Reinhold Kliegl, Ulrich Mayr, Ralf Th. Krampe

Time-accuracy functions for determining process and person differences : an application to cognitive aging

Time–Accuracy Functions for Determining Process and Person Differences: An Application to Cognitive Aging

REINHOLD KLIEGL, ULRICH MAYR, AND RALF TH. KRAMPE

Max Planck Institute for Human Development and Education, Berlin, Germany

A paradigm for the determination of time–accuracy functions (TAFs) for individual participants is introduced for two pairs of tasks differing in cognitive complexity, that is, word scanning vs cued recognition and figural scanning vs figural reasoning. TAFs can be used to test dissociations of cognitive processes beyond scale-related ambiguities of ordinal interactions. The approach is applied to examine the cognitive-aging hypothesis that a single slowing factor can account for interactions between adult age and cognitive task complexity. Twenty young and 20 old adults participated in 17 sessions. Presentation times required for 75, 87.5, and 100% accuracies were determined for each task with a variant of the psychophysical method of limits. Accuracy was fit by negatively accelerated functions of presentation time. State-trace analyses showed that different slowing factors are required for high- and low-complexity tasks. Relations to speed–accuracy and performance–resource functions are discussed. © 1994 Academic Press, Inc.

A typical question in the psychophysics of perceptual processes is how much of a particular resource is needed to achieve a specific level of accuracy? In many cognitive tasks, accuracy depends monotonically on the amount of available presentation time. In this article we treat presentation time as a resource in the psychophysical sense. We designed two pairs of tasks manipulating the level of cognitive complexity in processing verbal and figural material. For each of these tasks we determined for each participant the presentation time needed to achieve three different levels of accuracy, using a variant of the psychophysical method of limits. This criterion-referenced testing procedure yielded individual time–accuracy functions which reflect the induced complexity of cognitive processing as well as individual differences between persons.

We demonstrate that time–accuracy functions can be determined for a broad spectrum of experimental conditions encompassing cognitive pro-

cesses of scanning, episodic memory, and figural reasoning. The paradigm overcomes the traditional schism between performances assessed with latency and accuracy measures. Accuracy is modeled as a negatively accelerated function of presentation time covering the entire range of performance for each individual. Time–accuracy functions are used to address an important open question in the field of cognitive aging, namely the hypothesis that adult age differences in nonlexical tasks can be predicted without reference to cognitive content or task-specific processes (e.g., Cerella, 1990, 1991; Fisk, Fisher, & Rogers, 1992). Finally, our approach overcomes scale-related interpretational ambiguities associated with ordinal interactions in cognitive research (Loftus, 1978). This requires that time–accuracy functions are converted for state–trace analysis, that is equal-accuracy performance curves for two experimental conditions must be different for different experimental groups (Bamber, 1979).

TIME–ACCURACY FUNCTIONS

Time–accuracy functions (TAFs) specify how much presentation time is needed to achieve a particular accuracy level. The two hypothetical functions displayed in Fig. 1 can serve different purposes; they can describe differences between two tasks due to the type or complexity of cognitive processing involved, differences between two persons or groups due to characteristics such as ability or age, and changes within a person or group due to learning, development, or a shift in cognitive strategy. In general, the left curve reflects less resource-demanding performance because less presentation time is required to achieve the same accuracy.

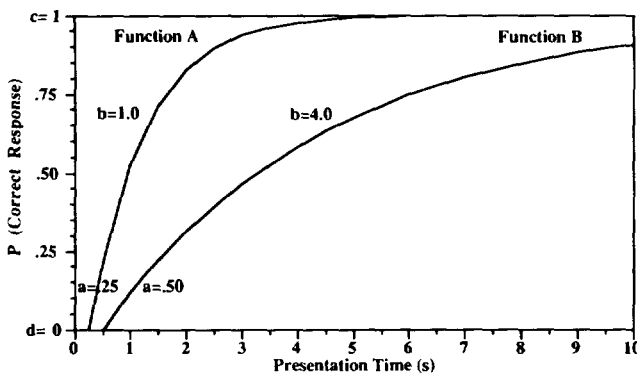The functions in Fig. 1 were generated on the assumption that the

FIG. 1. Two theoretical time–accuracy functions based on negatively accelerated exponential functions: $p = d + (c - d)(1 - \exp(-(t - a)/b))$.

relation between time t and accuracy p corresponds to a negatively accelerated exponential function

$$p = d + (c - d) * [1 - \exp(- (t - a)/b)], \tag{1}$$

for $a \geq 0$, $b > 0$, and $0 \leq d < c \leq 1$. This function has been used in speed–accuracy tradeoff research to summarize the relation between response latency and accuracy (e.g., Lohman, 1986, 1989; Wickelgren, 1977). It captures the fact that time demands for a proportionate reduction of error probability are linearly related to accuracy. Different from typical applications in speed–accuracy tradeoff research, t stands for presentation time—not for response latency. Information necessary to solve a given task (e.g., word pairs in a cued recognition task or figural objects in a scanning task) is presented for varying amounts of time; the accuracy achieved at each level of presentation time is the dependent variable. This represents a fundamentally different, psychophysical approach because presentation time is under experimental control, whereas in speed–accuracy tradeoff research response latency is a dependent variable. Our method is similar to the determination of performance curves for visual persistence or iconic memory (e.g., Loftus, Truax, & Nelson, 1987; Loftus, Duncan, & Gehrig, 1992). All three approaches share the conceptual interpretation of parameters: beyond a minimum presentation time (parameter a), performance approaches the asymptotic maximum accuracy (parameter c) at a constant rate of $1/b$. Thus, parameter b describes the degree of curvature of the function; the smaller the b, the steeper the function. The minimum level of accuracy expected (parameter d) is primarily a function of the response format (e.g., the guessing probability).

In the following set of tasks, the maximum performance was assumed to be 100% for all persons; the chance level of performance was 50% because a two-alternative forced-choice procedure was used; thus, parameters c and d were fixed at 1.00 and .50, respectively. Parameters a and b were estimated for each person in each of the four tasks. They reflect the minimum amount of time required to do better than chance (parameter a) and the time required to decrease the error probability $(1 - p)$ by a constant proportion (parameter b). We will refer to this parameter as processing time. In the past this parameter has frequently been specified as a rate (i.e., $b' = 1/b$). Our specification has the advantage that the same metric applies to parameters a and b; they are both scaled in units of time. Therefore, the parameters are also in the same metric as response-time tasks. This is relevant because hypotheses about the effects of experimental manipulations and age on cognitive processes are frequently tested with latency data. With this specification we hope to facilitate a transfer of familiar "latency schemas" (e.g., "old = slow"/

"young = fast," "complex = slow"/"simple = fast") to the estimates of processing times proposed here.

Parameter estimates should relate to experimental manipulations and individual differences in the same way response latencies do but with one difference. They are estimated with respect to a mathematical model that is assumed to hold for the entire range of presentation times and accuracy. In other words, the same cognitive mechanism is assumed to govern task performance at all levels of presentation time and accuracy within a task or within an individual. Therefore, different from latency measures, a significant difference in processing times (parameters b) implies a significant interaction between the main effect of task (or group) and presentation time when accuracy of performance is used as the dependent variable.

The negative exponential function can be derived from the discrete stochastic replacement model of information processing (Restle & Greeno, 1970).[1] According to the replacement model, wrong response tendencies are replaced by correct ones at a constant rate. As the pool from which these response tendencies are sampled is assumed to be fixed, the probability of sampling "novel" correct tendencies decreases linearly with time. Parameter b is a proxy of the time required to achieve a constant proportional reduction in error probability. Differences between tasks or individuals in this processing time are not necessarily the result of differences in the duration of "microlevel" processing steps. Depending on task or person characteristics, internal processing cycles could be differentially susceptible to error; for example, the need to repeat such cycles might lead to additional processing time for the same level of accuracy to be obtained. Moreover, although processing time is conceptually linked to the presentation phase it could reflect processing constraints of the response phase. For example in the case of an episodic memory task, a free recall format would lead to a less steep TAF than a cued recall format for the same learning material. Obviously, this lack of specificity about microlevel processing lags behind the detailed level of theorizing about many cognitive tasks. The advantage of the "generic processing description" is that it localizes differences between tasks and individuals at a level below the observed behavior, but at a level that (a) is conceptually shared by a wide variety of cognitive tasks, and (b) captures the dynamics of cognitive processing for the entire functional range between chance and perfect performance in relation to a meaningful physical resource such as "presentation time."

Using the time–accuracy paradigm, the present research tests the gen-

---

[1] We will consider alternative negatively accelerated functions discussed in the literature (i.e., hyperbola and power function) in a final set of analyses of this article.

erality of negatively accelerated exponential function (a) across low- and high-complexity tasks from two different domains (i.e., figural processing and verbal/memory) and (b) across generally low-achieving (old) and generally high-achieving (young) adults. Mapping performances across such a large spectrum of task complexity and individual differences onto a common metric at the level of process parameters was the theoretical goal and the methodological challenge of our study.

GENERAL AND PROCESS-SPECIFIC AGE DIFFERENCES

In the present experiment we used time–accuracy functions to examine the hypothesis that negative adult age differences in nonlexical cognitive measures can be explained by a task-independent slowing factor (Birren, 1965; Salthouse, 1985). Metaanalytic summaries of age differences in response–time tasks indicate that old adults are slowed relative to young adults by a proportional factor of about 1.6 (see Cerella, 1990, for a review). A graphic illustration of this finding is obtained by plotting old adults' against young adults' mean latencies of corresponding experimental conditions (Brinley, 1965). The slowing factor is reflected in the slope of the linear function. Assuming age equivalence in the processing algorithm, this function has been interpreted in terms of an unspecific reduction of processing speed affecting all cognitive processes to the same extent because old adults' mean performance can be predicted from the mean of young adults' performance without reference to the "cognitive content" of experimental conditions. Linear functions usually accounted for more than 90% of the variance; nonlinear power or quadratic functions provided a somewhat better fit, suggesting that age differences may increase somewhat for more difficult conditions (Cerella, 1990; Myerson, Hale, Wagstaff, Poon, and Smith, 1990). Empirical support for such general slowing models, however, was restricted to experiments using response times as dependent measures with equal accuracy for age groups. For example, complex reasoning or episodic memory, which are usually measured in terms of accuracy, were precluded from analysis. What has been missing is, in analogy to Cerella's (1990, p. 219) call for a "joint speed–accuracy platform," a joint time–accuracy platform for the investigation of cognitive adult-age effects.

We will show that TAFs permit a meaningful comparison of age effects across tasks in which presentation time is a critical resource. TAFs were obtained for old and young adults in word scanning, cued recognition, figural scanning, and figural reasoning tasks. Task selection was guided (a) by the general goal to represent a broader range of task conditions in one experiment than investigated previously (e.g., verbal and figural, simple and complex), (b) by past findings regarding large age effects in memory performance after mnemonic training (Baltes & Kliegl, 1992;

Kliegl, Smith, & Baltes, 1989), and (c) by a theoretical distinction between sequential and coordinative complexity (Mayr & Kliegl, in press).

Memory accuracy after mnemonic instruction emphasizing the construction of interactive images or thoughts has been shown to be strongly age-sensitive (Baltes & Kliegl, 1992; Kliegl et al., 1989). However, due to incompatible metrics of measurement it was difficult to relate these results to the typical "general slowing" age effects found in response time tasks. In two age-comparative studies, memory was also measured in the time domain by determining the amount of presentation time required for equal accuracy (Kliegl & Lindenberger, 1993; Thompson & Kliegl, 1991). Old–young ratios were always larger than 3.5, that is considerably larger than the typical 1.6 observed in response–time tasks. In the absence of a control task, however, it remains unclear to what degree differences in experimental format are responsible for the divergence. Thus, one important question of this study was whether mnemonic performance actually does constitute a "special" domain with respect to cognitive aging, a domain not covered by general slowing models based on nonlexical information processing tasks.[2] To this end, a cued-recognition (plus mnemonic instruction) and a word-scanning task were included in the present experiment. Basic perceptual–cognitive factors such as reading speed limit performance in both word scanning and cued recognition. We assumed that additional processing limitations related to the construction of interactive images or thoughts would emerge in the cued-recognition task.

The dissociation between sequential and coordinative complexity was proposed by Mayr and Kliegl (in press) and relates assumptions about working memory functioning to predictions about processing demands in rapid figural transformation tasks. Sequential complexity denotes conditions in which processing time requirements are assumed to depend only on the number of independent processing steps involved; in the context of aging research they should reflect a relatively basic-level slowing. In coordinative complexity conditions, processing steps are interrelated, creating simultaneous demands on storage and processing in working memory. Mayr and Kliegl concluded that coordinative processes are particularly age-sensitive for response latencies and criterion-referenced presentation times. Using TAFs and criterion-referenced testing we expected not only a replication of the dissociation between sequential and coordinative complexity, but also an estimation of relative age effects on the basis of TAF parameters. Conditions of sequential and coordinative complexity were generated in a figural transformation paradigm. Figural scanning (the sequential-complexity condition) required simple compari-

---

[2] There is already evidence, for example Lima, Hale, and Myerson (1991), that lexical tasks show less slowing than nonlexical tasks (e.g., lexical decision vs visual search tasks).

sons between two arrays of objects whereas figural reasoning (the coordinative-complexity condition) also involved the identification and consideration of a spatial transformation of one of the arrays. Identifying and keeping in mind the spatial transformation was assumed to increase the coordinative burden in reasoning as compared to the scanning condition.

In summary, two potential refutations of general slowing were implemented in two task pairs: old–young ratios of basic processing times should be larger for episodic memory and figural reasoning than for word scanning and figural scanning, respectively. For both task pairs the use of TAFs provides a homogeneous frame of measurement for experimental conditions that are usually difficult to compare.

PROCESS DISSOCIATION: STATE TRACES AND YOUNG–OLD FUNCTIONS

How can TAFs be used to test specific effects (process dissociations) against "general factor" models? The most frequent sources of evidence for decomposition in cognitive, developmental, and neuropsychological research are ordinal interactions. Unfortunately, these interactions can be interpreted as process dissociations only if a particular scale is assumed, that is, if they cannot be made to disappear by monotonic transformations of the dependent variable (Bogartz, 1976; Loftus, 1978). This argument applies to most experimental work but is exacerbated in research where nonexperimental variables such as age or pathology are involved because the possibility that an unspecific general factor debilitates (or enhances) performance is more obvious here. Dunn and Kirsner (1988) outlined an approach, the principle of reversed association, that establishes process dissociation under conditions of ordinal interactions. The minimum design required is a three-factorial (2 × 2 × 2) design. They start with the assumption that the relation between the theoretical efficiency of a cognitive process and task performance is monotonic for the three experimental manipulations—an undisputed assumption for many cognitive tasks. As a test of process dissociation, the levels of one of the three factors are used to define a two-dimensional space within which the four pairs of means associated with the three factors are plotted. Dunn and Kirsner (1988, p. 98) showed that the line connecting these four data points in two-dimensional space must be monotonically increasing or decreasing if, in principle, a general factor explanation can account for the pattern of means; process dissociation (reversed association) is indicated in a discontinuity (a dip) of this function.[3]

A similar line of reasoning underlies Bamber's (1979) state-trace anal-

---

[3] Strictly speaking a (2 × 3) design would suffice. In this case the process dissociation, if obtained, would be carried by the levels within an experimental manipulation.

ysis. A state trace is a graph displaying the covariation of two dependent variables. The young–old plots described in the last section represent a particular type of state-trace analysis. In these analyses the state trace (i.e., the young–old function) was determined by regressing old adults' means assembled from various experimental conditions (and experiments) on corresponding means of young adults (Brinley, 1965; Cerella, Poon, & Williams, 1980; Salthouse, 1978, 1992). There are several problems associated with this procedure, for example, the discarding of within-group variance and covariance, the lack of independence of observations, and possible confounds between type of task and task difficulty (cf., Fisher, 1992; Hertzog, 1991; Kliegl & Mayr, 1992; Mayr & Kliegl, in press). In particular, given that task difficulty (i.e., error probability or response latency) generally increases with task complexity, it is experimentally very difficult to select conditions in such a way that the confound between level of performance and task complexity can be avoided. This problem is also a serious limitation of the approach by Dunn and Kirsner (1988) and Bamber (1979). Basically, we were still faced with the problem that ordinal interactions frequently disappear with a suitable nonlinear transformation of the dependent variable.

The present approach avoids the problem of having to select levels of task difficulty that allow the identification of process dissociation via state traces. If complete TAFs are available for two experimental conditions, young–old functions can be derived that encompass the entire functional range between chance and asymptotic maximum performance. Young–old functions are equal-accuracy state traces because old adults' need of presentation time for a specific accuracy is plotted as a function of young adults' presentation time for the corresponding accuracy. For TAFs such as those shown in Fig. 1, the equal-accuracy function relating old adults' time demands to those of young adults is easily derived from Eq. (1). For equal probability p and equal parameters c and d and after solving for t_{old}, a typical "proportional slowing" function obtains

$$t'_{old} = k * t'_{young}, \tag{2}$$

where $k = b_{old}/b_{young}$ for $t'_{old} = (t_{old} - a_{old}) \geq 0$, and $t'_{young} = (t_{young} - a_{young}) \geq 0$. The linear translation of t_{old} to t'_{old} and of t_{young} to t'_{young} limits the young–old function to the range of presentation times for which equal accuracy can be achieved for both age groups beyond chance performance (see Fig. 1); it also defines a new coordinate system where young–old functions pass through the origin.

The central question of process dissociation in the present context is whether the same ratio k of this young–old function characterizes all conditions or whether the ratio k is task dependent (in other words, aging is general or aging is process specific). General slowing predicts the ratio

k of old and young processing times b to be invariant across tasks. As mentioned above, we expect particularly large age-related slowing for episodic memory and for figural reasoning. As we estimate b parameters for each person in each condition, the equality across conditions of the slowing ratio k can be tested with a standard repeated-measures analysis of variance using log-transformed processing times b as dependent variables; age differences in ratios are reflected in significant age × condition interactions for these log-transformed processing times b (Cohen & Cohen, 1975, p. 250). In other words, we test the specificity of age-related slowing in proportionate rather than absolute-difference measurement space.

As an illustration why this test cannot be performed for untransformed processing times, assume old adults' means of parameters b were 2s and 4s for a simple and a complex task, respectively, and corresponding values for young adults were 1s and 2s. The larger age difference for the complex task yields a significant interaction between age and task for untransformed processing times. In Eq. (2), however, these hypothetical parameter estimates lead to identical young–old functions for the simple and the complex task (i.e., for both tasks, $k = 2$). Thus, the interaction will *not* be significant if we analyze logarithms of parameters b; the interaction will be significant, however, with logarithms of b if the ratio k ($= b_{old}/b_{young}$) differs between tasks.[4]

The test of process-specific age differences (ANOVAs of logarithms of processing times) which we propose amounts to a test of the null hypothesis that one monotonic state trace is sufficient to describe the data. Rejecting the null hypothesis implies rejecting the simple general-factor explanation. Analogous to the principle of reversed association proposed by Dunn and Kirsner (1988), the minimum design for process dissociation requires three experimental factors: one factor is used to fix the dimensions of the space (e.g., old vs young adults), a second factor manipulates task difficulty (e.g., different levels of matched accuracy), and the third factor is hypothesized to carry the dissociation (e.g., episodic memory vs word scanning; figural scanning vs figural reasoning).

Different from our approach (which follows the practice of previous cognitive aging research), Bamber (1979) and Dunn and Kirsner (1988) defined the dimensions on the basis of the experimental conditions and plotted state traces for different groups of people (e.g., normal and amnesic persons). For the analysis proposed here, this means deriving equal-accuracy time demands for episodic memory and word scanning; and,

---

[4] For this analysis, estimation of parameter b in units of "time" yields statistically identical results to estimation of parameter b' as "rate" because in logarithmic space they are linearly related [i.e., $b' = 1/b$; $\log(b') = -\log(b)$].

similarly, equal-accuracy time demands for figural reasoning and figural scanning. Young and old adults cannot be characterized by a single state trace if the ratios $(b_{\text{cued recognition}})/(b_{\text{word scanning}})$ or $(b_{\text{figural reasoning}})/(b_{\text{figural scanning}})$ for old adults are significantly different from the corresponding ratios for young adults. The two types of analyses yield identical results.[5]

To summarize, the main goals of this experiment were to show: (a) that time–accuracy functions can be determined at an individual level for four different tasks covering a broad spectrum of cognitive processes, (b) that ratios of old and young processing times are larger in tasks involving episodic memory or coordinative complexity than in tasks involving sequential complexity, and (c) that the approach presented here offers tests of dissociations of cognitive processes which are not subject to scale-related interpretational ambiguities associated with ordinal interactions.

METHOD

Participants

Twenty young and 20 old persons participated in 17 experimental sessions. Young adults (10 male, 10 female) ranged in age from 21 to 29 years (M: 25.3 years). Old adults (9 male, 11 female) ranged in age from 65 to 84 years (M: 70.3 years). To assure comparability with previous research, the WAIS digit symbol substitution was administered. Young adults earned a score of 59.2 (SD: 7.4); the corresponding score for old adults was 45.9 (SD: 8.3). These values are typical of scores observed in US samples of cognitive aging research (e.g., Salthouse, Kausler, & Saults, 1988). There were no significant differences between age groups in ratings of life satisfaction, subjective physical health, and subjective mental health, indicating good to average scores relative to those of age peers. Participants were paid 20 DM (about $13) per session.

Schedule

Table 1 contains the experimental schedule. In sessions 1, 2, 15, and 16 demographic information was collected and various standard psychometric and experimental measures were taken. Session 2 involved instruction of cued recognition and figural reasoning. Blocks of three sessions for each of the four tasks were carried out in sessions 3 to 14. Session 17 was used for debriefing.

Tasks

Experimental tasks were implemented on Macintosh-II computers. Stimuli were presented on a 13″ monitor (black and white mode). Responses were collected on button boxes.

In word scanning, participants were shown four concrete nouns displayed in one line on the monitor (font: Times; size: 24). After the presentation time the words were masked and two words appeared in the lower half of the screen. One of these had been among the stimuli. The task was to push a button corresponding to the old word. After 16 items, the computer counted the correct responses and adjusted the presentation time for the next

---

[5] If $b_{11}/b_{12} > b_{21}/b_{22}$, then $b_{11}/b_{21} > b_{12}/b_{22}$ with first subscripts indicating complex (1) and simple (2) task conditions and second subscripts indicating old (1) and young (2) adults.

TABLE 1

Experimental Schedule

1	Questionnaire, vocabulary, digit–symbol substitution, simple and choice reaction time, reaction times for word scanning and figural scanning, demonstration and instruction of cued recognition and figural reasoning
2	Instruction and pretest with fixed presentation times for cued recognition and figural reasoning; 10 blocks with 16 items each
3	Word scanning: 20 blocks with 16 items; criterion: 87.5%
4	Word scanning: (1) 8 blocks with 16 items; 87.5%; (2) 8 blocks with 16 items; 100%; (3) 8 blocks with 16 items; 75%
5	Word scanning: like session 4; plus four lists with 80 ms presentation time per item
6–8	Figural scanning (analogous to word scanning)
9–11	Cued recognition (analogous to word scanning)
12–14	Figural reasoning (analogous to word scanning)
15–16	Post-test (similar to session 1 and 2)
17	Debriefing

block of items. Words ranged in length from four to nine letters. In each block the target word appeared equally often in each of the four positions. Presentation was initiated and terminated with a random-dot mask covering the region in which words were presented; durations of these masks were 750 and 500 ms, respectively. Targets and distracters were identical in the first two letters; their lengths differed at most by one letter. Each target word was used only once.

Cued recognition required participants to encode 16 location–noun pairs (e.g.. lake–tiger). Participants were instructed to use interactive imagery as described, for example, in Kliegl, Smith, and Baltes (1990) for the method of loci mnemonic. Subjects practiced this mnemonic technique in two lists. The same 16 locations (e.g., lake, fountain, church, zoo) were used on all lists with different orders during encoding and retrieval; locations were assigned equally often to serial positions within lists across the experiment. Concrete nouns intended for recall were used only once; to ensure concreteness of nouns only those denoting objects that could be touched were included. At retrieval, a location cue and two words from the current list were presented. The subject had to decide which of the two words had been paired with the location. Thus, during retrieval each noun appeared once as a target and once as a distracter. Participants were instructed how to use this constraint to rule out wrong pairs and infer correct response alternatives.

In figural scanning two arrays, each containing four objects, were presented. One of the objects differed in one of the four features: shape, size, shading of the margin, shading of inside area. The task was to identify the critical feature. In the example of Fig. 2, the upper left object had changed in the shading of its peripheral area. Presentation of items started with a 750-ms random-dot mask. After the presentation time had expired both arrays were masked and two features were offered. In this case, "A" for "Aussen" ["margin" in German], would have been the correct response; the distracter "F" [Form] suggested a change in the shape of one of the objects. Within each block critical features occurred equally often in the four positions of the array; each of the four features was equally often the critical one. The distracter feature was with equal probability one of the three remaining ones.

Figural reasoning was a variant of figural scanning. The only difference was that the right
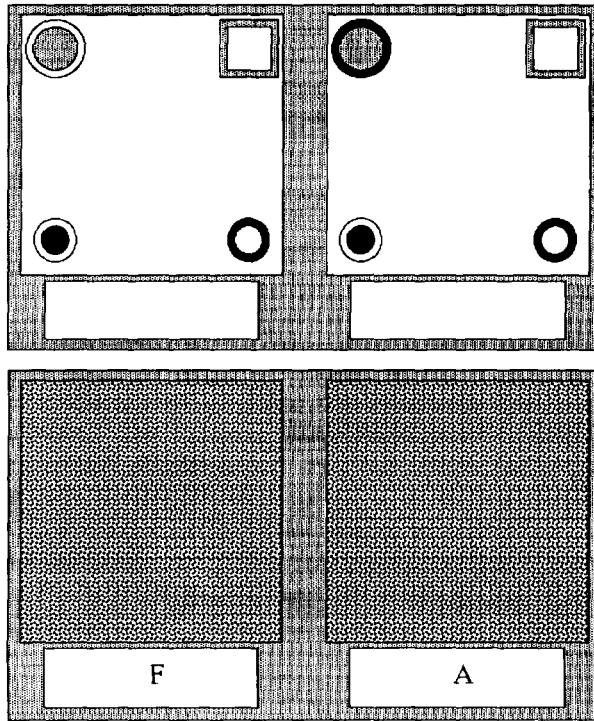
FIG. 2. Example of a figural scanning item.

array had been rotated with equal probability either 90 degrees clockwise or counterclock-wise. The subject still had to identify the feature of the object that had changed but in addition she or he also had to identify and take into account the rotation of the array.

Procedure

The data base for estimating individual time–accuracy functions was achieved through a three-session, criterion-referenced, adaptive testing module devoted to each of the four tasks. In the first session of the module the adaptive procedure was calibrated, and subjects were trained in the task. In the second and third session the actual testing occurred.

In the first session 20 blocks of 16 items were administered. The criterion regulating the adaptive process was 14 of 16 items correct (87.5%). The presentation time for the next block of items was determined by decreasing or increasing presentation time on a fixed scale between blocks of items depending on whether the criterion was met or not on the current block. The presentation times were constructed by taking 1000 ms as a starting point and going upward in 20% steps and downward in 16.67% steps. Thus, differences between presentation times were equal on a logarithmic scale. In total, 30 presentation times between 80 and 15,450 ms were used. In the calibration session, the adjustment of presentation times was based only on every other value of this table to achieve fast convergence.

In the second session presentation times were determined first for 87.5%, then for 100%, and finally for 75% with eight blocks each. Starting values for the 75 and 100% criterion were

three presentation time steps below and above the median of the last nine blocks at 87.5% of the first session. The third session was identical in format to session 2 except that at the end an additional four lists were administered as fast as the computer would permit (i.e., 80 ms per item) to obtain an empirical estimate of chance performance.

Subjects were instructed to always attempt 100% accuracy. They were also told that response time was not critical; they should maximize the number of correct responses. In the instruction sessions, time was also devoted to optimizing guessing strategies for the equal probability, two-alternative forced-choice situations.

RESULTS

Illustration of Data Protocol for One Participant

By using three different accuracy criteria (75%, 87.5%, 100%) the adaptive testing procedure automatically sampled—for each subject in each task—presentation times from the time-sensitive segment of the time–accuracy function without confronting subjects with noninformative presentation times (i.e., either too long or too short). Figure 3 displays the experimental protocol for one old adult in the figural reasoning task. In the top panel, logarithms of presentation times are plotted as a function of the 68 blocks of figural reasoning across the three sessions. Except for the first value, presentation times were selected according to whether or not the criterion was met on the preceding block. The corresponding accura-
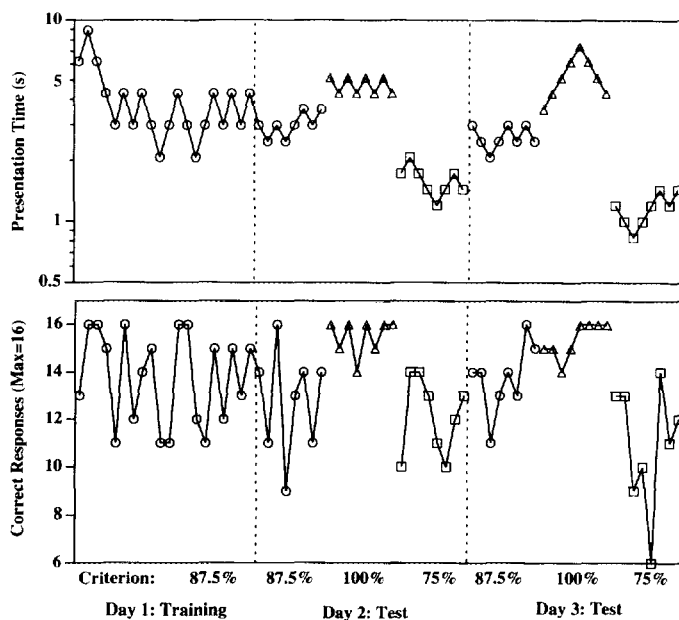

FIG. 3. (a) Criterion-referenced presentation times of a typical old adult in the figural reasoning task across three sessions; (b) corresponding accuracies.

cies are shown in the bottom panel of Fig. 3. Obviously, the presentation times increased with the accuracy criterion specified; for example, much longer presentation times were needed to meet the 100% criterion than the 75% criterion.

Estimation of Time–Accuracy Functions

For each block of items presentation time was the independent and accuracy was the dependent variable. Combining data from sessions 2 and 3, three data vectors were obtained: (a) the presentation times encountered during the adaptive procedure, (b) the number of items for each of these presentation times encountered during the adaptive procedure, and (c) the probability of correct responses associated with each presentation time. In Fig. 4 the probability of correct responses is plotted against presentation time for one young adult (top panels) and one old adult (bottom panels) in each of the four tasks; figural reasoning data for the old
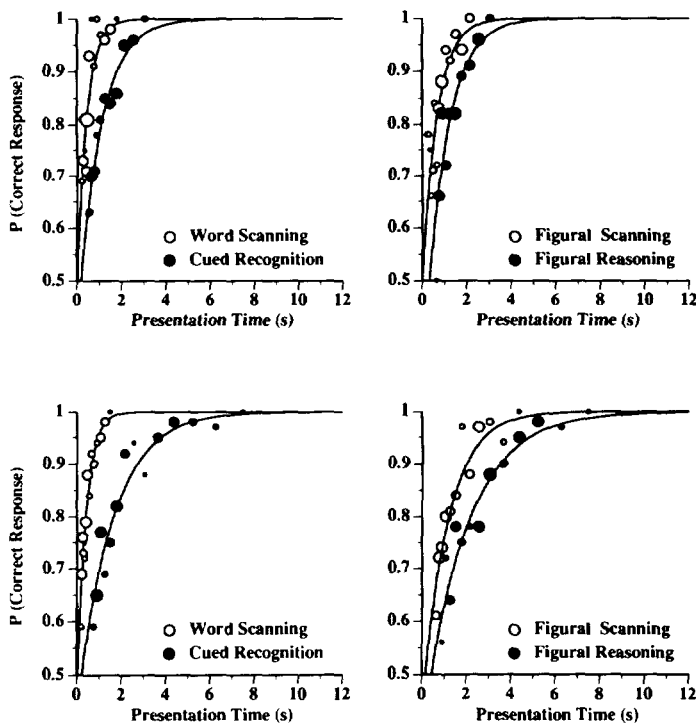

FIG. 4. (Top) Data of one young adult for the two low- and high-complexity task pairs. (Bottom) Data of one old adult. Continuous lines indicate best fitting negative exponential functions with two free parameters (intercept on x axis and slope).

adult are based on those displayed in Fig. 3. The size of the symbols codes the number of observations for a specific presentation time which differed across tasks and between persons because of the adaptive testing procedure.

The variance of the negatively accelerated functions considered in this article depends on the probability of correct responses [i.e., $N * p * (1 - p)$]. Therefore, a minimum χ^2 statistic was used as an estimator [i.e., $(n_o - n_e)^2/n_e$, where n_o and n_e are observed and expected frequencies]. This required the following steps: (a) responses were sorted into the cells of a Type of Response (correct, wrong) × Presentation Time contingency table; (b) χ^2 was computed by summing across the cells of the table; (c) parameters for which this statistic was at a minimum were determined (we used the CNLR-Module of SPSS-X to this end; SPSS, Inc., 1988). The degrees of freedom for a data set were equal to the number of different presentation times minus two (i.e., the free parameters). Although two cells (i.e., correct and wrong responses) were available for each presentation time, only one degree of freedom was associated with each because the number of items for a presentation time was fixed in the marginal. On the assumption that items are independent, the statistic is asymptotically χ^2 distributed. The continuous lines in Fig. 4 represent the best fitting two-parameter exponential functions; the goodness of fit statistics for the functions of Fig. 4 were around the median of the functions obtained in this experiment.

Goodness of Fit of Exponential Reference Functions

A total of 160 data sets (40 participants × 4 tasks) was available to evaluate the goodness of fit of the exponential function (Eq. 1). Results are summarized in the left part of Table 2. The lack of overall fit of the exponential function for the total data set ($p < .01$) was primarily due to the cued-recognition data, in particular those of young adults. When the worst four data sets (i.e., the only data sets of 160 for which the ratio of χ^2 and degrees of freedom was larger than 2.5 in the exponential function) were left out, the probability of the χ^2 statistic fell below the 5% level of significance.

The worst four data sets originated in two young and two old adults' cued-recognition condition. The poor overall quality of fit for young adults' cued-recognition data, however, was not just a problem of a few subjects; rather, individual χ^2s were significant at the 5% level for 8 of 20 young participants. We inspected plots of all functions but data were not suggestive of a different function—negative acceleration was obviously present. The problem of fit for young adults was related to larger variability compared to the other conditions. One reason for the divergence

TABLE 2

Goodness-of-Fit Statistics for Exponential, Hyperbolic, and Power Functions for Young and Old Adults in Four Tasks

	Exponential function				Hyperbolic function			Power function		
	df	χ^2	p	n_{05}	χ^2	p	n_{05}	χ^2	p	n_{05}
Word scanning										
Young adults	213	202.0		2	231.3		2	186.5		1
Old adults	224	212.1		1	198.0		0	162.2		0
Cued recognition										
Young adults	174	299.5	**	8	320.2	**	9	282.9	**	7
Old adults	222	268.8	*	2	251.5		3	240.3		3
Figural scanning										
Young adults	217	193.9		0	254.2	*	2	208.0		1
Old adults	181	170.2		0	248.9	**	3	214.1	*	2
Figural reasoning										
Young adults	194	203.2		0	268.0	**	4	228.9	*	3
Old adults	214	248.0		2	238.8		1	234.1		2
Total	1639	1795.7	**	15	2010.8	**	24	1756.8	*	19
− Worst four										
data sets	1601	1680.7		11	1903.0	**	20	1637.6		15

Note. df, degrees of freedom for χ^2 statistic; p, probability of χ^2 statistic (*$p < .05$; **$p < .01$); n_{05}, number of subjects (out of 20) with significant χ^2 statistic ($p < .05$). "Worst 4 [of 160] data sets" were cued recognition data from two old and two young adults.

could be due to the procedural peculiarity of the memory condition where an entire list of words had been presented before responses could be collected. Problems of exponential fit for cued recognition of old adults were due to two participants. They had the lowest memory scores and performed below 100% accuracy even when 15 s of presentation time (the maximum in this study) was available for each word pair. Nevertheless, when the asymptote (parameter c) was free to vary it still was estimated as 1.0. The slopes of these functions were the slowest and third slowest observed but not slow enough to qualify as outliers (that is, they were within three standard deviations of the mean). In the discussion we reconsider the theoretical validity of the exponential function for cued recognition.

We also carried out weighted least-squares analyses to obtain estimates of the amount of variance explained at an individual level. Median corrected R^2s were .90 for word scanning (range: .70 to .97), .88 for cued recognition (range: .61 to .98; plus the two old adult "problem cases": .23 and .50), .87 for figural scanning (range: .58 to .98), and .84 for figural reasoning (range: .61 to .98).

Analyses of Parameter Estimates

Estimates of parameters a and b of the exponential function were analyzed for effects of task complexity and age.[6] Young and old adults' means and standard deviations are listed in the top block of Table 3. In Fig. 5 corresponding time–accuracy functions for the four tasks based on the group means of these two parameters are displayed. Within each age group, curves were steeper for the low-complexity conditions of word scanning and figural scanning than the high-complexity conditions of cued recognition and figural reasoning. Moreover, in each condition older adults needed more presentation time to match the performance of young adults. The age difference in additional time needed was larger for high- than low-complexity tasks. The pattern of TAFs was in agreement with expectations about effects of task complexity and associated age differences.

The interpretation of TAFs was confirmed in two ANOVAs of parameter estimates a and b with age group (2) and the within-subject factor task (4); the α-level was set at 5%. Two orthogonal contrasts were specified to test the increase in complexity (a) from word scanning to cued recognition and (b) from figural scanning to figural reasoning. Of primary theoretical relevance are the processing times reflected in parameter b (see Table 3, bottom block). The main effect of age $[F(1,38) = 48.1, MSe = 1.48]$, both task contrasts $[F(1,38) = 22.4, MSe = 1.77; F(1,38) = 94.3, MSe = 0.24]$, and their respective interactions with age $[F(1,38) = 11.0, F(1,38) = 43.7]$ were significant. For equivalent proportional reduction of error probability, young adults needed less processing time than old adults, word scanning required less time than cued recognition, and figural scanning less time than figural reasoning. Age differences were larger for the complex condition of each task pair. Thus, the expected age × type-of-complexity interaction was demonstrated for each of the two contrasts in the estimates of processing time needed for converting presentation time into greater performance accuracy. Interactions between age and task for processing times imply a significant three-way interaction with additional differential effects of presentation time if accuracy were used as a dependent variable, accentuating age effects for long presentation times.

For parameter a, a significant difference between word scanning and cued recognition, $F(1,38) = 60.1, MSe = 0.03$, was obtained. A slight anomaly is apparent for process initialization of cued recognition, as old adults were estimated to need less time than young adults; this difference was almost significant, $t(38) = 2.00, p < .055$. It is the primary source of

---

[6] All the following analyses were also carried out for parameter estimates based on weighted least-squares without any substantive differences in results.

TABLE 3
Means and Standard Deviations of Parameter Estimates of Time–Accuracy Functions
for Power and Exponential Reference Functions

	Word scanning		Cued recognition		Figural scanning		Figural reasoning	
	M	SD	M	SD	M	SD	M	SD
Exponential function								
Parameter a								
Young adults	.0397	.0572	.4159	.2100	.0908	.1057	.1636	.1607
Old adults	.0733	.0835	.2612	.2758	.3711	.1944	.2694	.3919
Parameter b								
Young adults	.3608	.1320	.7823	.5215	.7342	.1832	1.0762	.3392
Old adults	.6488	.3335	3.0397	2.6692	1.4011	.3749	3.1991	1.2132
Power function								
Parameter a								
Young adults	.0857	.0636	.5841	.1961	.1866	.1391	.3513	.1524
Old adults	.1717	.0946	.7715	.3180	.7104	.2236	1.1444	.4827
Parameter b								
Young adults	.2540	.0809	.4220	.1779	.4506	.0865	.5673	.1221
Old adults	.3776	.1107	.9529	.4875	.6511	.0825	.9820	.1828

the interaction between the contrast of word scanning vs cued recognition
and age, $F(1,38) = 6.7$. The main effect for the second contrast, figural
scanning vs figural reasoning, and its interaction with age were not sig-
nificant.

Young–Old Functions and State-Trace Analyses

In the introduction we showed that even in the presence of an interac-
tion between age and task for processing times, there could be a single
monotonic function linking the processing times of both tasks. Can age
differences in these task pairs be explained by a single young–old function
or state trace? Our expectation was that different linear young–old func-
tions (Eq. 2) are required for tasks characterized by low complexity and
those of high complexity. From Table 3, it is apparent that old–young
ratios k of parameter b for low-complexity conditions (word scanning:
1.80; figural scanning: 1.91) were smaller than those for high-complexity
conditions (cued recognition: 3.89; figural reasoning: 2.97). To test wheth-
er these condition-specific age ratios were significantly different from
each other the log-transformed values of b parameters were analyzed.
The two critical age × task-contrast interactions were significant: for
word scanning vs cued recognition $F(1,38) = 11.9$, $MSe = 0.20$, and for
figural scanning vs figural reasoning $F(1,38) = 32.7$, $MSe = 0.03$. Thus,
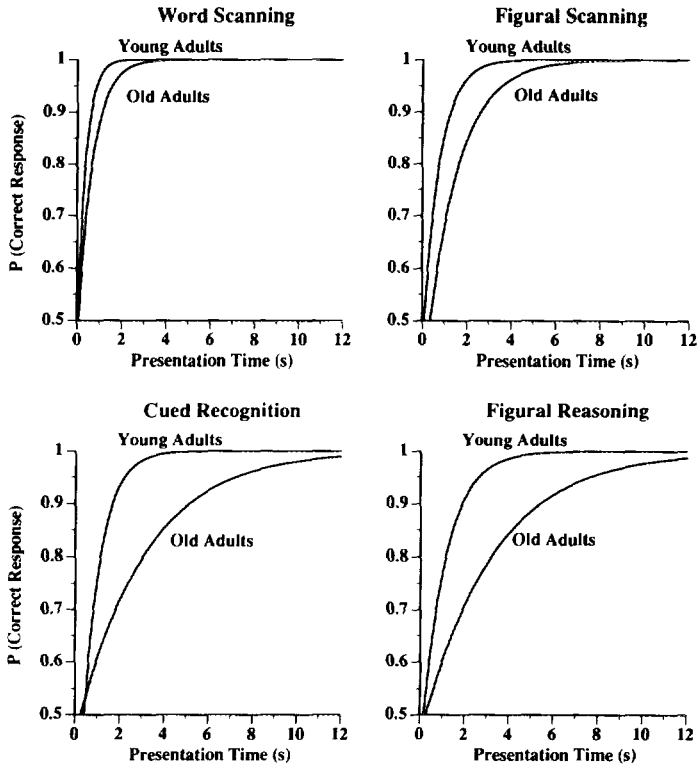the age differences reflected in the four tasks of this study cannot be

FIG. 5. Time–accuracy functions for young and old adults by condition. Functions are based on group and condition means of parameters *a* and *b* estimated with reference to the exponential function (see Table 2).

explained by a single linear function of cognitive slowing which ignores information about task complexity.

To illustrate the dissociation between low- and high-complexity conditions, young–old functions (top) and state traces (bottom) are shown for exponential TAFs in Fig. 6. The results support the hypotheses that episodic memory processes related to the construction of interactive images or thoughts and working memory processes requiring the coordination of information exhibit larger age effects than those obtained in the two low-complexity control conditions. These age × condition interactions are invariant against monotonic transformations despite their ordinal character.

Cognitive psychology is blessed with an abundance of ordinal interactions. State-trace analysis can help to separate those interactions that remain significant under monotonic transformations of the measure from those that, in principle, could be explained by a common mechanism
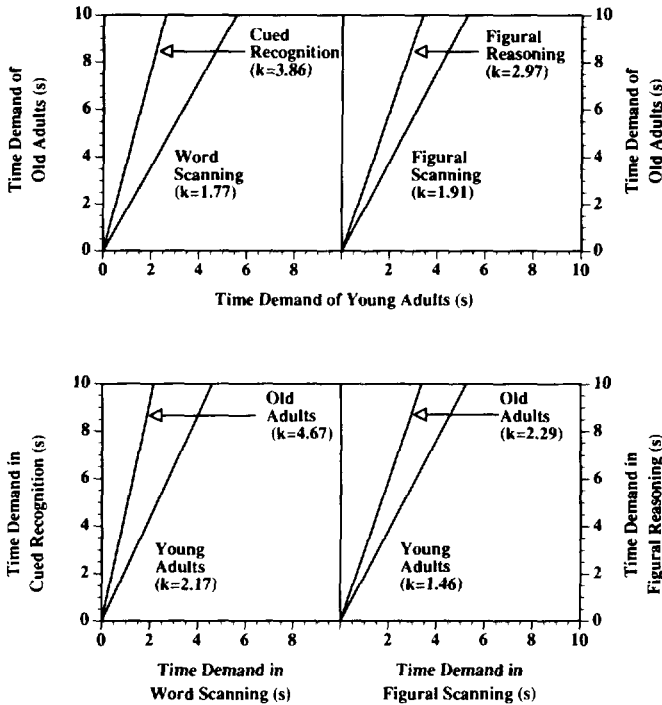
FIG. 6. Equal-accuracy young–old functions (top) and state traces (bottom) based on exponential time–accuracy functions (TAFs).

(Bamber, 1979; Dunn & Kirsner, 1988). For example, in the young–old space, state traces for word scanning and figural scanning were not significantly different from each other, neither were those of cued recognition and figural reasoning. Thus, the age differences in the four conditions of this experiment are compatible with two complexity-related mechanisms. In this respect, state-trace analysis can be thought of as a "factor analysis" for experimental research; it yields the minimum number of mechanisms required for the description of ordinal interactions (Dunn & Kirsner, 1988, p. 100).[7]

Alternative Model Equations for Time-Accuracy Functions

The exponential function is but one possibility to model the negatively accelerated relation between time and accuracy. The consideration of

[7] With a minimum of three different state traces, one can move one level up and ask whether state traces such as those displayed in Fig. 6 can be traced to a single second-level mechanism, that is to a state trace of a state trace.

alternative TAFs is important because we will show that they can lead to different state traces than the one derived for the exponential function. In particular, we consider a hyperbolic and a power function. Mazur and Hastie (1978) used the hyperbola and Newell and Rosenbl~ ~m (1981) the power function to model learning and skill acquisition. Tne difference between the exponential and the hyperbolic functions is that in the latter, in terms of the underlying stochastic model, wrong response tendencies are not replaced by correct ones but rather correct response tendencies are simply added to the pool of response tendencies; negative acceleration, then, is a consequence of the accumulation of correct response tendencies (Restle and Greeno, 1970). The difference between the exponential and the power function is that for the power function the processing time b is not a constant factor of performance but increases linearly with it. For skill acquisition, Newell and Rosenbloom (1981) describe the power function as an "exhaustion of exponential learning"; in the present context it reflects an exhaustion of proportionate reduction of error. Both Mazur and Hastie (1978) and Newell and Rosenbloom (1981) ended up favoring their respective alternative over the exponential function on empirical grounds. The hyperbola used by Mazur and Hastie (1978) is given as

$$p = d + (c - d) * [(t - a)/(t - a + b)]. \tag{3}$$

The power function we used is given as

$$p = d + (c - d) * [1 - (1 + t - a)^{-1/b}]. \tag{4}$$

Parameters a, c, and d retain the interpretation introduced for the exponential function. Also as before, parameter b is scaled in units of time and reflects the steepness of the curve; smaller values (or shorter processing times) reflect a steeper ascent. Thus at a general, descriptive level the conceptualization of processing outlined earlier for the exponential function also applies to these alternative functions but, as indicated above, there ~re processing-specific differences in how time is converted into accuracy.

The choice of the model function has important consequences for tests of process dissociation; different mathematical specifications of the TAF may imply different young–old functions or state traces. The young–old function for the hyperbola of Eq. (3) is identical to the one derived for the exponential TAF, that is the linear function of Eq. (2). For the power function of Eq. (3), however, old and young adults' demands of presentation time for equal accuracy lead to the power function

$$t'_{old} = (t'_{young})^k, \tag{5}$$

or, equivalently,

$$\log (t'_{old}) = k * \log (t'_{young}), \tag{5'}$$

where $k = b_{old}/b_{young}$ for $t'_{old} = (t_{old} - a_{old} + 1) \geq 0$, and $t'_{young} = (t_{young} - a_{young} + 1) \geq 0$. Thus, empirical evidence in favor of one of the alternative TAF specifications implies logical constraints on the state traces derived from them.

Goodness of fit. Goodness-of-fit statistics for the hyperbola and the power function are presented in the middle and right part of Table 2. In general, the exponential and the power function yielded a similar quality of fit whereas the hyperbola did clearly worse. As for the exponential case, young adults' cued recognition data were the primary source of misfit for the power function. In addition, old adults' figural scanning and young adults' figural reasoning data pose minor problems for the power function but the significance was only marginal ($p > .04$). When the four worst data sets (the same data sets identified for the exponential function) were left out the overall fit was acceptable for the power function ($p > .05$) but not the hyperbola. One statistic, finally, suggested that the exponential function may be superior to the power function, at least for figural scanning and figural reasoning conditions: in 60 of 80 intraindividual comparisons for these two conditions the χ^2 statistic was smaller for the exponential TAF. For word scanning and cued recognition the balance was 42 to 38 in favor of the power function. We conclude that the exponential function remains the best candidate for generating TAFs for a broad spectrum of cognitive tasks, but that the power function has too close a fit, in particular in word scanning and cued recognition, to consider the issue settled.[8]

Analyses of parameter estimates. The bottom block of Table 3 contains parameters a and b for the power functions. An analogous set of analyses as described for the exponential parameters led to the same conclusions: main effects of age in processing time implied larger effects of age on accuracy with an increase in presentation time. Moreover, the age difference in processing time was larger for the high-complex conditions.

State–trace analysis. Analysis of log-transformed processing times for the ratio of powers (Eq. 5) informs about process dissociation. As for the exponential function, ratios for low-complexity conditions (word scanning: 1.49; figural scanning: 1.44) were smaller than those of high com-

---

[8] The gradation of fit from exponential to power to hyperbolic function corresponds to the speed with which the asymptote is approached when the b parameter of the power function is smaller than one. As shown in the next section (Table 3), this was indeed the case for all four tasks.

plexity (cued recognition: 2.26; figural reasoning: 1.73). These slowing ratios are not to be confused with ratios characterizing proportional slowing such as those reported for the exponential TAFs; they are the powers in power functions. Despite their comparatively smaller values they will generate larger divergence than proportional slowing factors as presentation time is extended. The ANOVA of log-transformed b parameters yielded significant age $\times$ task-contrast interactions, $F(1,38) = 8.2$, $MSe = 0.09$, and $F(1,38) = 12.6$, $MSe = 0.01$. Thus, again, a single power function of cognitive slowing which ignores information about task complexity will not account for the age differences. Statistical comparisons of goodness of fit between alternative TAFs are limited because the functions are not nested within each other. Although we are not able to carry out a conclusive test of which TAF is most appropriate, the major results of an age-related dissociation of cognitive processing related to theoretical complexity of tasks generalizes at least across the functions considered here.

DISCUSSION

We demonstrated that TAFs can be used to model the relation between available processing resources and performance. Criterion-referenced testing was shown to yield complete functional relations between presentation time and the probability of correct responses for four different cognitive tasks at an individual level. Specifically, we determined functions for cognitively low-complex word scanning and figural scanning and cognitively high-complex episodic memory and reasoning tasks. Functions were estimated with reference to an exponential, to a hyperbolic, and to a power function. The exponential function was suitable for all four tasks with the power function a very close second in goodness of fit, and the hyperbola a slightly distant third.

We showed how TAFs can be used for tests of current models in the field of cognitive aging. Our findings replicate the well-established interactions between age and task complexity in the context of TAFs. In terms of presentation times needed for equal accuracy, age differences were larger for high-complexity than low-complexity tasks. Moreover, age differences increased within each task with the accuracy to be achieved. Of central theoretical interest is that not only absolute but also proportional age effects were larger in high-complexity tasks (cued recognition and figural reasoning) than in low-complexity tasks (figural and word scanning). This finding is counter to extant cognitive slowing models but is in agreement with studies showing large age effects in mnemonic performance (Kliegl et al., 1989; Thompson & Kliegl, 1991) and different slowing factors for sequential and coordinative complexity tasks (Mayr and

Kliegl, in press).[9] The twofold refutation of a simple slowing model considerably strengthens a modular view of cognitive aging. Apparently, for particular processes, such as elaborations based on mental images and coordination in working memory, effects of aging are stronger than for other, simpler information processing tasks.

The study of age effects on cognitive processing requires a large spectrum of complexity in various task domains. In this regard the joint time–accuracy platform provided by TAFs can avoid the shortcomings of earlier research focusing only on specific tasks; the method of limits (i.e., the systematic variation of accuracy criteria) constitutes an encompassing manipulation of task difficulty irrespective of type and domain of processing. An explication of potential relations between task-specific complexity, aging-related decline in processing time, and maximum attainable accuracy requires experimental paradigms such as the one introduced in this article; conclusions from traditional experiments based on ordinal interactions are always compromised by uncertainties regarding scale equivalence for different age groups and task-complexity conditions (Loftus, 1978).

In this respect the approach presented here reached beyond cognitive aging because the issue of process dissociation is of relevance for cognitive neuropsychology as well as for cognitive psychology in general. In the introduction we showed that state-trace analysis (Bamber, 1979), the principle of reversed association (Dunn & Kirsner, 1988), and young–old plots (Brinley, 1965; Cerella, 1990) all deal with the problem of delineating general and specific influences in cognitive processes. The present study differs from the earlier approaches in three important ways, all of which are related to the specification and determination of complete individual-based TAFs in each experimental condition. First, in previous applications, state-traces were inferred from a limited number of experimental conditions for different groups (for a recent application see Haist, Shimamura, & Squire, 1992). In the present approach, the complete functional range of accuracy between chance and asymptotic maximum performance enters the analysis; presentation time is specified as a continuous independent variable. This provides a much more solid basis for testing dissociations of cognitive processes. Second, we determined TAFs for individuals. Individual differences in these parameters capture dynamic aspects of cognitive functioning; traditional performance indicators are static descriptions. Such an experimental strategy may also help

---

[9] Mayr and Kliegl (in press) also showed that nonlinear young–old functions derived from assumptions about age differences in organizational overhead (Cerella, 1990) or information loss (Myerson et al., 1990) account for less variance than a bilinear model based on two complexity-specific proportional slowing factors.

focus theoretical analyses on the level of individual persons and holds promise for diagnostic purposes (Shallice, 1988). Third, in the present approach we started with an explicit assumption about the relation between presentation time and accuracy (i.e., exponential, hyperbolic, or power). State traces were analytically derived from this specification and therefore, conclusions about process dissociations are only valid with respect to the underlying TAF. At the same time, the approach provides a coherent framework within which extant models can be refined and alternative ones can be developed.

Limitations

The lack of fit of cued recognition data, in particular those of young adults, points to a possible limitation of the replacement model generating the exponential function. Cued recognition differed from the other conditions in an important way that may have jeopardized the validity of the assumption that cognitive processing is invariant across the entire range of accuracy (or presentation times). As distracters in the two-alternative, forced-choice response format were taken from the current list of words to be remembered, recognizing a distracter as a word that belonged to a different cue increases the probability of a correct response for words not remembered. This facilitation is more likely to occur for high levels of recall (e.g., the 100% condition of criterion-referenced testing) compared to low levels of recall. Moreover, even with explicit instruction old adults may be less able to use this strategy than young adults. The particularly low fit of young adults' memory data, then, may be a reflection of different cognitive processes entering at higher levels of functioning. The replacement model underlying the exponential function would need to be extended to take this possibility into account. One direction to explore would be the modeling of cascades of cognitive processes as envisioned by McClelland (1979). This approach assumes that cognitive performance on a task is the consequence of a number of parallel, possibly exponential component processes. One such additional process might describe the increasing benefits of sophisticated guessing for higher levels of accuracy. The integration of the two processes may lead to different forms of negative acceleration than considered so far.

Another weakness of this study was that participants knew the criterion for determining presentation time in the next block of trials or list of words. They were aware how their performance on the current block would affect the next one. We instructed them to try for 100% performance, but we cannot rule out some unwanted influences of the transparency of the test situation. Other procedures, such as double random staircases, continuous adjustments of presentation times after each trial,

and the simultaneous determination of several accuracy levels within a block of trials, will be given consideration in future studies.

For all tasks described in this article, performance depended on presentation time. This dependency is critical for this type of assessment. Thus, while the paradigm provides for a homogeneous measurement space for a broad spectrum of cognitive tasks, some of them were not represented (e.g., tasks probing and priming the structure of semantic memory). However, in principle, TAFs can be determined for any task which requires processing under experimentally controlled time constraints. This would also apply, for example, to the accuracy in synchronizing complex motor movements with an external auditory pulse rate.

We assumed that presentation time and accuracy are related by a negatively accelerated function. We investigated the exponential, the hyperbola, and the power function (Mazur & Hastie, 1978; Newell & Rosenbloom, 1981). For the tasks of this study, Mazur and Hastie's conclusion that the hyperbola would fit better than the exponential function could not be corroborated. One possible reason for this is that their analyses focused mostly on improvement across trials or sessions whereas we modeled the accumulation of information at the microlevel of a trial. The slight edge of the exponential over the power function for figural scanning and figural reasoning appears to hold up in ongoing research. For word scanning and cued recognition we could not detect an advantage of one over the other. The differences were not clear enough to warrant strong inferences about the underlying cognitive processes, although we feel fairly confident in rejecting the hyperbola. There are several reasons why it would be desirable to clarify these relations in future work. First, different functions imply specific assumptions about how a cognitive process model would convert presentation time into accuracy gains. Second, it is possible that different functions may be required for different tasks (or for different groups of people). Again, conclusive evidence along this line would provide important constraints for the construction of process models. Third, once the time–accuracy function is known, state traces can be derived analytically. If there were clear empirical evidence in favor of one of the TAFs across tasks and groups, logical constraints on the young–old functions rule out an entire family of alternative functions. For example, with clear evidence for a better fit of the exponential over the power function, the associated young–old function must be linear (i.e., Eq. 2); it cannot be a power function (i.e., Eq. 5). This is a point of contention in cognitive aging research because on the basis of several metaanalyses Myerson et al. (1990) claimed that young–old functions (within and across tasks) are best described by a power function (see also Hale, Myerson, & Wagstaff, 1987). Consequently, from this perspective, the power function should provide a better fit than the exponential or hyperbolic functions.

(Note, however, that there may be other TAFs than the power function that give rise to the young–old power function of Eq. (5), just as both the exponential and hyperbolic TAF implied the same young–old linear function of Eq. 2.) Obviously, the greater the validity of the time–accuracy function, the more conclusive will be tests of process dissociation.

The best strategy to get clarification in this matter is an increase in observational density. Also, subsequent to criterion-referenced testing, shorter and longer presentation times than those captured by the adaptive procedure should be administered to tie the function with respect to parameters indicating process initialization and asymptotic maximum performance. This way function-implicit differences in parameter b might be reflected in the goodness-of-fit; basically, the function should have as little room to move as possible. As a first step in this direction we had added a set of four "fast" blocks of items at the end of each task (see Table 1). Inclusion of these data, however, did not change any of the conclusions, probably because the time segment between zero and parameter a was not sampled systematically and frequently enough. While there is room for improvement in the estimation of time–accuracy functions, we do not think that these limitations were consequential for the theoretical issues addressed in this paper.

Cognitive Operating Characteristics

Psychometric functions describing perceptual thresholds in psychophysics or iconic memory (e.g., Loftus, Duncan, & Gehrig, 1992), the identification of test tones as a function of the silent interval between test and masking tones (e.g., Massaro, 1970; Massaro & Burke, 1991), or visual search times (e.g., Zacks & Zacks, 1993) are similar to the conceptual and methodological approach presented here. There are two other precursors of TAFs in the cognitive domain: the speed–accuracy tradeoff function and the performance–resource function. Indeed, Eq. (1) corresponds to Wickelgren's (1977) proposal (see also Lohman, 1986, 1989). A common assumption is that time limits will lead to a decline in performance. The main difference is that we are not determining the relation between response time and accuracy but between presentation time and accuracy. This has two important consequences: (1) on any trial, speed–accuracy research collects two performance measures (i.e., response time and accuracy), whereas in time–accuracy research presentation time is an independent variable and accuracy the sole dependent variable; (2) in the TAF-case, participants are instructed *always* to try for the maximum possible score—speed of responding is deemphasized. In speed–accuracy tradeoff research, subjects are either instructed to aim at a prespecified level of accuracy by increasing or decreasing their response speed or by inducing deadlines. Both procedures are problematic in practice because

participants' adherence to these instructions is not under experimental control. Another strong advantage of the TAF approach is its compliance with statistical assumptions. The usual estimation procedures assume error-free measurement of the independent variable. This is true for presentation times in time–accuracy functions but not for response times in speed–accuracy functions.

Two other advantages of TAFs over response–time tasks are worth mentioning in this context. First, we bridged the gap between typical response–latency and typical accuracy tasks. For example, performance in episodic memory was brought into the same measurement space as visual or verbal scanning. Thus, the scope of cognitive processes investigated in this paradigm is larger than that of typical response time tasks. Second, we measured cognitive processing avoiding a confound with a motor component (such as a button push) normally involved in response time tasks, that is the type of tasks normally included in metaanalyses. Both of these advantages are of relevance in developmental research where different slowing factors may prevail for simple and complex tasks and for central and peripheral processes (Cerella, 1985). For a comparison and discussion of the advantages of forced-choice procedures in combination with criterion-referenced presentation times over traditional response–latency assessments, see also Zacks and Zacks (1993).

The second relative of the time–accuracy function is the performance–resource function. In the paper which introduced this concept, Norman and Bobrow (1975) wrote:

> To determine a task's performance–resource-function we need to vary processing resources systematically while measuring some aspect of the task performance. This is not easy to do. In fact, we have been unable to find any instances of experiments in the literature that allow us to illustrate actual performance–resource functions. The major difficulty comes with the control of the resource allotment. (p. 52)

Norman and Bobrow (1975) were primarily concerned with cognitive time-sharing, that is, with problems of internal resource competition between several tasks. One criticism of their conceptualization was that performance–resource functions would need to be determined for each of the tasks in isolation and in combination in order to model competition of resources in this framework (Kantowitz & Knight, 1976). TAFs are single-task, single-subject operating characteristics with presentation time as an external resource; the paradigm is extendible to theoretical issues of cognitive time-sharing typically addressed by dual-task procedures. Also similar to the performance–resource functions envisioned by Norman and Bobrow, time–accuracy relations are not restricted to negatively accelerated functions. If a precise theoretical model about the relation of the external time resource to the accuracy of performance were available, in

principle, any function is imaginable (e.g., functions involving steps, plateaus, or segments with different accelerations). For each TAF, the state trace can be derived. This derivation may suggest statistical tests of the null hypothesis that the same state trace accounts for performance in different experimental conditions. Whether an ordinal interaction is compatible with a single underlying factor or not becomes a question that can be decided empirically.

In conclusion, we established (a) that time–accuracy functions can be determined at an individual level for tasks covering a broad spectrum of cognitive processes, (b) that ratios of old and young processing times depend on task-specific assumptions about the complexity of cognitive processing, and (c) that statistical tests for dissociations of cognitive processes avoid the scale-related interpretational ambiguities associated with ordinal interactions. The paradigm promises to be useful for investigating other person/group and process differences in cognition.

REFERENCES

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28,* 121–125.

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19,* 137–181.

Birren, J. E. (1965). Age changes in speed of behavior: Its central nature and physiological correlates. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system* (pp. 191–216). Springfield, IL: Charles C. Thomas.

Bogartz, R. S. (1976). On the meaning of statistical interactions. *Journal of Experimental Child Psychology, 22,* 178–183.

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system* (pp. 114–149). Springfield, IL: Charles C. Thomas.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98,* 67–83.

Cerella, J. (1990). Aging and information processing rates in the elderly. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 201–221). New York: Academic Press.

Cerella, J. (1991). Age effects may be global, not local. *Journal of Experimental Psychology: General, 120,* 215–223.

Cerella, J., Poon, L. W., & Williams, D. (1980). Age and the complexity hypothesis. In L. W. Poon (Ed.), *Aging and the 1980's* (pp. 332–340). Washington, DC: American Psychological Association.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences.* New York: Lawrence Erlbaum Associates.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review, 95,* 91–101.

Fisher, D. L. (1992). *Models of general, task-specific, and process-specific slowing.* Paper presented at the Fourth Cognitive Aging Conference, Atlanta, GA.

Fisk, A. D., Fisher, D. L., & Rogers, W. A. (1992). General slowing alone cannot explain

age-related search effects: Reply to Cerella (1991). *Journal of Experimental Psychology: General,* 121, 73–78.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 18, 691–702.

Hale, S., Myerson, J., & Wagstaff, D. (1987). General slowing of nonverbal information processing: Evidence for a power law. *Journal of Gerontology,* 38, 131–136.

Hertzog, C. (1991). Aging, information processing speed, and intelligence. In K. W. Schaie (Ed.), *Annual Review of Gerontology and Geriatrics* (Vol. 11, pp. 55–79). New York: Springer.

Kantowitz, B. H., & Knight, J. L. (1976). On experimenter-limited processes. *Psychological Review,* 83, 502–507.

Kliegl, R., & Lindenberger, U. (1993). Modeling intrusions and correct recall in episodic memory: Adult age differences in encoding of list context. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 19, 613–637.

Kliegl, R., & Mayr, U. (1992). Commentary. *Human Development,* 35, 343–349.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology,* 25, 247–256.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the locus and process of magnification of adult age differences during mnemonic training. *Developmental Psychology,* 26, 894–904.

Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging,* 6, 416–425.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition,* 6, 312–319.

Loftus, G. R., Duncan, J., & Gehrig, P. (1992). On the time course of perceptual information that results from a brief visual presentation. *Journal of Experimental Psychology: Human Perception and Performance,* 18, 530–549.

Loftus, G. R., Truax, P. E., & Nelson, W. W. (1987). Age-related differences in visual information processing: Qualitative or quantitative? In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 59–77). Norwood, NJ: Ablex.

Lohman, D. F. (1986). The effect of speed–accuracy tradeoffs on sex differences in mental rotation. *Perception and Psychophysics,* 39, 427–436.

Lohman, D. F. (1989). Estimating individual differences in information processing using speed–accuracy models. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology* (pp. 119–163). Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review,* 77, 557–567.

Massaro, D. W., & Burke, D. (1991). Perceptual development and auditory backward recognition masking. *Developmental Psychology,* 27, 85–96.

Mayr, U., & Kliegl, R. (in press). Sequential and coordinative complexity: Age-based processing limitations in figural transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Mazur, J. E., & Hastie, R. (1978). Learning as accumulation: A reexamination of the learning curve. *Psychological Bulletin,* 85, 1256–1274.

McClelland, J. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* 86, 287–324.

Myerson, J., Hale, S., Wagstaff, D., Poon, L. W., & Smith, G. A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review,* 97, 475–487.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum Associates.

Norman, D. A., & Bobrow, D. G. (1975). On data limited and resource limited processes. *Cognitive Psychology, 7*, 44–64.

Restle, F., & Greeno, J. (1970). *Introduction to mathematical psychology*. Reading, MA: Addison Wesley.

Salthouse, T. A. (1978). *Age and speed: The nature of the relationship*. Unpublished manuscript. Department of Psychology, University of Missouri, Columbia.

Salthouse, T. A. (1985). *A theory of cognitive aging*. Amsterdam: North Holland Press.

Salthouse, T. A. (1992). Shifting levels of analysis in the investigation of cognitive aging. *Human Development, 35*, 321–342.

Salthouse, T. A., Kausler, D. H., & Saults, J. S. (1988). Investigation of student status, background variables, and feasibility of standard tasks in cognitive aging research. *Psychology and Aging, 3*, 29–37.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge Univ. Press.

SPSS, Inc. (1988). *SPSS-X user's guide* (3rd ed.). Chicago.

Thompson, L. A., & Kliegl, R. (1991). Adult age effects of plausibility on memory: The role of time constraints during encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 542–555.

Wickelgren, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*, 67–85.

Zacks, J. L., & Zacks, R. T. (1993). Visual search times assessed without reaction times: A new method and an application to aging. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 798–813.