarchuk and Eaton (2004), who showed that parents' reports were concordant with home-visit observations. Data were collected for 22 additional stair climbing and sliding skills, but these were not included in the current study because they required access to special equipment not readily available on a daily basis.

Parents noted whether they had observed infants performing each skill at any point over the course of the day. The diaries provided space for additional written comments about observed skills that did not quite match criteria. Such comments about the first 2 participants—Karen E. Adolph's daughter and the son of another psychology professor—provided useful information for revising skill definitions and criteria. Only skills with uniform definitions and criteria were included in the final data set. Parents entered a question mark for days when they could not remember whether they had witnessed the skill or if they had forgotten to fill in their diaries. Parents also noted days when infants did not have normal access to the floor or to furniture (because of long car trips, camping trips, infants' illness, etc.) and, thus, were precluded from performing various skills because of situational factors.

We distributed diaries to parents each month and organized the diaries to minimize errors in parents' reports. Skills were grouped roughly by postural systems and order of appearance (sitting, prone/crawling, standing/cruising/walking). The first page of the diary contained sitting and early prone skills, the second page contained crawling and upright skills, and the third page contained stair climbing and sliding skills. More stringent criteria for specific skills (e.g., walking >3 m) followed more lenient criteria (e.g., walking <3 m). Some of the skills in our data set were ordered hierarchically, where demonstrated facility for a stricter criterion necessarily assumed facility under a more lenient criterion (noted by asterisks in the Appendix). For example, once a child can consistently walk 3 m or more, it is not necessary to also record walking fewer than 3 m. Therefore, after infants demonstrated

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

facility for at least 30 consecutive days with the stricter criterion, the entries for the lenient criterion were assumed to be present.

During monthly lab visits, researchers collected parents' completed diaries from the previous month, interviewed parents about diary entries (confirmed infants' expression of new skills, cessation of old skills, and question-mark and no-access days) and distributed a new diary for the current month. The interviewer reminded parents about the criteria for the various skills by using verbal descriptions, giving physical demonstrations of the behaviors, and directing them to the relevant definitions in the instruction manual.

Missing Data

Because our aim was to assess effects of sampling interval on characterization of the underlying developmental trajectories, it was especially important to maintain high confidence in the integrity of the time series. Days that parents noted with question marks and days in which infants had no access to the floor constituted missing data. Given that the aim of the study was to detect variability, we adopted a conservative strategy for interpolating over missing data. For each skill, a software program written in our laboratory searched for the first instance of existing data prior to the day for which data were missing and replaced the missing data entry with that notation. The assumption underlying the interpolation rule was that infants were likely to have continued doing what they last did until otherwise noted. At most, two consecutive days of missing data were reconstructed in this way. If a skill contained more than two consecutive days of missing data or if missing data constituted 5% or more of all entries, the time series was not used for further analyses.

Overall, each infant contributed 4–30 skills ($M = 23.73$ skills), for a total of 99,971 usable diary entries in the final data set across infants and skills. Several factors caused the large range in the number of skills that each infant contributed. For the first 2 infants in the sample, we revised the definitions and criteria for several skills and thus eliminated several time series collected under earlier definitions and criteria. For other infants, some of the time series included more than 5% missing data because of days noted with question marks, days when infants did not have access to the floor, and in the case of one infant, a lost month of entries. Across the sample, some infants never performed certain skills (e.g., never crawled >3 m). Finally, several time series were either cut short or were not performed by the infant who withdrew from the study because of a medical emergency.

Manipulation of Sampling Frequency

The critical tests involved varying sampling frequency, then interpolating over the intervening points. The actual daily data entries provided the smallest sampling interval. We wrote software to simulate the effect of sampling at longer intervals by systematically selecting observation points at 2- to 31-day intervals for each skill reported by each parent in the data set. For example, to simulate a 2-day sampling interval, the program selected every second data point; to simulate a 3-day sampling interval, the program selected every third data point. After resampling, removed days were replaced with interpolated values. The process continued for each of the remaining sampling intervals until the least frequent sample at 31-day intervals. Therefore, every simulated time series had the same number of days as the original.

When observation points are distributed in time, the specific day that each sample is collected can vary (e.g., sampling at a 2-day interval could be initiated on the first available day that measurements were collected and on all odd numbered days thereafter, or the sample could begin on the second day of data collection and proceed on all even numbered days). Failure to take phase into account would allow for random sampling effects to influence the overall trajectory, particularly if performance of the skill is variable. For example, the singular occurrence of a skill on Day 31, but not on surrounding days, would appear as a stage-like transition a month earlier if sampling at 30-day intervals beginning on Day 1 (i.e., 1, 31, 61, 91. . .) compared with the same rate of sampling beginning on Day 2 (i.e., 2, 32, 62, 92. . .). To allow for random variation due to the phase of sampling, we made the final data set exhaustive and included all possible phases at each sampling interval (e.g., 30 phase sets were created for each skill when sampling was conducted at 30-day intervals).

The resulting final data set included the original data collected daily and data sets resulting from sampling at 2- to 31-day intervals at all possible phases. After sampling each simulated series of observations, the software program interpolated over missing values by filling in daily values based on the last available observation point. Although it would have been possible to use an alternative rule, such as retroactively filling in missing data according to the next available data point, we adopted a conservative assumption that a binary function continues on the same trajectory until a demonstrated instance of a change. Because each of the original time series resulted in 495 additional sampled time series, the original data set of 261 (Infant × Skill) time series yielded a final data set of 129,456 unique time series.

Results

Effects of Sampling Interval on Observed Trajectories

We assessed the effect of variations in sampling interval on the shape of the observed trajectory by counting the number of transitions between absent and present for each time series. A single transition would represent an abrupt step-like trajectory from absent to present, as exemplified by Infant 11, who began standing on one day and stood every day thereafter (see Figure 2A, which is also depicted as the data point nearest the origin in Figure 3A). Alternatively, multiple transitions would represent a variable trajectory between absent and present, as exemplified by Infant 7, who vacillated 21 times between standing and not standing (see Figure 2B and top data point in leftmost panel of Figure 3A).

Of the 261 time series in the data set, only 15.7% showed single abrupt transitions (either onsets alone or a single onset and offset) at a 1-day sampling interval. For the remaining 84.3% of time series, the daily diary data showed variable trajectories, ranging from 3 to 72 transitions during the acquisition period ($M = 13.37$ for those time series showing variable trajectories). Inspection of all time series revealed that variable trajectories were characteristic of all infants and skills. Between 65% and 100% of the time series for each infant showed multiple transitions, regardless of sex. Similarly, between 67% and 100% of the time series for each skill (expressed by at least 2 infants) showed multiple transitions, regardless of the kind of skill, the strictness of the criterion for judging skill occurrence, or the average age at which the skill was expressed.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

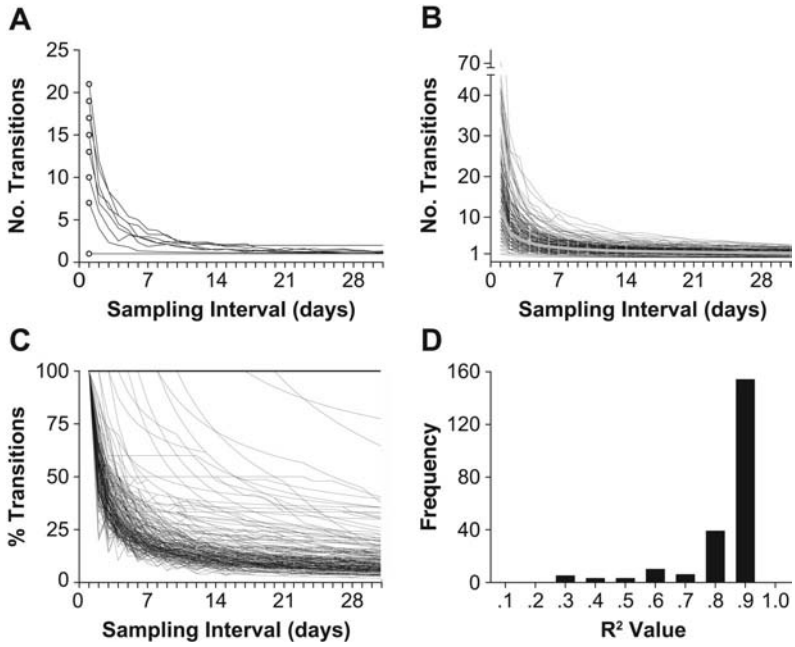

Figure 3. Effects of sampling interval on sensitivity to variability in developmental trajectories. A: The number of observed transitions between absence and presence for one skill (standing). Each curve represents data for 1 of the 8 infants for whom we had a complete time series. Open symbols depict data when the skill was sampled daily; lines show data averaged across all possible phases at each of the 1- to 31-day sampling intervals. Note that the data point nearest the origin represents the stage-like data from Infant 11, shown in Figure 2A. The other seven data points show data for variable trajectories from other infants, including the top data point depicting Infant 7, shown in Figure 2B. B: Number of observed transitions, presented as in Figure 3A, for all 32 skills. The thick gray line represents the mean trajectory across all 261 time series. C: The same data presented in Figure 3B expressed as a percentage of observed transitions recorded at daily intervals. The horizontal line at 100% represents the 41 time series with only one abrupt transition from absent to present (15.7% of all time series). Most time series consisted of variable trajectories when measured daily, but more than 75% of transitions were not detected when sampled at weekly intervals. D: Distribution of R^2 values for inverse power functions fit to each of the 240 time series with multiple transitions. Most time series were best described by an inverse power function, indicating that modest increase in small sampling intervals (<1 week) resulted in a sharp decline in the ability to detect transitions.

The consequence of larger sampling intervals was to obscure the true shape of the developmental trajectory. Across the 32 motor skills in the data set, sampling at the simulated rate of once per week caused 51.2% of the 220 variable time series to show a single transition from absent to present. At the simulated rate of once per month, 91.4% of the variable time series appeared to involve a single, abrupt transition. Overall, with monthly sampling, 242 (92.7%) of the 261 time series appeared to follow step-like trajectories (compared with the 15.7% based on daily samples), yielding a very different picture of developmental change from that of the daily data.

How quickly did we lose the picture of developmental change? Given fewer observations at larger sampling intervals, one would expect a general loss of sensitivity to detect variability. In fact, sampling at progressively larger intervals carried a tremendous cost: Sensitivity to detect variability in the time series declined dramatically each time we widened the sampling interval by 1 day. Figure 3A illustrates this precipitous drop-off in sensitivity for each child for one skill, standing (also represented in Figure 2). Of the 8 children depicted in the graph, only 1 (Infant 11) exhibited a step-like transition from absence to presence when standing was indexed daily. By the time sampling approached 14-day intervals, however, the child with 21 transitions (Infant 7) was indistinguishable from the child with a single transition.

The black curves in Figure 3B show that the dramatic decrease in sensitivity was evident for all of the 32 skills in the data set. The gray curve in the figure shows the group average across sampling intervals. Although infants averaged 11.47 transitions (SD = 10.48) in their actual daily diaries across all 32 skills, sampling once per week yielded only 2.51 transitions (SD = 2.10), on average, and sampling once per month yielded an average of only 1.20 transitions (SD = 0.83). The drop-off in sensitivity is more evident in Figure 3C, which depicts these same data expressed as a percentage of the number of transitions observed with daily sampling. As shown by the concentration of trajectories in the lower left of the figure, for most time series, fewer than one in four transitions (25%) were detected when sampling at larger than 1-week intervals. Moreover, time series with frequent transitions were disproportionately mischaracterized. The only trajectories that were depicted accurately at larger sampling intervals were the 41 time series (15.7% of all time series) with only one abrupt transition from absent to present, shown by the superimposed horizontal lines at 100%.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Each day that the sampling interval widened resulted in fewer transitions detected. To quantify how quickly sensitivity to variability was lost, we fit a variety of mathematical functions to the data shown in Figure 3B. The loss of sensitivity to detect transitions was best described by an inverse power function, meaning that the rate of loss of sensitivity was greatest at the smallest sampling intervals and declined as intervals grew larger. As shown in Figure 3D, most of the R^2 values exceeded .8 for power functions fit through the data for each of the 240 time series with multiple transitions at each of the 31 possible phases.

Effects of Sampling Interval on Estimated Onset Ages

Developmental researchers rely on onset age—the earliest date at which children can consistently and reliably express a particular motor or cognitive skill—as a primary index of developmental progress. As the foregoing discussion of sampling intervals suggests, measuring developmental change at long intervals is likely to result in greater error in identifying the onset of skill performance than is measuring at shorter intervals. We sought to quantify the expected magnitude of error in estimating the onset ages by calculating the deviation between the date determined by a particular sampling frequency and the date determined by daily sampling.

When we sampled at 31-day intervals, each unique phase set provided a separate estimate of the onset age and, thus, a distribution of 31 different estimates of the error of measuring onset age relative to daily samples. In addition to phase differences, if onset is determined by a criterion other than first day of expression, the specific pattern of days in which a skill occurs within a variable trajectory can influence the age identified as the onset of stable performance. Because most time series were variable, we sought to vary the particular sequence of days in which the skill was expressed to obtain a larger and more robust set of time series. By creating variants through a constrained randomization procedure, we provided more time series for analysis at all sampling intervals, including short intervals that provided fewer estimates of onset age (e.g., sampling on alternate days provides only two estimates, one for each phase). In applying a randomization procedure, however, it was important to constrain the procedure to local sequences within the time series, thereby leaving the overall arc of the trajectory the same.

We used a Monte Carlo randomization procedure to introduce slight variations in the dates on which skills were expressed. In a typical randomization procedure, the sequence of events in the entire time series is shuffled, producing a random reordering of the data set (e.g., Johnson et al., 1996; Kleven, Lane, & Robinson, 2004). Clearly, if the sequence of events in the original time series were completely randomized, no developmental pattern could be discerned. To preserve the overall developmental pattern while creating random variations in the daily events, we constrained randomization to restrict the temporal range within which shuffling occurred. A similar procedure was applied by Loreau (1989), who constrained randomization on a seasonal basis to maintain biological realism in a model of annual activity cycles and ecological competition.

To implement our randomization procedure, we parsed each binary time series, after simulated sampling and interpolation, into a sequence of bins comprising 14 consecutive days. We selected the size of this bin (14 days) after exploring alternative bin widths, and we chose it to provide a diversity of permutations while introducing minimal error in the overall developmental profiles. Within each bin, daily events were randomly resampled without replacement, creating a sequential permutation of the original bin (Crowley, 1992). Although the specific dates of occurrence were reordered within bins throughout the time series, the sequence of 14-day bins was not modified. Thus, for a time series of daily samples spanning a year, there would be twenty-six 14-day bins and therefore $(14!)^{26}$ possible permutations. We selected 25 randomly generated time series from this set of possible permutations for each unique phase set for further analysis. This approach resulted in the creation of many alternative time series that differed in the specific dates that skills were expressed but preserved the same general developmental trajectory.

The 129,456 Density $\times$ Phase combinations and 25 randomization procedures applied at each simulated sampling interval resulted in a total of 3,236,400 time series of skill performance. For each of these time series, we applied an objective algorithm to identify the onset age based on the earliest age at which the skill was consistently and reliably performed. Determination of the onset age is straightforward when the underlying developmental trajectory is a step function, because skill performance exhibits a single transition from absence to presence in the infant's repertoire (see Figure 2A). However, objectively defining skill onset is more problematic when the skill is performed on one day and not on the next (see Figure 2B).

In determining onset age, one might simply report the first day on which the skill was observed. In some developmental research, however, the first date of observation is not used as the criterion for onset because a singular performance followed by weeks of no expression may be interpreted as anomalous or unrepresentative of a stable ability. Other criteria (e.g., skill must be expressed on 3 consecutive days) are also arbitrary and seem to lead to exceptions and additional criteria requiring qualitative inspection of each time series. In lieu of these options, and to provide an automated method of determining the onset of stable performance of each skill that could be applied to 3 million time series, we applied an objective algorithm to summate over variable periods of skill expression until a criterion level of skill performance was reached.

The algorithm we used to objectively determine the onset age consisted of an activation function that summated across consecutive days of variable skill performance. Although it was inspired by the rules of summation that generate action potentials in a neuron, this function did not involve processing of the data with an artificial neural network but acted more generally as a smoothing function over periods of variable expression. The critical parameters of the activation function were a decay rate (d), a criterion onset threshold (T_{on}), and a criterion inactivity threshold (T_{off}). Activity accumulated or decayed in the function over successive days following Equation 1, where A_t is the accumulated activity at time t, E is the value of the event at time t (1 = *skill present*, 0 = *skill absent*), and d is decay rate, which specifies the amount of activity that carries over from one day to the next:

$$A_t = E_t + dA_{t-1}. \quad (1)$$

With this simple smoothing rule, each day that a skill was performed added activity to the function (much like a small synaptic potential contributes to the net depolarization of a neuron), but activity decayed from one day to the next. When the skill was performed over consecutive days, the function approximated a logarithmic function, and activity summated toward an asymptote that represented consistent and reliable performance. Over a span

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

of days when the skill was not performed, activity decayed toward zero as a negative exponential function. When the cumulative activity, as determined by the particular pattern of skill expression over successive days, exceeded the criterion-onset threshold (T_{on}), the skill was considered stable, and the onset age was determined by tracing the rising slope of activity back to the preceding minimum below the inactivity threshold (T_{off}; see Figure 4A). In practice, this algorithm identified the first day a skill was expressed in cases in which there was a single step-like transition from one day to the next, and it consistently identified a date between the first day a skill was expressed and the asymptote in trajectories with periods of variable expression.

We systematically explored the effects of varying different parameters in this function with a subset of the data to maximize the number of time series for which an objective onset age could be determined. To confirm the validity of this function, all four authors visually examined representative graphs of the time series to identify an age by consensus for the onset of stable and consistent performance. The subset of time series included skills that exhibited sudden onset from one day to the next and skills that showed protracted periods of intermittent expression before skills were consistently expressed. We then adjusted parameters of the activation function to identify the same ages in the exemplar trajectories. For the results reported below, we used a decay rate of 0.8, an upper onset threshold of 75% of asymptote, and a lower inactivity threshold of 10% of asymptote as optimal for identifying onset ages across all types of trajectories. With these settings, we identified onset ages for 3,045,764 time series (94.1%). In most instances, failure to identify an onset age by these objective criteria was due to the infrequent expression of the skill (on 5 or fewer days) in the time series (and, thus, insufficient activity accumulated to exceed the onset threshold).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

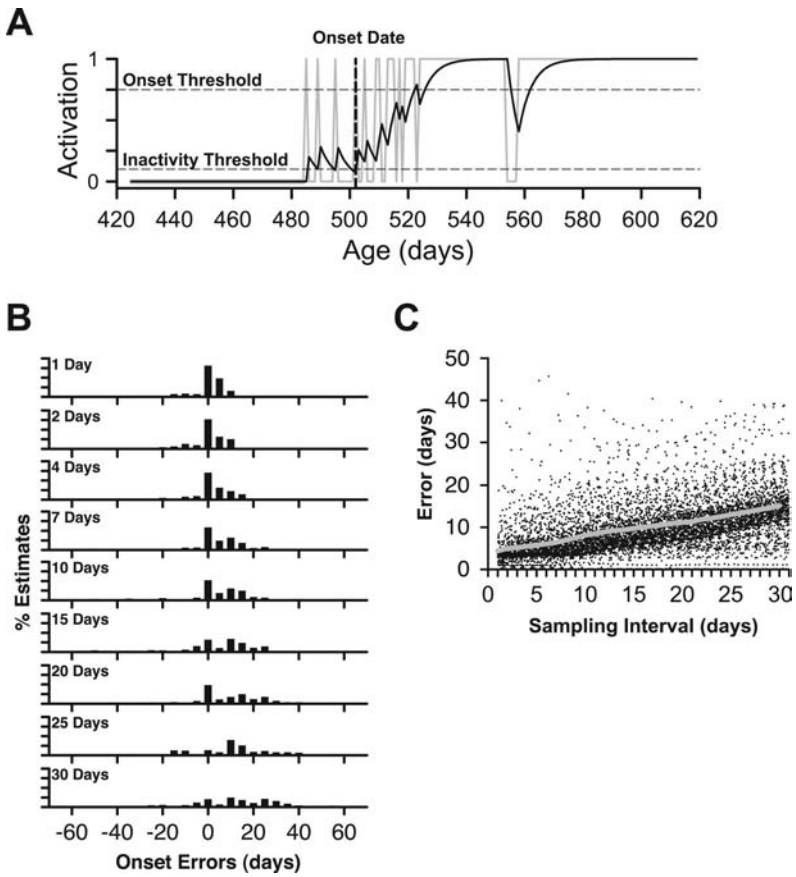

Figure 4. Effects of sampling interval on estimates of onset ages. A: Neurally inspired activation function and resulting estimate of the onset age applied to the daily data shown in Figure 1B for standing in Infant 7. The onset age is determined by identifying the first instance of activity that exceeds a criterion threshold, then tracing the function back to the preceding period of inactivity. In this case, the function identifies an onset age at 501 days (shown as vertical dashed line). B: Histograms showing errors in estimates of the onset age for one skill, standing, in all 8 of the infants for whom time series were available. The *y* axis is expressed as a percentage of total estimates. Note that larger sampling intervals result in a greater range of errors, a general increase in the magnitude of errors, and a tendency for errors to be shifted toward later ages. C: Number of days that estimates of onset ages deviated, either earlier or later, from estimates derived from daily sampling. Data are presented for all available skills for each child (261 time series) as a function of the sampling interval; the superimposed gray line shows the mean absolute error resulting from sampling at different intervals.

For each child and each skill, we used the activation function to identify an onset age from the original daily diary data. Then, we compared the original onset ages with estimated onset ages for all of the other time series generated by the randomization procedure at each of the simulated sampling intervals and phases. Figure 4B shows a series of histograms charting the distributions of error estimates for one representative skill, standing, in all 8 of the infants for whom we had useable data. As revealed by reading down the column of histograms, the magnitude of error increased systematically with larger sampling intervals. As sampling interval increased, the distributions progressively shifted to the right, reflecting delays in identifying onset. At the most extreme, the estimated onset age was delayed by 55 days.

The pattern of increasing error exemplified by standing was characteristic across the entire data set (Figure 4C). With daily sampling, the magnitude of error introduced by the Monte Carlo randomization procedure averaged 4.31 days ($SD = 1.97$) across infants and skills. In other words, constraining our randomization of skill sequences within 14-day bins resulted in relatively small variations in onset age. However, with progressively longer sampling intervals, the average magnitude and range of error in estimating onset ages increased sharply. For example, sampling at weekly intervals resulted in a mean absolute error of 6.31 days ($SD = 4.43$), and absolute errors >14 days occurred in 7.7% of estimated onset ages. (Errors larger than about 14 days can seriously compromise theorizing about motor skills.) Sampling at 20-day intervals resulted in a mean absolute error of 11.06 days ($SD = 7.74$), and absolute errors >14 days in 21.4% of estimates. At a 30-day sampling interval, the mean absolute error compared with daily samples was 15.06 days ($SD = 9.86$), and absolute errors >14 days constituted 37.5% of estimates. At the most extreme, the estimate of onset age differed from the actual onset age calculated from daily sampling by 109 days.

Moreover, errors were not distributed symmetrically around the daily estimates of skill onset; most errors were greater than zero, indicating a delayed estimate of the onset age. Sampling at longer intervals resulted in estimates that were increasingly delayed. When sampled at 2-day intervals, 19.5% of estimates were delayed relative to the actual onset age, compared with 20.1% occurring earlier and 60.4% occurring on the correct date. Sampling at weekly intervals resulted in 34.3% of all estimates occurring later than the actual onset age. At 30-day sampling intervals, delay errors increased to 59.0% of all estimates. For all skills, acceleration errors did not change across sampling intervals. But delay errors increased with longer sampling intervals: The rate of increase followed a power function, $R^2 = .96$.

Discussion

A fundamental goal of developmental science is to understand change processes. To achieve this goal, researchers need accurate pictures of the shape of change, and such pictures require repeated observations. Most developmental researchers, however, do not conduct longitudinal and microgenetic studies because repeated observations are difficult and expensive to collect. The problem is compounded because overly large sampling intervals distort depictions of developmental change by obscuring important fluctuations in the data: Trajectories charted with binary data appear more abrupt than they really are, and trajectories charted with interval or ratio data smooth over important irregularities such as regressions and sudden changes in the rate of change.

The present study addressed the problem of selecting sampling intervals for developmental data by assessing the empirical costs of sampling at progressively larger intervals. The aim was not merely to confirm the loss of detail with coarser sampling but to determine how quickly depictions of development may be altered by sampling data at the rates typically used by developmental researchers. We compiled an illustrative data set of 32 infant motor skills and sampled daily to provide a fine-grained depiction of developmental change. We used real, rather than hypothetical, data to ensure that our sampling regimes incorporated actual patterns of variability into depictions of the shape of developmental change. Most skills showed a period of variability (vacillating between occurrence and absence) before acquiring a stable period of daily expression. When we simulated sampling at longer intervals (2–31 days), the picture of a variable acquisition period was quickly lost, so that skills with variable trajectories showed a single, step-like transition. Other critical aspects of the trajectories were also distorted: Most skills showed large delays in estimating onset ages.

Daily Changes in Infant Motor-Skill Acquisition

A surprising finding that emerged from analyses of the original, daily time series was the large number of transitions preceding stable performance. The widespread practice of using point-onset dates for motor skills (e.g., Adolph et al., 2003; Campos et al., 2000; Frankenburg & Dodds, 1967) presupposes that most skills appear suddenly and are consistently expressed thereafter. However, in the current study, a variable acquisition period characterized most skills for every infant across the entire age range. For example, infants averaged 14.57 transitions for standing ($SD = 4.96$), as illustrated in Figure 2, and 13.37 transitions across all skills ($SD = 10.35$). Is it really possible that infants vacillate between occurrence and absence of a skill on a day-to-day basis? Perhaps the variability is just noise and is not developmentally significant.

Several factors lend assurance that the diary reports were reliable indicators of daily performance. First, the parents were a select group of observers. Eight infants had parents who were professors or doctoral students and who had conducted behavioral research of their own. In addition, parents were carefully trained on the criteria for each skill and understood the importance of noting question-mark days when they had insufficient data to mark a skill present or absent. Most parents spontaneously annotated their diaries when infants' performance did not meet a criterion (e.g., number of crawling steps, seconds of independent sitting), suggesting that they took the criteria seriously and were eagerly waiting for performance to reach threshold. A second factor that inspires confidence in the daily data is a reliability study: A less select group of 95 parents provided reliable reports of sitting, crawling, standing, and walking skills using a daily checklist diary designed after the one used here (Bodnarchuk & Eaton, 2004). Home visits by experimenters blinded to the diary entries yielded concordant data for 11 of 12 measures. A third factor concerns the directional bias of parents' errors. If parents did err, the most likely errors were false positives. That is, observing infants pass a criterion on one day may have biased parents to produce "present" responses on the following days. False positives, however, would produce fewer transitions in the time series for any skill, suggest-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ing that the number of transitions reported here are, if anything, an underestimate of the true day-to-day variability.

Why then might infants have failed to express sitting, crawling, walking, or other basic motor patterns after demonstrating the ability to do so? Variable acquisition periods cannot be attributed to a lack of opportunity. We only analyzed skills that could be performed in a normal home environment (with floor, furniture, etc.) and that did not require special equipment or resources (e.g., stairs). Moreover, we eliminated days when the family situation precluded access to the floor (traveling, illness, etc.). Variable acquisition periods also cannot be explained as an artifact of low base-rate levels of performance. Nearly all (94%) of the 261 time series eventually reached a stable pattern of daily performance, suggesting that infants were highly motivated to perform the indexed skills.

A remaining possibility is that variable acquisition periods reflect a biological reality: As infants acquire new motor skills, they perform close to the limits of their abilities, much like athletes struggling to meet their personal best during competition. In early periods of skill acquisition, infants' peak skill level is far below the criterion level, and on a binary scale, the skill is considered absent. At later periods, as infants' abilities hover around the criterion threshold, their top level of performance exceeds criterion on some days but not others, resulting in variable trajectories. Eventually, infants' peak skill level comfortably surpasses threshold, and skills are expressed on a consistent, daily basis. To achieve a more stringent criterion for the same skill (e.g., walks >3 m vs. walks <3 m), infants must acquire a still higher level of peak performance.

This peak-performance interpretation implies that at least for gross motor skills, over the first year and a half of life, infants continually push the envelope of possibility by attempting actions that they have not quite mastered. Like Vygotsky's (1978) concept of a zone of proximal development, day-to-day variability in motor-skill performance may reflect periods of development when infants are operating close to their limits; they are most disrupted by perturbations and can benefit most from external support. This account also accords with previous proposals that motor skills are more unstable and sensitive to context when they first appear in infants' repertoires (Thelen, Fisher, & Ridley-Johnson, 1984; Robinson & Smotherman, 1992; Garciaguirre, Adolph, & Shrout, 2007). As infants' peak abilities expand, performance improves, and skills are expressed for longer durations and under more variable and challenging circumstances.

A question that arises about daily variation in infants' motor skills is whether even smaller sampling intervals would have revealed something additional. As in the example of physical growth, where episodic growth across days encompassed episodic growth across minutes, like a set of nested Russian dolls, smaller, meaningful units of motor action are nested within daily samples. For example, nested within the stuttering day-to-day trajectory of performance in crawling and walking, infants also show a variable trajectory in their expression of locomotion. On the scale of minutes and seconds, infants vacillate between short bouts of locomotion and longer periods of rest (Adolph, Badaly, Garciaguirre, & Sotsky, 2008; Badaly & Adolph, 2008; Chan, Lu, Marin, & Adolph, 1999). Variable expression from step to step produces a temporally distributed and spatially variable practice regimen that is most effective in promoting motor learning (Adolph & Berger, 2006). The intervening rest periods provide time to consolidate effects of practice and to renew infants' motivation. Thus, intermittent rest periods may be especially important when infants must operate at peak performance simply to execute crawling or walking steps.

These theoretical speculations about variable acquisition periods, however, depend on the characterization of the developmental trajectory. Without evidence for variable acquisition periods, the foregoing discussion of theoretical implications for motor skill acquisition would be moot. And without sampling at a rate that renders the same picture as the daily data, there would be no evidence for variable acquisition periods. Instead, we would be constructing an account to explain step-like transitions in the development of motor skills. A similar dilemma is posed for sampling development in other domains.

Empirical and Theoretical Costs of Sampling Decisions

Given the long history of microgenetic research (Vygotsky, 1978), methodologists' exhortations to select sampling intervals for reasons other than convenience, tradition, or intuition (Wohlwill, 1970, 1973) and formal demonstrations that long sampling intervals can compromise conclusions about development (Boker & Nesselroade, 2002; Collins, 2006), one might expect that developmental research would reflect the same care in choice of sampling regime as in experimental design. Unfortunately, it does not. The general principle that one must take sampling interval seriously in designing developmental studies is not reflected in current practice.

It is possible that general awareness about sampling on a developmental time scale has not yet filtered down to the rank and file. As Collins and Graham (2002) commented, a similar situation prevailed 40 years ago for the use of power analyses to determine sample size: Originally, power was a concept that statisticians worried about, but it was not widely applied in actual research settings. Now researchers routinely use power analyses to design their experiments as they balance the practical demands of minimizing sample size while avoiding the empirical and theoretical pitfalls of Type II errors.

How quickly we lose the picture of developmental change. How seriously must developmental researchers consider the problem of selecting a sampling interval? In previous studies of infant motor-skill acquisition that are touted in the literature as heroic examples of microgenetic research, observations were collected at weekly or monthly intervals, two of the larger intervals among our simulated sampling frequencies (Adolph, 1997; Corbetta & Bojczyk, 2002; Thelen et al., 1993; Thelen & Ulrich, 1991; Vereijken & Thelen, 1997). In the current study, daily sampling revealed that 84% of time series exhibited a variable pattern of emergence. When we simulated sampling infants' daily motor performance at larger intervals, the picture of day-to-day variability was quickly lost. When sampled once per week, fewer than half of these time series appeared variable, and when sampled monthly, only 9% appeared variable. In other words, sampling motor skills once a month caused 75% of the developmental trajectories to erroneously look abrupt and step-like, thus characterizing 93% of the entire time series with step-like trajectories. It should come as no surprise then that researchers typically consider the first appearance of motor skills to be the onset of a stable period of expression.

The shape of developmental change was not just distorted at the largest sampling intervals. Relatively small increases in interval length resulted in unexpectedly large decrements in sensitivity to variability. In fact, an inverse power function accurately described the rate of loss of sensitivity in portraying actual developmental trajecto-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ries. These findings indicate that, in the realm of motor development, the ability to detect variable developmental trajectories drops off extremely rapidly at sampling intervals longer than 2 to 3 days. It is the rapidity of this drop-off in sensitivity that is counterintuitive, not the fact that infrequent sampling generally reduces precision.

A second aspect of developmental profiles that was significantly affected by different sampling rates was estimates of onset ages. Increasingly large sampling intervals caused an increased rate of errors in estimating the earliest age of stable expression for motor skills. With 1-month sampling intervals, the average absolute error was 15 days, and 59% of errors were biased toward delays. In areas such as infant motor development and language acquisition, where skills appear and disappear in relatively quick succession, errors of this magnitude are likely to have serious consequences for both theory and application in studies of development. Erroneous onset ages carry concomitant costs for estimating durations of experience (e.g., how long a child has been walking or talking), developmental sequences (e.g., the ordering of motor and linguistic events), and the duration of stable periods (e.g., telegraphic speech, overregularization of verb tense).

Risks of oversampling. Of course, frequent sampling also carries potential costs. As others have pointed out (Cohen, 1991; McCartney, Burchinal, & Bub, 2006), substantial practical costs can be incurred by dense sampling. Collection of behavioral data, particularly in experimental settings, often entails considerable time, effort, and expense that may present logistical difficulties. Frequent sampling may have adverse effects on subject recruitment and attrition because demands on participation can be considerable and onerous. Repeated testing can alter participants' responses to the experimental condition, although this problem can be addressed explicitly by including a control group sampled less frequently. Dense sampling over a long period exacerbates problems of data management and methods for summarizing and analyzing data.

But does oversampling carry the risk of misrepresenting developmental trajectories, thereby causing researchers to misinterpret the research findings? Many time-based phenomena are evident only when assessed on the appropriate time scale. For example, it might be difficult to discern a 24-h circadian rhythm while viewing an activity record plotted on a time scale of seconds or to detect the day-to-day variability in infants' acquisition of crawling and walking over a trajectory that included the bout–rest periods of locomotion on a time scale of seconds. More generally, researchers might fail to detect a developmentally significant pattern on a larger time scale that is obscured by abundant low-level variability or noise in a densely sampled time series.

The interpretive problem, however, arises solely from failure to adequately summarize data obtained from dense sampling. There are no intrinsic interpretational problems that arise from sampling frequently, because any time series can be resampled at a reduced rate or smoothed to faithfully represent patterns at a lower grain of resolution. In fact, researchers routinely oversample physiological and movement data and then apply various smoothing functions to reduce noise and to detect underlying patterns in the data. In other words, researchers can recover the developmental pattern from oversampled data, but the converse is not true: Researchers cannot recover the developmental pattern from data sampled with overly large intervals.

Moreover, as illustrated by the findings in the present study, variable developmental trajectories are not an inevitable consequence of high sampling rates. Although we found that infant motor development is most often characterized by variable trajectories, the data also demonstrated that 15.7% of the daily time series showed a sudden, step-like transition, with the skill appearing from one day to the next. Contrary to the notion that high sampling rates might create the false impression of variability, only with sufficiently frequent sampling is it possible to refute the possibility that a developmental trajectory is variable and that a step function is a more accurate depiction of the underlying pattern. This fact is well appreciated by evolutionary scientists, who acknowledge the need for much finer resolution in the fossil record, on a geological time scale, to distinguish between competing theories of evolutionary change, such as gradualism versus punctuated equilibrium (Gould & Eldredge, 1993; Gingerich, 2001).

Beyond binary data. Can conclusions regarding the effects of sampling interval generalize beyond the specifics of the data set reported here? Sitting, standing, walking, and so on were scored as binary data (either present or absent over the course of each day), and all skills reached a level of stable, daily performance. Likewise, skills in other domains can be expected to attain stable, daily performance (e.g., correct production of words, learning the multiplication tables). Skills such as crawling and cruising (and in other domains, weaning from breastfeeding, the ability to distinguish speech sounds outside the native language, etc.) also attain stable offset periods, where children never produce them again. But what of skills scored as a binary process with base rates between 0 and 1? Symbolic play, for example, might achieve a stable base rate during the preschool years between .8 and 1, and professional hitting averages in baseball only rise to the neighborhood of .2 to .3. Going in the other direction, crying begins at 1 for newborns, but thankfully decreases to a base rate closer to 0. How does a base rate less than 1 (and for offsets, a base rate greater than 0) affect the optimal selection of sampling interval?

A simple Markov switching model can help to clarify the issue of generalization to skills with intermediate base rates. Even a high base rate will result in some days when the skill is not expressed. Figure 5 provides an illustration. The black curves represent data from three hypothetical time series; the gray curves represent a 15-day moving average that smoothes over the same data. Suppose that the developmental trajectory involves a step-like switch from an early period of absence ($p_E = 0$) to a later period of probabilistic occurrence ($p_L < 1.0$). For instance, as shown in Figure 5A, a sudden, step-like shift from absence to a .95 probability of daily expression would result in an average of 5 days when the skill was not expressed and 11 concomitant transitions (between absence and presence and vice versa) within a 100-day period that actually represents the stable base rate of the skill. Under these conditions, sampling on a daily basis would reveal occasional transitions in the stable base rate, which might be misidentified as a variable period of acquisition. In such a situation, it might seem preferable to sample less frequently, say, once a week or once a month, to reduce the chance of erroneously attributing transitions to a variable acquisition period, rather than to a stable, more mature period with a base rate <1. Obviously, if fewer samples were collected, fewer false transitions would be detected.

However, a reduced sampling rate would not provide a more accurate measure of the developmental profile. Instead, it would fail to identify the correct onset age in a step-like trajectory, and it would decrease the estimate of the number of transitions during the acquisition period for time series with variable trajectories. In contrast, for step-like trajectories, dense sampling would pinpoint

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

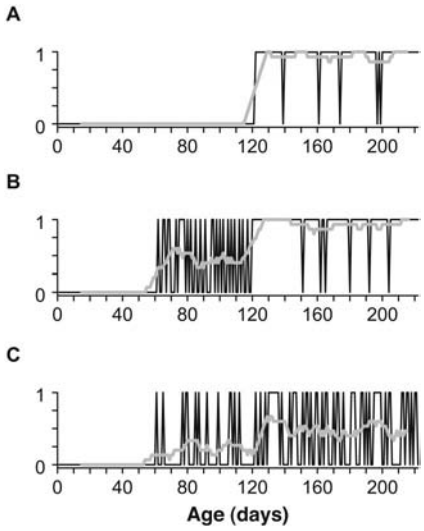

Figure 5. Simulated developmental trajectories (dark lines) generated by a simple Markov switching model. In each graph, the first 60 days represents a period in which the behavior of interest is not yet expressed ($p = 0$), and the final 100 days represents a period of consistent expression in which the behavior occurs at a stable rate <1. A: A stage-like trajectory involving an abrupt transition from absence (extended through the first 120 days) to a high base rate of occurrence ($p = .95$) during the period of stable expression. B: Trajectory involving an intervening acquisition period (from Day 61 to Day 120) before achieving a stable period with a high base rate ($p = .95$). During the acquisition period, behavior is generated by randomly switching between the early regime (absence) and the later period of stability (high base rate). C: Trajectory involving an intervening acquisition period before a stable period with a lower base rate ($p = .5$). Regime switching occurs in the same way as in B. In all three graphs, the thicker gray line shows a 15-day moving average that depicts the same data; in B and C, this smoothing function visually demarcates the variable acquisition period from the later period of stable expression.

the onset age. For variable trajectories, dense sampling would allow researchers to distinguish a variable acquisition period from a postacquisition period with a stable base rate <1, using the difference in the number of transitions (or some other measure of variability, as revealed by a smoothing function) as an index.

For example, Figure 5B presents the same two-state model (p_E and p_L) as in Figure 5A, but it is now separated by a 60-day window (representing a variable acquisition period) in which the underlying process randomly shifts between p_E and p_L. If p_L is high ($p_L > .8$), then the number of transitions detected by daily sampling will be greater during the variable acquisition period than during the later period of stable expression. But, as shown in Figure 5C, if p_L is low ($p_L < .5$), then the number of transitions during the variable acquisition period will be fewer than the number observed after the onset of stable expression. In both cases, a simple smoothing function can reveal differences in the level of expression during the variable-acquisition and stable periods. Thus, the difference in the number of transitions over the entire time series provides a clue as to whether the change from absence to stable expression is step-like or variable. Even though the absolute number of transitions can be inflated during acquisition for skills with base rates <1, only dense sampling can reveal differences in the rate of expression when expression of the skill is probabalistic.

Recording skills with greater precision does not alleviate the need for dense sampling to characterize the developmental trajectory. The findings regarding sampling intervals should also generalize from the binary data presented here to more precise levels of measurement (ordinal, interval, and ratio scales). Ordinal data, for instance, present much the same problem as do binary data for dealing with variable trajectories. On an ordinal scale, sampling less frequently poses the risk of missing periods of vacillation between higher and lower levels of performance, periods of inconsistent fluctuation between several levels, periods of consistent expression at intermediate levels, or wholesale reversals in levels of performance.

Similarly, when skills are indexed with interval or ratio data, infrequent sampling over development can lead to decreased sensitivity to detect periods of stability and instability (in single observations, means, and measures of variability, such as the coefficient of variation). Interval and ratio data are conducive to interpolation and curve fitting that smooth over variations in the developmental trajectory, so researchers may fail to detect brief episodes of improvement (as in the case of physical growth) or decrement, interruptions in the expression of skills, spikes in activity, accelerations (e.g., the vocabulary explosion) and decelerations, or other changes in rate that are evident only when sampling on a finer time scale. Moreover, because sensitivity to random sampling error is greater with larger sampling intervals, estimation of onset ages on the basis of achievement of a criterion level of performance, as measured with ordinal or interval data, also would be subject to the same types of errors that we have described for binary data.

Theoretical consequences. Developmental trajectories provide more than empirical summaries of change over time: Historically, evidence that cognitive, perceptual, social, or motor skills exhibit particular developmental trajectories (step-like, variable, linear, episodic, U-shaped, etc.) has stimulated some of the most important theories in developmental psychology. The concept of developmental stages illustrates the profound influence of empirical claims about developmental trajectories on theoretical work about developmental change. Stage theories have enjoyed a long and influential history in developmental psychology (e.g., Baltes, Reese, & Nesselroade, 1977; Brainerd, 1978; Piaget, 1954). Although the concept of developmental stages encompasses qualitative changes, hierarchical reorganizations, universal sequences, and so on (Fischer & Silvern, 1985), typically, a central feature of stage theories involves the timing of development—extended, stable periods interrupted by shorter periods of developmental change. Rapid, stage-like transitions from one stable pattern of performance to the next are characteristic of phenomena ranging from moment-to-moment fluctuations in sleep and waking states in rat pups (Blumberg, Seelke, Lowen, & Karlsson, 2005) to patterns of change on an evolutionary time scale (Gould & Eldredge, 1993).

Considerable effort has been devoted to constructing formal models that can account for abrupt transitions between developmental stages. For example, theoretical accounts of stage-like cognitive development include simulations using connectionist models (McClelland, 1989) and rule-based approaches (Siegler, 1976), mathematical models based on catastrophe theory (van der Maas & Molenaar, 1992), dynamic systems theory (Thelen & Smith, 1994), and other mathematical frameworks (van Rijn, van Someren, & van der Maas, 2003). For instance, according to catastrophe theory and dynamic systems theory, enhanced variability is a hallmark of transitions between stable attractors (Kelso,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

1995; Thelen & Smith, 1994), such as successive stages (Raijmakers & Molenaar, 2004). Thus, accurate assessment of the amount and timing of variability is critical for empirically evaluating such models of cognitive development.

The present study suggests that researchers' ability to accurately characterize variable and step-like trajectories in development is profoundly affected by sampling rate, and either trajectory may be inferred erroneously as an artifact of inadequate sampling. Models of developmental change become moot if the empirical evidence cannot distinguish among alternative trajectories. That is, without an appropriate sampling interval, researchers would not be able to detect a sufficient amount of variability to distinguish between punctate onset dates (Wimmers et al., 1998), instability around times of transitions (Kelso, 1995), expression of partial knowledge (Munakata, McClelland, Johnson, & Siegler, 1997), or other patterns of skill onset.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Sampling Development

How can developmental researchers avoid the pitfalls of undersampling? Unfortunately, formal principles such as the Nyquist theorem are not applicable to developmental time series because developmental trajectories can assume many different shapes, few of which are periodic or conform strictly to mathematical functions. Instead, sampling rates must be determined empirically on the basis of the questions being addressed and developmental processes being studied. Building on previous work (Siegler, 2006; Thelen & Ulrich, 1991), the present study suggests some precepts to guide the empirical enterprise of identifying optimal sampling rates to accurately capture the shape of developmental change:

1. *Determine the base rate.* In most cases, skills of interest to developmental psychologists eventually reach a level of stable, consistent performance. Estimating the typical rate at which the skill is expressed is important in planning how to sample the acquisition period and/or the more mature, stable period. For skills with stable periods of occurrence (or stable periods of absence, in the case of skills that disappear from children's repertoires), determining the base rate of occurrence depends on the number of observations collected, not on the rate of sampling. If the base rate is less than 1, applying a smoothing function may be useful in determining an average rate of occurrence (see Figure 5).

2. *Find the acquisition period.* For most kinds of skills, researchers are likely to initiate a study with some knowledge of the timeframe encompassing significant development. Some aspects of children's behavior emerge over a span of weeks; other aspects may require years. A preliminary investigation with economical sampling (at monthly intervals or longer) may be useful to identify the approximate age range for the acquisition period and thereby narrow the span of time requiring more detailed examination. Note that the initial characterization of the developmental trajectory from the preliminary study is unlikely to reveal a detailed and accurate picture of the shape of developmental change but may be necessary in planning more detailed sampling in future efforts.

3. *Sample as small as you can.* If the objective is to accurately portray the shape of a developmental process, it is crucial to sample data at the minimum, practicable interval, especially over the ages spanning the acquisition period. Researchers should consider the default rate for most kinds of child behavior to be daily sampling: absence or presence over the course of the day for skills indexed with binary data; highest level attained over the course of the day for skills indexed with ordinal data; and a summary score such as the mean, sum, or coefficient of variation over the course of the day for skills indexed with interval or ratio data. One reason to consider daily sampling a privileged sampling interval is that it reflects the nearly ubiquitous influence of 24-h circadian rhythms on human psychological functions. Skills expressed each day are interrupted by sleep each night, during which the day's activities and experiences may be absent, suppressed, forgotten, or consolidated (e.g., Stickgold, 2005). A second reason to consider daily sampling privileged is that the present study and others (e.g., Ganger and Brent's 2004 study of the vocabulary explosion and Lampl and colleagues' work on physical growth—Johnson et al., 1996; Lampl et al., 2001; Lampl et al., 1992) serve as demonstration proofs of important day-to-day changes in multiple domains of development. By sampling less frequently than every day, one risks losing the shape of those trajectories. Sampling multiple times each day may provide additional or converging insights into development, as in the cases of infant walking and physical growth. However, multiple samples per day also introduce variability that may not be meaningful because circadian rhythms affect patterns of performance by changing children's behavioral state, motivation, and opportunity for performance.

4. *Look before the onset.* To satisfy the objective of describing the entire trajectory, especially the shape of the acquisition period, researchers need to focus attention on the ages when the skill is first expressed. A preliminary investigation using coarse sampling should be useful for obtaining an initial estimate of an onset age, but as the findings of the present study show, estimates of onset ages based on infrequent sampling are likely to produce large delay errors, and such errors increase with larger sampling intervals. In trajectories based on monthly samples, for example, estimates of onset age are three times more likely to occur after the actual onset age. Therefore, the earliest expression of the skill is likely to occur before the earliest onset age identified by relatively infrequent sampling. As a consequence, more dense sampling efforts should include ages prior to the crude estimate of onset.

5. *Look for changes in variability.* For skills indexed by binary data, trajectories may be step-like or variable. The latter will show fluctuations prior to attaining a stable level of performance. If the base rate of occurrence is high but <1 during the period of stable performance ($>.8$), then a variable acquisition period will likely consist of an increased number of transitions. In contrast, if the base rate is low ($<.5$), then a variable acquisition period should show a lower number of transitions relative

to the later period of stability. As shown in Figures 5B and 5C, application of a simple moving average or similar smoothing procedure can reveal periods of enhanced variability, regardless of the base rate. After smoothing, variable periods appear at a lower rate of occurrence relative to later ages. Thus, smoothing techniques can be useful in demarcating changes in the level of variability of performance, which can help researchers verify that they have distinguished the acquisition and stable periods.

Concluding Remarks

Historically, much of developmental research has resembled the old saw about the man who lost his keys in a dark alley and turned his attention to searching for them on the street, "where the light was much better." Understanding child development has benefited from descriptions of age-related differences and demonstrations of the surprising abilities of infants. But understanding the process of developmental change requires more. It requires solid empirical foundations built upon accurate depictions of change over time. The implications of our analysis of sampling intervals would appear to offer a bleak view of methodological difficulties, even greater than those already recognized by researchers engaged in longitudinal and microgenetic research. The payoff for dealing with the thorny methodological difficulties of sampling rate is that accurate descriptions of developmental trajectories will be instrumental to advancing theories of development. It is simply necessary for understanding the shape of developmental change.

References

Adolph, K. E. (1997). Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development, 62*(3, Serial No. 251).

Adolph, K. E. (2005). Learning to learn in the development of action. In J. Lockman & J. Reiser (Eds.), *Action as an organizer of learning and development: The 32nd Minnesota Symposium on Child Development* (pp. 91–122). Hillsdale, NJ: Erlbaum.

Adolph, K. E., Badaly, D., Garciaguirre, J. S., & Sotsky, R. (2008). *15,000 steps: Infants' locomotor experience.* Manuscript in preparation.

Adolph, K. E., & Berger, S. E. (2006). Motor development. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology*: *Vol. 2. Cognition, Perception, and Language* (6th ed., pp. 161–213). New York: Wiley.

Adolph, K. E., Vereijken, B., & Shrout, P. E. (2003). What changes in infant walking and why. *Child Development, 74,* 474–497.

Badaly, D., & Adolph, K. E. (2008, March). *Walkers on the go, crawlers in the shadow: 12-month-old infants' locomotor experience.* Poster presented at the International Conference on Infant Studies, Vancouver, British Columbia, Canada.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods.* Belmont, CA: Wadsworth.

Bates, E. (1993). Comprehension and production in early language development: Comments on Savage-Rumbaugh et al. *Monographs of the Society for Research in Child Development, 58*(3–4, Serial No. 233).

Bloom, P. (2004). Myths of word learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 205–224). Cambridge, MA: MIT Press.

Blumberg, M. S., Seelke, A. M. H., Lowen, S. B., & Karlsson, K. Æ. (2005). Dynamics of sleep-wake cyclicity in developing rats. *Proceedings of the National Academy of Sciences, 102,* 14860–14864.

Bodnarchuk, J. L., & Eaton, W. O. (2004). Can parent reports be trusted? Validity of daily checklists of gross motor milestone attainment. *Applied Developmental Psychology, 25,* 481–490.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research, 37,* 137–160.

Brainerd, C. J. (1978). The stage question in cognitive-developmental theory. *Behavioral and Brain Sciences, 1,* 173–181.

Burchinal, M., & Appelbaum, M. I. (1991). Estimating individual developmental functions: Methods and their assumptions. *Child Development, 62,* 23–43.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. C. (2000). Travel broadens the mind. *Infancy, 1,* 149–219.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development, 61*(1–2, Serial No. 246).

Chan, M. Y., Lu, Y., Marin, L., & Adolph, K. E. (1999). A baby's day: Capturing crawling experience. In M. A. Grealy & J. A. Thompson (Eds.), *Studies in perception and action V* (pp. 245–249). Mahwah, NJ: Erlbaum.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23,* 43–71.

Clark, E. V., & Hecht, B. F. (1983). Comprehension, production, and language acquisition. *Annual Review of Psychology, 34,* 325–349.

Clifton, R. K., Muir, D. W., Ashmead, D. H., & Clarkson, M. G. (1993). Is visually guided reading in early infancy a myth? *Child Development, 64,* 1099–1110.

Cohen, P. (1991). A source of bias in longitudinal investigations of change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 18–25). Washington, DC: American Psychological Association.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology, 57,* 505–528.

Collins, L. M., & Graham, J. W. (2002). The effects of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. *Drug and Alcohol Dependence, 68,* S85–S93.

Corbetta, D., & Bojczyk, K. E. (2002). Infants return to two-handed reaching when they are learning to walk. *Journal of Motor Behavior, 34,* 83–95.

Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences, 10,* 131–138.

Crowley, P. H. (1992). Resampling methods for computation-intensive data analysis in ecology and evolution. *Annual Review of Ecology and Systematics, 23,* 405–447.

Darwin, C. R. (1974). A biographical sketch of an infant. In H. E. Gruber & P. H. Barrett, *Darwin on man* (pp. 464–474). New York: Dutton. (Reprinted from *Mind: A Quarterly Review of Psychology and Philosophy, 2,* 285–294, 1877)

Dasen, P. R. (1984). The cross-cultural study of intelligence: Piaget and the Baoule. *International Journal of Psychology, 19,* 407–434.

Dromi, E. (1987). *Early lexical development.* New York: Cambridge University Press.

Elman, J. (2003). Development: It's about time. *Developmental Science, 6,* 430–433.

Fischer, K. W., & Silvern, L. (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology, 36,* 613–648.

Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology, 2,* 421–453.

Flynn, E. (2006). A microgenetic investigation of stability and continuity in theory of mind development. *British Journal of Developmental Psychology, 24,* 631–654.

Frankenburg, W. K., & Dodds, J. B. (1967). The Denver Developmental Screening Test. *Journal of Pediatrics, 71,* 181–191.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology, 40,* 621–632.

Garciaguirre, J. S., Adolph, K. E., & Shrout, P. E. (2007). Baby carriage: Infants walking with loads. *Child Development, 78,* 664–680.

Gingerich, P. D. (2001). Rates of evolution on the time scale of the evolutionary process. *Genetica, 112,* 127–144.

Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58,* 1523–1531.

Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Studies in the development of behavior and the nervous system* (pp. 1–35). New York: Academic Press.

Gould, S. J., & Eldredge, N. (1993). Punctuated equilibrium comes of age. *Nature, 366,* 223–227.

Granott, N., & Parziale, J. (2002). *Microdevelopment: Transition processes in development and learning.* Cambridge, United Kingdom: Cambridge University Press.

Held, R., Birch, E., & Gwiazda, J. (1980). Stereoacuity of human infants. *Proceedings of the National Academy of Sciences, 77,* 5572–5574.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: A review of methodological issues. *Psychology and Aging, 18,* 639–657.

Johnson, M. L., Veldhuis, J. D., & Lampl, M. (1996). Is growth saltatory? The usefulness and limitations of frequency distributions in analyzing pulsatile data. *Endocrinology, 137,* 5197–5204.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Kleven, G. A., Lane, M. S., & Robinson, S. R. (2004). Development of interlimb movement synchrony in the rat fetus. *Behavioral Neuroscience, 118,* 835–844.

Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science, 6,* 133–139.

Lampl, M., Johnson, M. L., & Frongillo, E. A. (2001). Mixed distribution analysis identifies saltation and stasis growth. *Annals of Human Biology, 28,* 403–411.

Lampl, M., Veldhuis, J. D., & Johnson, M. L. (1992). Saltation and stasis: A model of human growth. *Science, 258,* 801–803.

Loreau, M. (1989). On testing temporal niche differentiation in carabid beetles. *Oecologia, 81,* 89–96.

Majnemer, A., & Barr, R. G. (2005). Influence of supine sleep positioning on early motor milestone acquisition. *Developmental Medicine & Child Neurology, 47,* 370–376.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4, Serial No. 228).

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58,* 110–133.

McCartney, K., Burchinal, M. R., & Bub, K. L. (2006). Best practices in quantitative methods for developmentalists. *Monographs of the Society for Research in Child Development, 71*(3, Serial No. 285).

McClelland, J. L. (1989). Parallel distributed processing: Implications for cognition and development. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 8–45). Oxford, United Kingdom: Clarendon Press.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science, 317,* 631.

McNeil, N. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology, 43,* 687–695.

Muller, G. B. (2003). Embryonic motility: Environmental influences and evolutionary innovation. *Evolution & Development, 5,* 56–60.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104,* 686–713.

Noonan, K. J., Farnum, C. E., Leiferman, E. M., Lampl, M., Markel, M. D., & Wilsman, N. J. (2004). Growing pains: Are they due to increased growth during recumbency as documented in a lamb model? *Journal of Pediatric Orthopedics, 24,* 726–731.

Nyquist, H. (2002). Certain topics in telegraphic transmission theory. *Proceedings of the IEEE, 86,* 280–305. (Reprinted from *Transactions of the AIEE,* pp. 617–644, 1928)

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development, 65,* 1228–1238.

Peterson, C. C., & Siegal, M. (1999). Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological Science, 10,* 126–129.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Popper, K. (1959). *The logic of scientific discovery.* New York: Harper and Row.

Price-Williams, D., Gordon, W., & Ramirez, M. (1969). Skill and conversation: A study of pottery-making children. *Developmental Psychology, 1,* 769.

Raijmakers, M. E. J., & Molenaar, P. C. M. (2004). Modeling developmental transitions in adaptive resonance theory. *Developmental Science, 7,* 149–157.

Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology, 28,* 406–413.

Robinson, S. R., & Smotherman, W. P. (1992). Fundamental motor patterns of the mammalian fetus. *Journal of Neurobiology, 23,* 1574–1600.

Shannon, C. E. (1998). Communication in the presence of noise. *Proceedings of the IEEE, 86,* 447–457. (Reprinted from *Proceedings of the IRE, 37,* 10–21, 1949)

Shultz, T. R. (1998). A computational analysis of conservation. *Developmental Science, 1,* 103–126.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8,* 481–520.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (2006). Microgenetic analysis of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, Perception, and Language* (6th ed., pp. 464–510). New York: Wiley.

Smotherman, W. P., & Robinson, S. R. (1995). Tracing developmental trajectories into the prenatal period. In J. -P. Lecanuet, N. A. Krasnegor, W. P. Fifer, & W. P. Smotherman (Eds.), *Fetal development: A psychobiological perspective* (pp. 15–32). Hillsdale, NJ: Erlbaum.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437,* 1272–1278.

Thelen, E. (1984). Learning to walk: Ecological demands and phylogenetic constraints. *Advances in Infancy Research, 3,* 213–260.

Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., & Zernicke, R. F. (1993). The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development, 64,* 1058–1098.

Thelen, E., Fisher, D. M., & Ridley-Johnson, R. (1984). The relationship between physical growth and a newborn reflex. *Infant Behavior and Development, 7,* 479–493.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Thelen, E., & Ulrich, B. D. (1991). Hidden skills: A dynamic systems analysis of treadmill stepping during the first year. *Monographs of the Society for Research in Child Development, 56*(1, Serial No. 223).

van der Maas, H. L. J., & Molenaar, P. C. M. (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review, 99,* 395–417.

van Rijn, H., van Someren, M., & van der Maas, H. (2003). Modeling

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

developmental transitions on the balance scale task. *Cognitive Science, 27,* 227–257.

Vereijken, B., & Thelen, E. (1997). Training infant treadmill stepping: The role of individual pattern stability. *Developmental Psychobiology, 30,* 89–102.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes* (M. Cole, V. John-Teiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Webb, S. J., Long, J. D., & Nelson, C. A. (2005). A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science, 8,* 605–616.

Wimmers, R. H., Savelsbergh, G. J. P., Beek, B. J., & Hopkins, B. (1998). Evidence for a phase transition in the early development of prehension. *Developmental Psychobiology, 32,* 235–248.

Wohlwill, J. F. (1970). The age variable in psychological research. *Psychological Review, 77,* 49–64.

Wohlwill, J. F. (1973). *The study of behavioral development.* New York: Academic Press.

Appendix

Skills Analyzed From Daily Diaries

Skill	Description
Sits (propped on hands)	Sits on floor for ≥30 s, with legs outstretched, using hands for support.
*Sits (hands free)	Sits on floor for ≥30 s, with legs outstretched, without using hands for support.
Sitting to prone	Shifts from sitting position with legs outstretched to prone position.
Prone to sitting	Shifts from prone or crawling position into sitting position with legs outstretched.
Kneel to stand (holding)	Shifts from kneeling, sitting, or crawling position to standing position by holding onto furniture to pull body upright.
*Squat to stand (hands free)	Shifts from kneeling, sitting, or crawling position into a squat, and then stands up without pulling upright on furniture.
Stands (holding)	Balances upright for ≥3 s by holding onto furniture for support.
*Stands (hands free)	Balances upright for ≥3 s without holding onto furniture for support.
Stand to sit (holding)	Shifts from upright to sitting position while holding onto furniture for support.
*Stand to sit (hands free)	Shifts from upright to sitting position without holding onto furniture for support.
Rolls front to back	Shifts from lying prone to lying supine.
Rolls back to front	Shifts from lying supine to lying prone.
Torso raised (propped on arms)	Pushes head and chest off floor by propping on forearms or hands while lying prone.
*Torso raised (1 arm free)	Pushes head and chest off floor by propping on 1 arm and using the other hand or arm to reach or manipulate objects.
Rocks on hands and knees	Rocks ≥2 oscillations while balanced on hands and knees.
Turns 180° prone	Pivots in place ≥180° while on belly or hands and knees.
Crawls belly (<3 m)	Crawls forward <3 m, before stopping, with belly resting on floor for duration of each crawling cycle.
*Crawls belly (≥3 m)	Crawls forward ≥3 m without stopping, with belly resting on floor for duration of each crawling cycle.
Crawls intermittent belly (<3 m)	Crawls forward <3 m, before stopping, with belly alternately raised in air and resting on floor during each crawling cycle.
*Crawls intermittent belly (≥3 m)	Crawls forward ≥3 m without stopping, with belly alternately raised in air and resting on floor during each crawling cycle.
Crawls hands and knees (<3 m)	Crawls forward <3 m, before stopping, balancing on hands and knees for duration of each crawling cycle.
*Crawls hands and knees (≥3 m)	Crawls forward ≥3 m without stopping, balancing on hands and knees for duration of each crawling cycle.
Crawls hands and feet (<3 m)	Crawls forward <3 m, before stopping, balancing on hands and feet for duration of each crawling cycle.
*Crawls hands and feet (≥3 m)	Crawls forward ≥3 m without stopping, balancing on hands and feet for duration of each crawling cycle.
Cruises 2 hands (<3 steps)	Takes <3 upright steps, torso sideways, holding onto furniture for support with both hands.
*Cruises 2 hands (≥3 steps)	Takes ≥3 upright steps, torso sideways, holding onto furniture for support with both hands.
Cruises 1 hand (<3 steps)	Takes <3 upright steps, torso frontward, holding onto furniture for support with 1 hand.
*Cruises 1 hand (≥3 steps)	Takes ≥3 upright steps, torso frontward, holding onto furniture for support with 1 hand.
Walks supported (2 hands held)	Walks with both hands held by caregiver, supporting own weight.
*Walks supported (1 hand held)	Walks with 1 hand held by caregiver, supporting own weight.
Walks (<3 m)	Walks independently <3 m.
*Walks (≥3 m)	Walks independently ≥3 m.

Note. An asterisk denotes a skill with stricter definition than preceding skill.

Received January 10, 2006
Revision received February 26, 2008
Accepted February 27, 2008 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.