Testing-the-Limits and Experimental Simulation: Two Methods to Explicate the Role of Learning in Development

Ulman Lindenberger, Paul B. Baltes
Max Planck Institute for Human Development and Education, Berlin, Germany

Key Words
Development · Learning · Mechanism · Plasticity · Research methods · Simulation · Testing-the-limits

Abstract
Developmental psychology needs a coalition of different research methods and strategies to achieve a more complete picture of the range of development and its causes. In this article, two such strategies are described – testing-the-limits and the experimental simulation of development. The aim of testing-the-limits research is to identify and decompose age-graded differences in upper limits of plasticity (the norm of reaction or the zone of potential development). The experimental simulation of development consists of a coordinated sequence of theory-guided research steps. Its key features are identification of causal mechanisms through the arrangement of experimental conditions that mimic (simulate) variations in developmental phenomena, and subsequent evaluation of the evidence in naturalistic settings. The two approaches are discussed and illustrated with examples from life span cognitive and social psychology.

Developmental psychology aims at the description, explanation, and modification of intraindividual change and of interindividual differences in intraindividual change [Baltes et al., 1988]. In trying to attain these goals, and akin to other ‘historical’ disciplines in general, it is faced with a special challenge. Some of its important phenomena and mechanisms are not easily amenable to observation, manipulation, and control [Baltes et al., 1988; Wohlwill, 1973]. In fact, it is fundamental to ontogenesis that its constituent components are complex and robust and, moreover, that its time course extends over decades. This is especially true if development is conceived of as a lifetime (life span) phenomenon [Baltes, 1987; Caspi, 1987], and is placed into the stream of two temporal contexts – ontogenesis and evolutionary-historical change [Baltes, 1968; Riegel, 1976; Schaie, 1965]. As a consequence, the standard ‘descriptive’ assessments used in most cross-sectional or longitudinal research are rarely sufficient in themselves to make clear the sources, direction, and variability (modifiability) of a given developmental phenomenon.
The methods presented in this article—testing-the-limits [Baltes, 1987; Kliegl and Baltes, 1987] and experimental simulation of development [Baltes and Goulet, 1971]—are both motivated by the rationale to move from description to explanation and modification. An additional motivation is to create a methodological instrumentarium that aspires to a fruitful coalition among descriptive, experimental, and causal modeling procedures [Collins and Horn, 1991; Kruse et al., 1993; Nesselroade, 1991; Willett and Sayer, 1994; Wood and Brown, 1994]. The two methods differ, however, in historical and theoretical backgrounds.

The testing-the-limits approach, in which large amounts of training and/or practice are invested to gain insights into an individual's latent performance potential, has two roots. One is the use of a testing-the-limits method in clinical-diagnostic psychology, where researchers such as Klopf and Rey applied follow-up procedures to examine whether subjects would demonstrate higher levels of performance under conditions that are more supportive than the original (standard) testing context [Schmidt, 1971]. The second, more developmental root is located in the gestalt and cultural-historical traditions of European developmental psychology, which include Vygotsky's conception of the zone of proximal development and Werner's microgenetic method [Vygotsky, 1962; Werner, 1948, Brown, 1982]. Its primary focus is the exploration of latent potential and the scope of what, at given points during ontogenesis, is possible and impossible in principle. The developmental interest in experimental simulation [Baltes and Goulet, 1971], in contrast, is linked historically to experimental child psychology [Baer, 1973; Baltes and Lerner, 1980; Reese and Liss, 1970]. As a method, the experimental simulation of development represents the continuing effort to understand sources of development in a manner that approaches classical experimentation and control. Together, the methods are useful in clarifying the intersection between learning and development.

We begin with the testing-the-limits approach, using research on adult age differences in maximum limits of memory functioning for illustration [Baltes and Kliegl, 1992; Kliegl et al., 1989, 1990; Lindenberger et al., 1992; Kliegl and Lindenberger, 1993]. We then outline the experimental simulation approach and provide an example from social aging research [Baltes, 1988, 1995] to illustrate its application.

**Testing-the-Limits: The Search for Plasticity and Its Limits**

As the term implies, testing-the-limits is aimed at two goals: (a) the search for limits of performance potential or plasticity (e.g., the identification of what is possible and impossible), and (b) the componential analysis of individual differences near limits of functioning. Not unlike the microgenetic approach to the study of change [Kuhn and Phelps, 1982; Siegler and Crowley, 1991], testing-the-limits is based on the assumption that changes occurring on different time scales share important features [Werner, 1948]. Therefore, the detailed analysis of time-compressed (i.e., accelerated or decelerated) developmental-change functions is assumed to enhance our understanding of the mechanisms and the range of medium- and long-term developmental changes. Specifically, the systematic analysis of intraindividual change processes under varying conditions of support or deprivation is meant to provide insights into individuals' 'range of plasticity', 'norm of reaction', 'zone of development', or 'latent potential'.

Why is a focus on plasticity a critical feature of developmental research [Baltes, 1987; Lerner, 1984]? Foremost, because phenotypic manifestations of development are
but one sample from an unknown population of possible outcomes or, biologically speaking, just one instantiation of the norm of reaction of the genotype. In other words, development, as contextualists and historically minded scholars forcefully argue, is a continuously evolving phenomenon. This characteristic applies to sociocultural evolution as well as to ontogenesis. Testing-the-limits is one crucial strategy to elucidate the range of the population of possible developments. If the focus is on the application of learning principles such as massed practice, testing-the-limits is also a powerful tool for specifying the role of learning in development. In this regard, certain testing-the-limit approaches have much in common with research on expertise [Chi, 1978; Ericsson and Smith, 1991].

In the next section, the testing-the-limits approach is illustrated with an example from research on cognitive aging. In this field [Baltes, 1993; Baltes and Labouvie, 1973; Labouvie-Vief, 1977; Salthouse, 1991], much of the early research was essentially descriptive in nature. The dominant focus was on findings of decline. There were reasons to assume, however, that these findings were not the best indicators of latent cognitive potential in old age because older adults – due to the lack of a ‘positive culture’ of old age and negative cohort differences in education and work-associated cognitive stimulation – may experience major practice deficits. Therefore, the finding of aging losses based on standard one-time assessments may not reflect an age-graded reduction in the latent competence, or as life span scholars put it, in the range of plasticity. Cognitive-training research was carried out to test this proposition. The most radical form of this training research is testing-the-limits.

Research Example: Adult Age Differences in Plasticity and Limits of Memory Functioning

In the field of cognitive aging, testing-the-limits has been used primarily in the study of adult age differences in ‘fluid’ intelligence [Cattell, 1971; Horn, 1982], or what we call the ‘mechanics’ of the mind [Baltes, 1993]. Our research in this field has been guided by two complementary propositions [Baltes, 1987; Baltes et al., 1984]. First, it is assumed that cognitive plasticity, or the potential to profit from experience, is greater than observed in descriptive cross-sectional or longitudinal research. Thus, healthy older adults were expected to have ‘reserve capacity’ that can be activated through training and other performance-optimizing conditions. In contrast, the second proposition is that older adults may not have the same amount of reserve capacity as young adults. Specifically, losses in upper limits of plasticity were postulated as a hallmark of biological aging [Baltes, 1987], and testing-the-limits was proposed to test whether aging-induced changes in the neurophysiological status of the aging brain would lead to a robust reduction in upper limits of performance. In other words, this second proposition states that the role and power of learning as inducer of manifested potential is limited by age-associated variations in plasticity that are largely biological in nature (e.g., the norm of reaction of the genotype).

An initial line of cognitive-training research [Baltes, 1987; Baltes and Willis, 1982; Willis, 1987] provided support for the first proposition. Healthy older adults have in fact greater cognitive potential than standard one-time assessments suggest. This work, however, did not permit insight into age differences in upper limits of developmental potential. Therefore, the initial focus on demonstrating the continued existence of old-age plasticity needed to be complemented by an investigation of age differences in upper limits of plasticity. A series of testing-the-limits studies was initiated at the Berlin Max
Planck Institute to achieve this goal [Kliegl and Baltes, 1987]. The primary domain of these studies involved measures of episodic and working memory, which we consider to be an important component of fluid intelligence [Kyllonen and Christal, 1990].

One series of training studies [Baltes and Kliegl, 1992; Kliegl et al., 1989, 1990; Lindenberger et al., 1993] was conducted with the Method of Loci, a mnemonic technique for the ordered recall of word lists [Bower, 1970]. In the Method of Loci, subjects over-learn a set of locations or loci in an invariant order. After acquisition of this cognitive routing map, to-be-learned items are sequentially associated with each of the locations of the map by means of a mental image. To recall the items, subjects ‘visit’ each location, try to remember the mental image, and retrieve the target item. Correlations between memory performance with the Method of Loci and tests of spatial visualization, perceptual speed, and working memory were moderate to high, suggesting that mnemonic performance is indicative of interindividual differences in the fluid ability domain [Kliegl et al., 1990; Lindenberger et al., 1992].

Identification of Age Differences in the Range (Plasticity) of Performance

The key feature of testing-the-limits research in the memory domain, not unlike research on expertise [Ericsson and Smith, 1991], was massed practice and intensive tutoring to provide subjects with what we considered an ‘optimal’ activation of latent potential. In one of our studies [Baltes and Kliegl, 1992], subjects participated in 38 experimental sessions distributed over about one year.

Figure 1, taken from Baltes and Kliegl [1992], displays the main findings from this study. The two guiding propositions – continued existence of reserve capacity coupled with aging-induced reductions in maximum performance potential – were borne out by the data. First, as illustrated by the training gains of older adults depicted in figure 1, most elderly persons were quite able to learn the Method of Loci and, by using it, to perform outside the usual range of performance in recalling the word lists. This finding is
consistent with the proposition that some developmental reserve capacity (plasticity) is preserved in old age. It also supports earlier findings from other domains of fluid intelligence [Baltes and Lindenberger, 1988; Baltes and Willis, 1982; Willis, 1987]. After activation of such reserve capacity, older adults are capable of reaching much higher levels of performance than obtained under standard assessments.

In addition, however, the data also provide strong support for the proposition of an aging-associated reduction in maximum limits of plasticity (in the zone of development or the norm of reaction). Even after very extensive training, older adults were not able to attain the level of performance reached by young adults after very few training sessions. The negative age difference was substantial, resistant to extensive practice, and applied to all subjects. As can be seen in figure 1, none of the older adults (who were positively selected to begin with) performed above the mean of young adults.

*Componential Analysis of Age Differences Near Limits of Functioning*

After the demonstration of age differences in limits in functioning, the second goal of testing-the-limits research is to decompose its sources. With respect to the aging losses in memory functioning, this componential work is similar to other research in the field of cognitive aging [Salthouse, 1991]. At the same time, it profits from the general gain in clarity and robustness associated with the testing-the-limits approach.

One finding emanating from the componential analysis of age differences in Method of Loci performance is that proactive interference is a special problem of the aging memory [Kliegl and Lindenberger, 1993]. In three experiments using an A-B, A-B paradigm [Kausler, 1994], the presentation times of location-noun combinations of a 30-item list were individually adjusted so that both young and older adults were operating at an accuracy level of 50% (e.g., on average, both young and older adults recalled 15 of the 30 nouns at their correct locations). Young adults needed about 1.8 s and older adults about 5.3 s per location-noun combination to work at this prespecified level of performance. This substantial age difference in time demands confirms the well-established finding that older adults need more time than young adults to encode and process task-relevant information [Lindenberger et al., 1993].

In addition, an error analysis revealed a more specific age difference. As can be seen in figure 2, older adults were more likely than young adults to erroneously recall nouns that had been correct in earlier lists, that is, to show intrusion errors or proactive interference. Because the subjects had much prior experience in the use of the Method of Loci, we can conclude that the age-differential susceptibility to proactive interference observed in this series of experiments is robust against large amounts of instruction and practice [Kliegl and Lindenberger, 1993]. The more general implication from the componential analysis of age differences in memory functioning is that older adults have a specific and experience-resilient deficit in relearning, revising, or updating memory [Baltes, 1993].

In sum, research on adult age differences in upper limits of memory functioning with the Method of Loci demonstrates the general usefulness of the testing-the-limits paradigm. Simple one-time observations often may be insufficient to answer a question that lies at the heart of much developmental theory and research—what is possible in development and what is not. Producing knowledge about the limits in performance potential (plasticity) and age-associated changes in plasticity carries specific advantages. First, it generates a phenomenon with a high degree of robustness and clarity that is not easily modifiable. Second, it allows for componential analysis of age-differential processes and mechanisms. Finally, due to reliance on experiential factors (e.g., instruction
Fig. 2. The componential analysis of age differences in memory performance near limits of functioning points to an aging-based increase in proactive interference. The figure displays criterion-referenced presentation times (left panel), correct recall (middle panel), and intrusion errors from earlier lists (right panel) as a function of list and age group. Adapted from Klugl and Lindenberger [1993, experiment 1].

and massed practice), it helps to identify the 'limits' of the possible role of learning as a contributor to development.

The Experimental Simulation of Development

The experimental simulation of development offers another methodological perspective [Baltes and Goulet, 1971]. Unlike testing-the-limits, its primary focus is not on plasticity (i.e., possible range) but on sources and mechanisms of development, and their operation in situ, topics that raise the issue of external validity [Baltes et al., 1988].

In a general sense, simulation or modeling techniques are theory-testing devices that combine the venerable 'thought experiments' of classical philosophy with the methodological diversity of modern science. Accordingly, there is not just one type of simulation but a great variety of paradigms and methods. Within psychology, one may set apart two broad classes of paradigms: simulation via modeling (e.g., computer simulations of psychological phenomena) and simulation via experimentation (e.g., experiments proper or observation and intervention in situ).

According to our definition [Baltes and Goulet, 1971; Baltes et al., 1988], the experimental simulation of development refers to the theory-guided arrangement of experimental conditions in successive studies that simulate or mimic variations in a given developmental phenomenon. In most cases, the main goal is to understand the sources and mechanisms of naturally occurring age changes in behavior. As a research strategy, the product consists of a coordinated sequence of seven steps which, however, do not need to be performed in the sequence specified (see table 1). A developmental phenomenon is considered to be well understood if knowledge based on all steps is available.

In step one, the target phenomenon is defined and described (e.g., Piaget's age sequence of cognitive operations, Erikson's life stages). In step two, a hypothesis or causal
<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Definition and description of target developmental phenomenon to be studied.</td>
</tr>
<tr>
<td>2</td>
<td>Postulation of a causal hypothesis or causal structure about underlying mechanisms and contextual conditions.</td>
</tr>
<tr>
<td>3</td>
<td>Experimental manipulation of relevant variables in the laboratory.</td>
</tr>
<tr>
<td>4</td>
<td>Test of experimental data against target phenomenon: isomorphism check.</td>
</tr>
<tr>
<td>5</td>
<td>Reexamination of causal hypothesis or causal structure (confirmation/rejection/modification) and search for alternative explanations.</td>
</tr>
<tr>
<td>7</td>
<td>Evaluation of external validity: interventive evidence.</td>
</tr>
</tbody>
</table>

Modified after Baltes et al. [1988].

Structure about the mechanisms underlying the phenomenon is formulated (e.g., increase in working memory capacity or assimilation/accommodation; psychosocial theory). In step three, variables assumed to affect the relevant mechanisms are manipulated in an experimental setting. The precise nature of the relationship between manipulated variables and hypothesized causal mechanisms depends on the theoretical assumptions made in step two.

In step four, the outcome of experimental manipulations is compared with the data that led to the formulation of the hypothesis. The goal of this comparison process, or 'isomorphism check' [Baltes et al., 1988], is to examine the extent to which the manipulation of variables in the experiment corresponds to age-graded variations in the target phenomenon. This isomorphism check, for instance, addresses the question of whether there is measurement equivalence between the naturally observed and the experimentally produced behavior. The comparison process may lead to the confirmation, modification, or rejection of the original hypothesis(es), and to a search for alternative causal mechanisms (step five).

Finally, it is necessary to move to systematic tests of external validity. Step six investigates whether variations in variables (processes) selected for experimentation indeed occur in natural settings. Step seven, the last step in the sequence, examines whether manipulation of these variables in situ results in expected outcomes.

We illustrate the experimental simulation of development with a program of research that has fully exploited the entire scope of this approach – the work on dependency (lack of autonomy) in old age conducted by M. Baltes and colleagues [1988, 1995]. In its substantive focus, this research program provides evidence for the importance of social-environmental factors in the emergence of dependent behavior in old age. Specifically, older adults' social partners have been found to reinforce dependent and ignore independent behavior. The relevance of this 'dependence-support, independence-ignore' script in the etiology of dependent behaviors is convincingly demonstrated through the manipulation of script-relevant processes in experimental and naturalistic interventions.

Historically, this program of research originated within the theoretical and methodological framework of operant developmental psychology [Baer, 1973]. The basis for the operant analysis of reciprocal influences between organism and environment is the patterning of contingencies between the organism and its environment (the sequence of SD
These contingency patterns are investigated under three different approaches, which, in concert, are assumed to provide a more complete picture of the etiology of a given behavior—observational-operant, experimental-operant, and operant-ecological intervention [Baer, 1973; Baltes and Baltes, 1982]. These three approaches map easily onto steps in the experimental simulation of development (table 1). Specifically, the experimental-operant approach corresponds to step three (manipulation of relevant variables in the laboratory), the observational-operant to step six (identification of variations in relevant variables in the natural environment), and the operant-ecological intervention approach to step seven (manipulation of relevant variables in situ).

The starting point (step one in table 1) for the research program on old-age dependency was the observation that many older adults display major deficits in autonomy in the domain of self-care [Baltes, 1988; Baltes and Barton, 1979]. In addition to biological aging loss, social-environmental factors were assumed to play a major role in this age-associated emergence of dependent behavior. Specifically, it was hypothesized (step two) that the learning contingencies of the social environments of older adults reinforce dependent rather than independent behavior and thereby accelerate the emergence and solidify the existence of dependent behavior patterns.

To examine this hypothesis, a series of experimental laboratory studies were conducted to examine the effects of operant treatments (stimulus control, practice, reinforcement schedules) on self-care behaviors in older adults. Thus, the hypothesized causal mechanism—differential reinforcement schedules for dependent and independent behavior—was systematically manipulated to examine whether it produces the expected outcome—a shift in the relative frequency of dependent and independent behaviors (step three). Most of this research was done in the 1970s using operant single-subject ABAB reversal designs with various reinforcement and stimulus control procedures.

In many instances, older adults' dependent behaviors were found to be reversible within the context of an experimental setting, supporting the notion that environmental factors (i.e., behavioral contingencies) exert some influence on their frequency of occurrence [Baltes and Barton, 1979]. Thus, the effects of treatment conditions corresponded to (were isomorphic to) variations in the target phenomenon (step four), which lent some support to the original hypothesis that social-environmental factors contribute to the etiology of dependent behavior in old age (step five).

After this demonstration of behavioral plasticity under controlled experimental conditions, a second phase of research was initiated. The purpose was to study the existence of corresponding antecedent-consequent contingencies surrounding self-care behaviors in the natural environment (steps six and seven). For this purpose, the flow of social interactions between elderly people and their caretakers was recorded in situ by means of a behavior-coding system and analyzed in terms of event-based behavioral contingencies. Using this theory-guided approach in a variety of settings, it was consistently found that the self-care related behavioral contingencies in the real-life ecologies of older adults corresponded to the postulated explanatory causal model (step six). Specifically, older adults' dependent self-care behaviors were consistently rewarded through complementary supportive reactions on the side of caretakers. Older adults' independent self-care behaviors, in contrast, rarely produced performance-enhancing social contingencies. This selective presence of two complementary scripts of behavioral contingencies (dependence-support, independence-ignore) was found to be generalizable across different institutions and to be largely insensitive to individual differences in need of care [Baltes and Wahl, 1992] (fig. 3).
The ultimate step in this program of research (step seven) was to manipulate the relevant causal variables (i.e., the reinforcement schedules of the social partners of elderly persons related to dependent and independent behaviors) in situ, that is, within the natural environment. Recently, an intervention of this kind has in fact been undertaken [Baltes et al., 1994]. In three different nursing homes, a staff-training program focusing on basic behavioral principles and behavior modification was implemented to reduce the prominence of the 'dependence-support, independence-ignore' script. Observational data on staff-resident interactions in the context of self-care were collected before and after the intervention. Consonant with the guiding hypothesis that social contingencies contribute to the etiology of dependent behavior in old age, findings revealed (a) a decrease in dependence-supportive and an increase in independence-supportive behavior among trained staff members, and (b) a concomitant increase in independent behaviors on the part of elderly nursing home residents.

Taken together, experimental findings regarding the malleability of dependent behavior in older adults, observational findings on naturally occurring contingencies, and results of interventions in situ strongly suggest that the social ecologies of older adults are a powerful regulator of dependent behavior in old age. At the same time, the journey into careful observation of systemic in situ contingencies also produced an unantici-
pated new insight, pointing to the need to consider additional contingencies and target behaviors in order to understand the systemic properties of the dependence/independence phenomenon. Specifically, it was found that the prominence of the dependence-support, independence-ignore script does not imply that older adults are passive recipients of help and care, and that the script's meaning is only about dependency. Rather, older adults were shown to utilize their own dependent behavior as an instrument for the control of specific positive events such as social contact. In that sense, the reinforcement schedule associated with the dependence-support script comes not only with a cost, but also with gains—the cost to become more and more dependent, but the concurrent gains associated with achieving a predictable social world and receiving social attention [Baltes, 1995; Baltes and Wahl, 1992].

Finally, the existence of surplus aging-related losses in independence beyond those associated with learning-based accounts also illustrates that development is not identical to learning. As to the aging-associated dynamics between dependence and independence, conditions of learning are but one, albeit major, contributing factor.

**Conclusion**

Ontogenetic change evolves from complex interactions between organismic and environmental conditions. To unravel the underlying processes and understand their outcomes, it is necessary to move beyond more 'descriptive' traditional longitudinal and cross-sectional research methods to those that allow for greater experimental control over antecedent conditions [Baltes et al., 1988; Kruse et al., 1993].

The focus of this article has been two methods that emerged in the context of life span theory and research, with its distinct emphasis on long-term ontogenetic change and the role of age-based organismic resources and constraints. The goal of the testing-the-limits method is to explore the range of possible developments across the life span in order to advance our knowledge about what is possible and impossible in development. For instance, knowledge about the upper limits of cognitive functioning and robust (seemingly irreversible) age differences in these limits yields information about the relative power of learning as a contributor to development. The experimental simulation of development complements this interest in plasticity with a more direct quest for the identification of causal mechanisms and a concern for external validity, to ascertain whether laboratory-based findings are operative in the everyday context of life-span development. In our opinion, joint use of the two approaches will lead to a more profound understanding of both the conditions and the range (plasticity) of human development.

**References**


Paul B. Baltes, Max Planck Institute for Human Development and Education, Lentzeallee 94, D-14195 Berlin-Dahlem (Germany)