CHILD DEVELOPMENT

Child Development, January/February 2012, Volume 83, Number 1, Pages 367–381

Children Are Not Like Older Adults: A Diffusion Model Analysis of Developmental Changes in Speeded Responses

Roger Ratcliff and Jessica Love
The Ohio State University

Clarissa A. Thompson
The University of Oklahoma

John E. Opfer
The Ohio State University

Children (n = 130; M_{age} = 8.51–15.68 years) and college-aged adults (n = 72; M_{age} = 20.50 years) completed numerosity discrimination and lexical decision tasks. Children produced longer response times (RTs) than adults. R. Ratcliff's (1978) diffusion model, which divides processing into components (e.g., quality of evidence, decision criteria settings, nondecision time), was fit to the accuracy and RT distribution data. Differences in all components were responsible for slowing in children in these tasks. Children extract lower quality evidence than college-aged adults, unlike older adults who extract a similar quality of evidence as college-aged adults. Thus, processing components responsible for changes in RTs at the beginning of the life span are somewhat different from those responsible for changes occurring with healthy aging.

Across a wide range of cognitive tasks, children's responses are consistently slower than those of adults. Whether mentally rotating an alphanumeric character or a flag (Hale, 1990; Kail, 1986), making a same–different judgment (Bisanz, Danner, & Resnick, 1979), judging the direction of an arrow (Hale, 1990), detecting an auditory tone (Manis, Keating, & Morrison, 1980), or just visually fixating on a light (Luna, Garver, Urban, Lazar, & Sweeney, 2004), response times (RTs) decrease from early childhood to adulthood. A major question is whether these RTs decrease due to the maturation of some central limiting mechanism (Kail, 1988, 1991; Luna et al., 2004), due to skill transfer (Stigler, Nusbaum, & Chalip, 1988), or due to factors such as semantic knowledge and strategy use that are correlated with age (Chi, 1977; Roth, 1983).

To address these issues, researchers have sought a common metric for measuring relative differences in speed across highly diverse tasks, with most investigators using m, the slope of the linear regression function in which younger subjects' responses are plotted against older subjects' responses (Brinley, 1965; Hale, 1990; Kail, 1991, 1986; Luna et al., 2004). The slope of this ''Brinley function'' is attractive because it appears to provide a precise quantification of age differences from early childhood to late adulthood, is highly consistent across tasks, and allows the differentiation of developmental functions from practice functions. For example, in his review of 72 published studies, Kail (1991) found that children's and adolescents' RTs increased linearly with adults' RTs in corresponding experimental conditions (R^2s ranged from .91 to .99), and the slope of the Brinley function declined exponentially, with 3- to 4-year-olds' responses 3.1 times slower than college students, 7-year-olds' 2.4 times slower, 10-year-olds' 1.8 times slower, and 14-year-olds' 1.3 times slower.

However, recent work in cognitive aging has shed new light on the function's traditional interpretation as ''the rate for processing the information load'' (Jensen, 2006, p. 82). As with children, older adults consistently perform more slowly than college-aged adults, and RTs of older and college-aged adults show a linear relation when plotted against one another (e.g., Myerson, Ferraro, Hale, &

Preparation of this article was supported by NIA Grant R01-AG17083 and NIMH Grant R37-MH44640 to the first author. We would like to thank Russ Childers for his programming assistance, as well as the parents, teachers, administrators, and students from St. Agatha Catholic School, Columbus Academy, Upper Arlington School District, Worthington School District (all from Ohio), and Bentworth School District (Pennsylvania).

Correspondence concerning this article should be addressed to Roger Ratcliff, The Ohio State University, Department of Psychology, 291 Psychology Building, 1835 Neil Avenue, Columbus, OH 43210. Electronic mail may be sent to ratcliff.22@osu.edu.

 0009-3920/2012/8301-0026
DOI: 10.1111/j.1467-8624.2011.01683.x

Lima, 1992). Within the framework of evidence accumulation models such as the diffusion model, Ratcliff and colleagues (Ratcliff, Spieler, & McKoon, 2000; Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2001; Thapar, Ratcliff, & McKoon, 2003) showed that the Brinley plot slope could be explained by more conservative decision criteria for older relative to younger participants rather than reduced evidence during the decision.

The diffusion model (Figure 1, top panel) represents decision making as a noisy accumulation of evidence over time. It divides the decision process into several components, including the quality of information extracted from stimuli (the drift rate, v); the criteria used to make a decision (0 and a, starting at z); nondecision time including encoding, memory access, and response execution (T_{er}); and estimates of the trial-to-trial variability in these components of processing. These components together produce traditional measures of processing speed, as well as predictions for accuracy and RT distributions for correct and error responses. The model has been found to offer an accurate

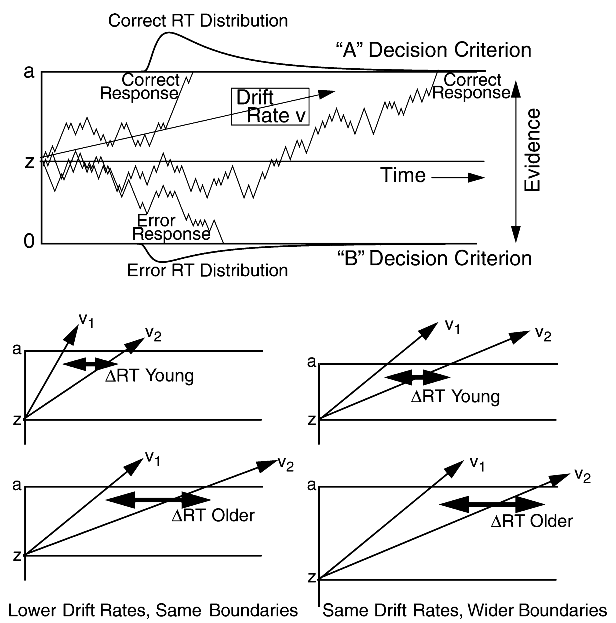

Figure 1. An illustration of the diffusion model.
Note. Parameters of the decision process are shown in the bottom panel: a = boundary separation, z = starting point, and v = drift rate. The bottom panel illustrates two different ways that young and older participants might differ in processing, yet produce the same difference in reaction times (RTs) between two conditions. The left-hand side shows an overall decrease in drift rates with the difference in drift rates held constant (although this distance could also increase). The right-hand side shows a larger boundary separation for older relative to college-aged adults.

explanation of accuracy and RT data when applied to a wide variety of perceptual and memory tasks, and populations that include aphasics, depressed and anxious participants, and participants under the effects of sleep deprivation or reduced blood sugar (Geddes et al., 2010; Ratcliff, Perea, Colangelo, & Buchanan, 2004; Ratcliff & Van Dongen, 2009; White, Ratcliff, Vasey, & McKoon, 2010).

Within this framework, the slope of the Brinley plot is a measurement of relative *spread* in RTs across conditions that can come about in different ways. Figure 1 shows two such scenarios. The left side shows the same distance between the starting point and the decision boundary, but lower drift rates for older relative to young adults. The right side shows the same drift rates for older and young adults, but with a greater starting point to boundary distance for older adults. In both cases, older adults have a larger difference in mean RTs between the two conditions. In addition, for the lower drift-rate case (left side), there would be a substantial difference in accuracy between the two conditions (e.g., Thapar et al., 2003), but in the equal drift-rate case (right side), this difference would be much smaller, with older adults performing only a few percentage points better because they allow more time for evidence accumulation (Ratcliff et al., 2001).

This reappraisal of the Brinley function has a number of implications for cognitive development. First, it means that the developmental patterns revealed by Brinley functions provide only a minimal constraint on explanations of age-related differences in RTs (see also Fisher & Glaser, 1996). Second, it means that the target for theories of cognitive development must be an understanding of the cognitive mechanisms underlying both accuracy and the entire RT distributions for correct and error responses for a cognitive task, not just the mean correct RTs that are plotted in Brinley functions (see also Siegler, 1987). Once the mechanisms are understood, it is then possible to examine how these mechanisms might change with age.

The diffusion model has been applied previously to data from older and younger adults for both tasks described in this article. For the numerosity discrimination task, Ratcliff et al. (2001; see also Ratcliff, Thapar, & McKoon, 2010) used Ratcliff's diffusion model (see Ratcliff, 1978; Ratcliff, Van Zandt, & McKoon, 1999) to model data from two experiments with college-aged adults and 60- to 74-year-old adults. Participants were instructed to decide whether the number of asterisks on a computer screen was ''large'' or ''small'' (Experiment 1) or whether the distance between two dots was ''large'' or ''small'' (Experiment 2), and responses and latencies were collected. Results showed that only some aspects of the response process were affected by age; namely, there was a 50-ms slowdown in the nondecision component of the RT, and the older adults adopted higher (more conservative) decision criteria. Critically, older adults did not show a decrease in the drift rate, the quality of evidence extracted from the arrays of asterisks or pairs of dots. Thus, although older adults did show slower (and more variable) RTs across conditions, there is no evidence that a deficit in central processing abilities is to blame. In fact, when both age groups were specifically encouraged to adopt a less conservative response style (i.e., when speed rather than accuracy was stressed in the instructions), the difference in RTs between the older and younger participants decreased. In a lexical decision task, where letter strings were presented and participants had to respond ''word'' or ''nonword,'' Ratcliff and colleagues (Ratcliff, Perea, et al. 2004; Ratcliff, Thapar, et al. 2004) also found an effect of age on the nondecision component and on boundary setting, but not an effect of age on the drift rate.

Applications of the diffusion model to data from older and young adults have thus posed serious challenges to theories of cognitive aging that propose that a decline in mental processing speed is responsible for the slowdown seen in RTs. In this article, we applied the diffusion model to children's data for the first time with three goals in mind. First, we wished to see whether the diffusion model could adequately fit children's speeded response data. Second, we wanted to determine which components of the decision process are responsible for the widely reported decreases in RTs associated with development. Third, we address the possibility that the same decision components are responsible for slowing on both ends of the developmental U-shaped curve. As Cerella and Hale (1994) comment in their review, ''The parallel between [the] conclusions on children, and those drawn earlier on elderly adults, is striking. They suggest that human cognition can be regarded quite generally as a rate-limited process, and that both maturation and senescence can be characterized as a transformation on processing rate'' (p. 127). Given the long-standing interest in this possibility—it has been described as ''one of the main questions for empirical inquiry'' (Jensen, 2006, p. 97)—we wanted to know whether children are slower than young adults because they are worse at extracting information from stimuli, or whether, like the older adults in Ratcliff and colleagues (2001; Ratcliff, Perea, et al. 2004; Ratcliff, Thapar, et al. 2004), they are slower for other reasons. In the sections that follow, we begin with a description of the diffusion model, and then we present two experiments. For both experiments, we first report the data using traditional methodologies, and we then follow with the diffusion model's analysis. To preview, our data suggest that children, unlike older adults, do in fact extract information from the stimulus (or from memory) at a lower rate than do college-aged young adults.

Why the Diffusion Model: Limitations of Traditional Analyses

Both of the tasks described in this article—a numerosity discrimination and a lexical decision task—are speeded response tasks that involve fast, two-choice decisions about stimuli that vary along a single dimension. Both tasks produce three dependent measures: accuracy, correct RTs, and error RTs. Researchers traditionally have focused either on accuracy data or on correct RTs, neglecting RTs for error responses. By excluding one or more of the dependent measures altogether, explanations are incomplete. Furthermore, sometimes RTs show a deficit, but accuracy does not. In the absence of a processing model, this can be seen as contradictory.

In the present study, we focus on the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008), a member of the class of sequential-sampling models. It uses correct and error RT distributions as well as accuracy information from behavioral data to estimate the relative contributions of each of the components involved in the task. In a speeded two-choice decision task, participants first encode the stimulus (e.g., build a visual representation of the asterisks or letter string), use encoded information to make a decision, and then make a key press.

As mentioned earlier, the actual decision process is represented as a noisy accumulation of evidence over time. For the numerosity discrimination task described in Experiment 1, a large positive or negative drift rate would indicate that the encoded representation of the asterisks is good evidence for either a ''large'' or ''small'' number of asterisks, respectively, whereas a smaller absolute value drift rate would indicate that the encoded representation provides poorer evidence. For the lexical decision task described in Experiment 2, a high positive or negative drift rate would indicate that the encoded representation of the letter string was either a very good or very poor match, respectively, to lexical memory. Thus, if one experimental condition is more difficult than another (e.g., the nonwords are pronounceable rather than random letter strings in lexical decision), the data in this condition would reflect a value of drift rate nearer zero. Because the drift rate represents the evaluation of evidence, it is the processing component that aligns most closely with traditional ideas of central processing. However, outside of the model, it is not sufficient to make RT predictions.

Within each trial, there is noise, or variability, in the information accumulation process, represented by the highly variable line in Figure 1. Thus, processes with the same mean drift rate may reach the same boundary at different times, producing RT distributions, or even reach different boundaries, thereby producing errors. In this way, the diffusion model predicts the positively skewed RT distributions that characterize behavioral data.

The diffusion model also models differences in participants' individual response styles, specifically addressing speed–accuracy trade-off and response bias. A conservative response style, which tends to produce slower but more accurate responses, is indexed by a large distance between the boundaries (a–0 or a in Figure 1; see also Ratcliff et al., 2001; Thapar et al., 2003). A liberal response style, producing quicker but less accurate responses, is represented with a small distance between the boundaries (Figure 1, bottom right panel). If a participant is biased toward one or the other responses, this is modeled by positioning the starting point (z) closer to that boundary (Ratcliff & McKoon, 2008; Voss, Rothermund, & Voss, 2004; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). Eliminating effects of an individual's response style allows for an uncontaminated estimate of the individual's information accumulation from stimuli (v).

Participants are assumed not to be able to set parameters at identical values from trial to trial. To model this, there are also three variability parameters, which represent variability in components of processing across trials: s_t is the range in T_{er}, s_z is the range of the starting point, and η is the standard deviation in the mean drift rate (for more on the variability parameters, including how the variability assumption allows the model to account for correct vs. error RT differences, see Ratcliff & Tuerlinckx, 2002; Ratcliff & McKoon, 2008).

Experiment 1: Numerosity Discrimination

Experiment 1 used a numerosity discrimination paradigm (e.g., Espinoza-Varas & Watson, 1994; Lee & Janke, 1964; Ratcliff et al., 1999; Ratcliff et al., 2001, 2010). Participants were instructed to decide whether a display of asterisks presented on a laptop contained a "small" or "large" number of asterisks, and they were presented with feedback after each trial. Responses and RTs were recorded. This task was chosen because it is very similar to experiments in Ratcliff et al. (2001, 2010), which allows us to compare any differences between children and young adults with those found between young adults and older adults.

Method

Participants. Three groups of students participated in the asterisk task during the middle and end of the academic year: 44 second and third graders (M = 8.51 years, SD = 0.64; 22 females), 36 fourth and fifth graders (M = 10.24 years, SD = 0.60; 21 females), and 24 ninth and tenth graders (M = 15.68 years, SD = 0.76; 14 females). Students were predominantly Caucasian and were recruited from elementary and high schools located in middle-class neighborhoods in central Ohio and southwestern Pennsylvania. Additionally, a comparison group of 31 college-aged adults participated in partial fulfillment of the requirements for an introductory psychology class at The Ohio State University. Although we do not have demographic information about the specific young adults who participated in our study, they were selected from a subject pool that was 62% female and had a mean age of 20.50 years (SD = 3.3).

Stimuli. White asterisks were displayed against a black background in a 10 × 10 grid in the upper left corner of a laptop computer monitor, subtending a visual angle of 4.30° horizontally and 7.20° vertically. There were 30 blocks, each of which contained 40 trials. For each block, the number of asterisks presented on a given trial was selected randomly without replacement from a sample of the numbers 31–70. Thus, each of the arrays in the 31–70 range was presented 30 times across each experimental session. The display positions for the asterisks in a given trial were selected randomly from the 100 possible positions in the 10 × 10 character grid. For arrays that contained 31–50 asterisks, the correct response was "small," whereas for arrays that contained 51–70 asterisks, the correct response was "large." The entire session, which consisted of 1,200 trials (i.e., 30 blocks with 40 trials each), lasted approximately 50 min. Between each block of trials, participants were encouraged to take a brief rest break if desired.

Procedure. Elementary school students (Grades 2–5) were tested in small groups of two or three in a quiet room at their school. Students were told they were going to play a computer game in which they would pretend they were workers at a chocolate factory. Their task was to determine whether each bag of chocolates (displayed on the screen) contained a small or large amount of candy. Children were told that the candy would look like tiny stars (i.e., asterisks), and were then shown examples of ''small'' (21 asterisks) and ''large'' (100 asterisks) amounts of candy on a paper printout that remained permanently displayed so that children could reference the examples at any time. Children were instructed to press the *Z* key with their left index finger if the bags contained a small amount of candy, to press the *?* key with their right index finger if the bags contained a large amount of candy, and to make their best guess if they were having difficulty in deciding.

Children were also told that if they correctly classified the bag as ''small'' or ''large,'' a smiley face would appear on the screen, but if they had classified the bag incorrectly, a frowning face would appear (using accurate feedback, as in Ratcliff et al., 2010). Students were instructed to earn as many smiley faces as possible, to make their judgments as quickly as possible, and to keep their index fingers over the appropriate answer keys throughout the game.

High school students and college-aged adults were given the same basic set of instructions as the elementary school students, but they did not hear the chocolate factory cover story. High school students were also told that they would receive corrective feedback in the form of smiley and frowning faces, but college-aged adults received feedback in the form of ''correct'' and ''error'' messages. As with the younger children, high school students and adults were instructed to make their best guess for difficult problems, to respond as quickly as possible, and to keep their index fingers over the appropriate keys throughout the experiment.

For all age groups, a trial began with the presentation of a 10×10 grid of asterisks. The asterisks remained on the screen until the participant made a judgment. After the participant indicated his or her response, the screen was erased and feedback appeared on the screen for 500 ms. The next trial began 400 ms later.

Results

We present the data analyses in two parts: (a) traditional analyses of mean RTs and accuracy, and (b) corresponding diffusion model parameters. We examined RT and accuracy for correct ''small'' and ''large'' responses and found that, for all age groups, they did not differ from each other systematically. For example, the probability of responding ''small'' to 31 asterisks was about the same as the probability of responding ''large'' to 70 asterisks. This allowed us to collapse the conditions for displaying the data (but not fitting the model).

Accuracy and mean RT data for Experiment 1 is reported in Figure 2. Overall, accuracy values were well above chance across ages, suggesting that even the youngest children (Grades 2–3) were capable of performing the task. A mixed-design 4×4 ANOVA was performed on correct mean RTs and accuracy values, with age (Grades 2–3, Grades 4–5, Grades 9–10, and college) as the between factor and task difficulty (1–4, see the note in Figure 2) as the within factor.

Response times. For correct mean RTs, the main effect of age was significant, $F(3, 128) = 88.91$, $MSE = 5.57 \times 10^6$, $p < .001$, with RTs decreasing with age. The main effect of task difficulty was significant as well, $F(3, 384) = 112.9$, $MSE = 2.45 \times 10^5$, $p < .001$, with RTs increasing with difficulty. The interaction between age and difficulty was not significant, $F < 1$, $MSE = 1{,}810$, $p = .586$.

Accuracy. For accuracy, both main effects were significant as well. The main effect of age was significant, $F(3, 128) = 14.34$, $MSE = 0.260$, $p < .001$, with accuracy increasing as age increased, and the main effect of task difficulty was significant, $F(3, 384) = 1632.4$, $MSE = 2.21$, $p < .001$, with accuracy decreasing as difficulty increased. The interaction between age and difficulty also reached significance, $F(9, 384) = 4.39$, $MSE = 0.006$, $p < .001$, with task difficulty affecting the older age groups (Grades 9–10 and college) more than the younger age groups.

Application of the diffusion model to experimental data. The diffusion model was fit to each individual participant's data, and parameter values were averaged across all participants in an age group and condition. For Experiment 1, all RTs < 300 ms and > 3,000 ms were excluded to eliminate fast guesses and slow outliers (about 9% of data for Grades 2–3 [40% of the eliminated data were fast guesses by just 4 participants], 5% of the data for Grades 4–5, 2.5% of the data for Grade 9, and 1% of the data for college-aged adults). For Experiment 2, the lexical decision experiment, the cutoffs were the same, and eliminated < 2% of the third graders' data and < 1% of the college-aged adults' data. The model was fit to the data by minimizing a chi-square

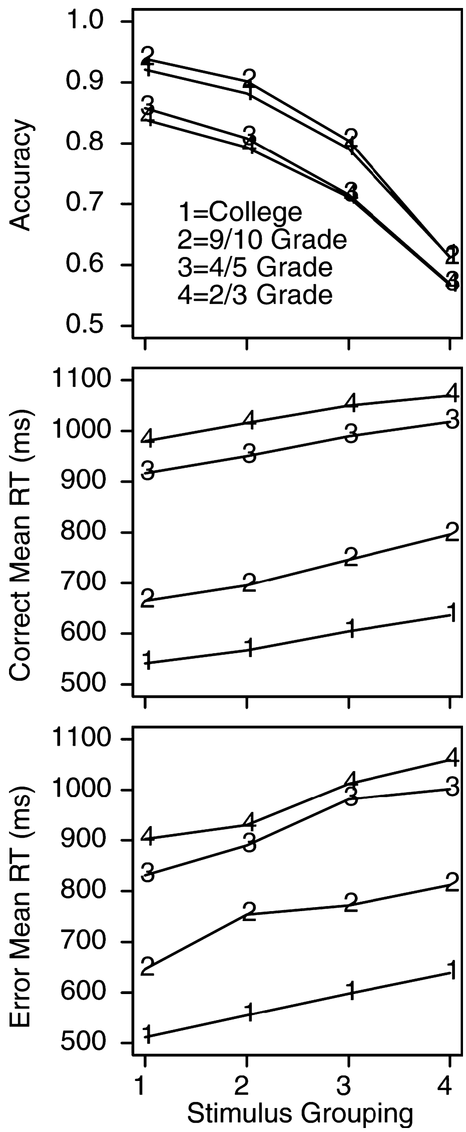

Figure 2. Accuracy and mean correct and error reaction time (RT) across participants for Experiment 1.
Note. Stimulus Group 1 collapsed ''small'' responses to 31–35 asterisks with ''large'' responses to 66–70 asterisks, Group 2 collapsed ''small'' responses to 36–40 asterisks with ''large'' responses to 61–65 asterisks, Group 3 collapsed ''small'' responses to 41–45 asterisks with ''large'' responses to 56–60 asterisks, and Group 4 collapsed ''small'' responses to 46–50 asterisks with ''large'' responses to 51–55 asterisks.

value with a general SIMPLEX minimization routine (Nelder & Mead, 1965) that adjusts the model parameters to find values that provide the minimum chi-square value (for a full description, see Ratcliff & Tuerlinckx, 2002). The data entered into the minimization routine for each experimental age/condition group were the RTs at each of the five quantiles for correct and error responses (representing the times at which .1, .3, .5, .7, and .9 of the responses had terminated), the accuracy values, and the number of observations.

The chi-square values used in the minimization routine provide a quantitative assessment of goodness of fit. Goodness of fit was measured with $(K \times 11) - M$ degrees of freedom, where K was the number of experimental conditions (eight for Experiment 1 and two for Experiment 2) and M was the total number of model parameters. For Experiment 1, mean chi-squares for each age group are between 114 and 139 (see Table 1), just above the critical value, 97.4 ($df = 76$). Mean chi-square values for Experiment 2 range from 29 to 40, just above the critical value of 22.4 ($df = 13$). However, the just significant values of chi-square do not necessarily reflect large inconsistencies in the model's predictions of the data (see Ratcliff, Perea, et al. 2004; Ratcliff, Thapar, et al. 2004; for a demonstration of how relatively small systematic deviations can lead to large increases in chi-square). These fits are comparable in quality to previous applications of the diffusion model (e.g., Ratcliff et al., 2001, 2010). In Figure 3, quantile probability plots comparing actual data to theoretical fits of the diffusion model are shown for each age group for Experiment 1. The model fits the data reasonably well for both experiments, so the parameter values can be meaningfully interpreted.

Diffusion model parameters. Mean values and standard deviations for the drift-rate parameters across participants in Experiment 1 are displayed in Table 1. Drift rates for individual participants are plotted in Figure 4. We will focus here on the parameters that are important for understanding developmental differences between childhood and young adulthood: the drift rate for each difficulty level (v), the criterion used to make a decision (the boundary separation [a]), and the nondecision components (T_{er}). We also address the variability parameters for each.

Most notably, the drift rates at each of the four difficulty levels did increase with age. A mixed-design 4×4 ANOVA run on the drift-rate parameters, with age (Grades 2–3, Grades 4–5, Grades 9–10, and college) as the between factor and difficulty level as the within factor, found that the main effect of age was significant, $F(3, 128) = 33.82$, $MSE = 0.661$, $p < .001$. The main effect of difficulty was also significant, $F(3, 384) = 434.4$, $MSE = 1.00$, $p < .001$. The interaction between age and task difficulty was significant as well, $F(9, 384) = 23.56$, $MSE = 0.0545$, $p < .001$, with task difficulty having a stronger effect on the drift-rate parameter among students in Grades 9–10 and college than among

Table 1
Mean and Standard Deviations in Diffusion Model Parameters Across Participants for Experiments 1 and 2

Expt.	Age group	a	z	T_{er}	η	s_z	s_t	p_o	v_1	v_2	v_3	v_4	v_c	χ^2
Numerosity	2–3	0.176	0.085	0.448	0.070	0.080	0.324	0.004	0.135	0.109	0.074	0.021	0.001	138
M	4–5	0.165	0.080	0.493	0.090	0.087	0.340	0.007	0.183	0.145	0.094	0.024	−0.008	125
	9–10	0.144	0.068	0.422	0.157	0.073	0.204	0.011	0.359	0.293	0.184	0.061	−0.009	114
	College	0.117	0.057	0.375	0.113	0.070	0.175	0.002	0.350	0.279	0.177	0.062	0.005	139
Numerosity	2–3	0.021	0.010	0.104	0.077	0.044	0.106	0.007	0.072	0.068	0.047	0.026	0.020	28
SD	4–5	0.030	0.015	0.069	0.096	0.036	0.065	0.013	0.119	0.094	0.102	0.018	0.026	42
	9–10	0.038	0.017	0.058	0.079	0.043	0.094	0.019	0.151	0.116	0.064	0.029	0.030	26
	College	0.020	0.012	0.032	0.061	0.031	0.029	0.010	0.116	0.090	0.057	0.024	0.028	120
Lexical *M*	Children pronounceable nonwords	0.119	0.063	0.528	0.049	0.057	0.246	0.049	0.233	−0.234				30
	Children random letter nonwords	0.158	0.093	0.528	0.077	0.069	0.216	0.081	0.178	−0.178				29
	College pronounceable nonwords	0.107	0.054	0.385	0.074	0.048	0.154	0.034	0.360	−0.329				30
	College random letter nonwords	0.125	0.073	0.425	0.138	0.050	0.159	0.078	0.316	−0.317				40
Lexical *SD*	Children pronounceable nonwords	0.015	0.012	0.054	0.044	0.027	0.060	0.088	0.064	0.054				16
	Children, random letter nonwords	0.034	0.021	0.055	0.051	0.040	0.087	0.153	0.048	0.072				10
	College pronounceable nonwords	0.018	0.013	0.030	0.065	0.025	0.033	0.017	0.059	0.069				18
	College random letter nonwords	0.027	0.019	0.029	0.073	0.022	0.054	0.250	0.081	0.077				20

Note. a = boundary separation, z = starting point, T_{er} = nondecision component of response time, η = standard deviation in drift across trials, s_z = range of the distribution of starting point (z), p_o = proportion of contaminants, s_t = range of the distribution of nondecision times. Drift rates for numerosity task are grouped as in the caption to Figure 2 and v_c is the drift criterion (drift rate for 31–35 asterisks is minus the drift rate for 66–70 m asterisks minus v_c). For lexical decision, drift rate v_1 is for words and v_2 is for nonwords. χ^2 is the chi-square goodness-of-fit measure with critical values of 97.4, df = 76, for numerosity discrimination and 22.4, df = 13, for lexical decision.

younger students. The 4×1 between-factor ANOVAs were run to see if age had a significant effect on any of the remaining model parameters of interest. There was a significant effect of age on boundary separation, a, $F(3, 128) = 31.01$, $MSE = 0.0227$, $p < .001$, with boundary separation decreasing with age. There was also a significant effect of age on the nondecision components, T_{er}, $F(3, 128) = 14.59$, $MSE = 0.0808$, $p < .001$, with the nondecision components also decreasing as age increased.

There was also a significant effect of age on drift-rate variability, η, $F(3, 128) = 6.47$, $MSE = 0.0415$, $p < .001$, with η increasing with age, and of variability in the nondecision component, s_t, $F(3, 128) = 35.15$, $MSE = 0.226$, $p < .001$, with s_t decreasing with age. There was no significant effect of age on the variability in starting point, s_z, $F(3, 128) = 1.22$, $MSE = 1.88 \times 10^{-3}$, $p = .304$. Overall, these results suggest that, from childhood to early adulthood, there are changes to all aspects of the decision process.

Discussion

Traditional analyses thus show strong effects of age on both RT and accuracy measures of performance. Collapsing across difficulty levels, mean RTs decrease from 979 ms (Grades 2–3) to 956 ms (Grades 4–5), to 748 ms (Grades 9–10), and to 604 ms (college). This speedup is accompanied by higher accuracy. These data are consistent with the developmental trends generally cited as evidence for an increase in the speed of processing.

The diffusion model analysis of data from Experiment 1 reveals that multiple components of the decision process are responsible for the speedup in RTs. Some of the age-related differences between

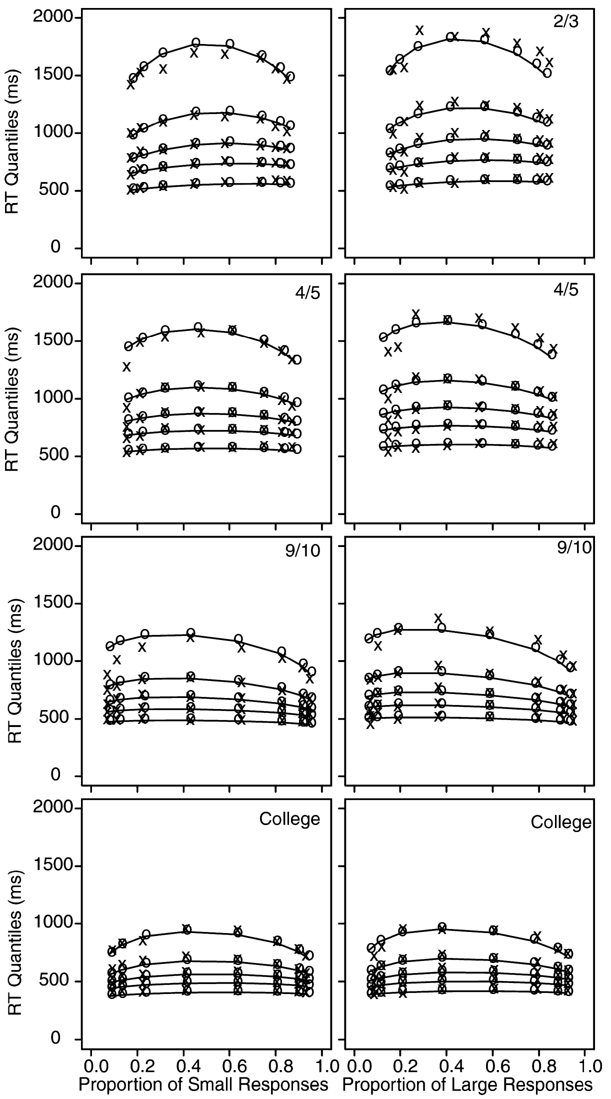

Figure 3. Quantile probability plots for the three tasks for data averaged over all subject groups.
Note. The *x*s are the data and the *o*s are the predictions joined by the lines. The five lines stacked vertically above each other are the values predicted by the diffusion model for the .1, .3, .5, .7, and .9 quantile reaction times (RTs) as a function of response proportion for the conditions of the experiments.

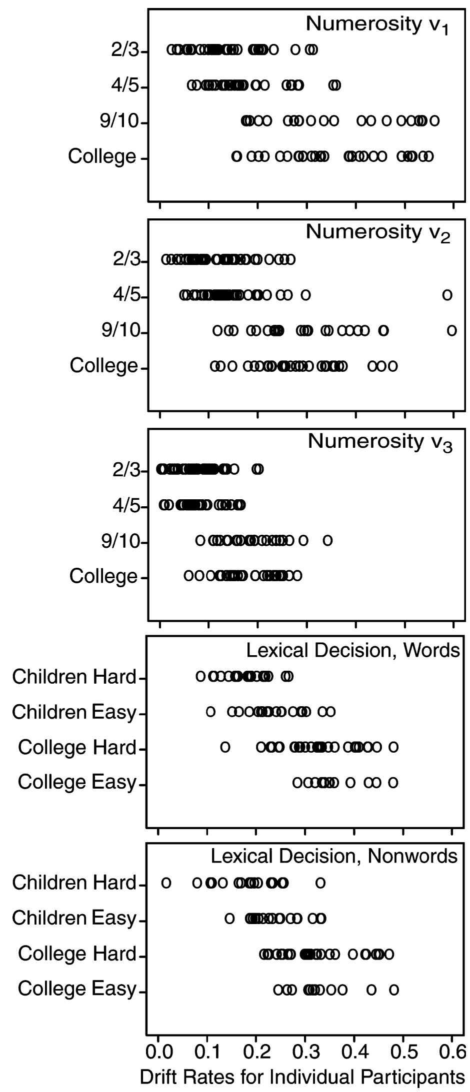

Figure 4. Drift rates for individual participants for Experiments 1 and 2.
Note. The plots for v_4 are not shown because performance is near floor for all individuals.

children and adults are explained by wider boundary separation, *a*, and longer nondecision components, T_{er}, for children relative to adults. Younger children have a more conservative response style, waiting for more evidence before making a response, and also spend more time on aspects of the response process that are not related to decision making, such as stimulus encoding and response execution.

However, the large significant effect of age on the drift-rate parameter suggests that the age difference in RTs between children and adults is also partly explained by differences in the ability to process information. More specifically, age differences in drift rate indicate that older adults extract numerosity information that is 2.5 times higher in quality than the information extracted by second- and third-grade children. Interestingly, the drift-rate parameter reaches adult-like values by Grades 9–10, although overall RTs are still slower (see Table 1). As Figure 4 demonstrates, the entire distribution of drift rates increases with age (i.e., it is not the case that this effect is driven only by some participants). Thus, results from Experiment 1 suggest that children do experience a large developmental increase in their ability to process meaningful, decision-related information in a numerosity discrimination task, unlike older adults who show no change in decision-related information (Ratcliff et al., 2001, 2010).

Experiment 2: Changes in Speed of Lexical Decisions

Experiment 2 used a lexical decision paradigm, where participants were instructed to decide whether a string of letters presented visually was a real word in English or not. The lexical decision paradigm has been used extensively in the adult literature, but it has not been used as often with children due to fears that developmental differences in performance may simply reflect children's reduced lexical knowledge. However, a number of standard lexical decision effects have been demonstrated with children, suggesting that the paradigm is more sensitive (and children's lexical knowledge more impressive) than is generally credited. For instance, Betjemann and Keenan (2008) found differences in phonological and semantic priming between children with reading disabilities and age-matched controls (average age: 11.5 years) using a visual lexical decision task. Duñabeitia and Vidal-Abarca (2008) and Acha and Perea (2008) also used the task to find robust neighborhood and length effects in children as young as third graders. The lexical decision task was chosen, in part, because of how different it is from the asterisk task; fitting data from two very different tasks (one of which may be especially challenging for children) provides a stringent test of the applicability of the diffusion model to children's data.

Experiment 2 had two conditions. In one condition, the ''easy'' condition, the nonwords consisted of random letter strings. In the second condition, the ''hard'' condition, the nonwords were pronounceable. The English words used were the same in both conditions. Feedback in the form of smiley and frowning faces was provided after each trial, and responses and response latencies were collected.

Method

Participants. Twenty-six predominantly Caucasian third graders (M = 9.29 years, SD = 0.36; 14 females) were recruited from two elementary schools located in middle-class neighborhoods in central Ohio to participate in the lexical decision task. A group of 41 college-aged adults from the same population as Experiment 1 served as a comparison group. These adults participated in partial fulfillment of the requirements of an introductory psychology class at The Ohio State University.

Stimuli. Letter strings, constituting words or nonwords, were displayed against a black background in the upper left corner of a laptop computer. The 352 English words used in Experiment 2 were selected nonrandomly from the first 400 of 1,000 Instant Words (Fry, 2004), which lists the most common words used in teaching children how to read, write, and spell. Words were selected from Fry (2004) in order to be as recognizable as possible to third graders, and indeed children's lexical decision accuracy (above 90% overall), as well as informal questioning after each session, suggests that they were sufficiently well known to the children. Words were selected nonrandomly so as to eliminate words containing apostrophes (e.g., *I'll*) and one- and two-letter words (e.g., *a*, *of*, *to*, *he*, *at*, *be*) in the test list. The words ranged from three to eight letters in length. For the third graders, they were divided randomly into two lists of 176 words, and the lists were counterbalanced across session order. For the college-aged adults, all 352 words appeared in a single session. Both third graders and young adults saw each English word exactly one time.

Each test list comprised an equal number of words and nonwords. For the easy condition, nonwords were random letter strings and unpronounceable (e.g., *xlsvrz*), whereas for the hard condition, nonwords were pronounceable in English, and were typically created by replacing one or more of the vowels in a standard English word (e.g., *merch*), or one or more of the consonants (e.g., *bipper*). Both types of nonwords ranged from three to eight letters in length.

Procedure. For the convenience of classroom teachers, students participated in two separate sessions, each lasting approximately 20 min. For half of these students (n = 13), both sessions were in the same condition (either easy or hard), and for the other half, the first session was in the easy condition and the second session was in the hard condition. The young adults completed both test lists in a single session that lasted approximately 40 min. Twelve young adults received the easy condition, and the remainder received the hard condition.

The third graders were tested one at a time for two sessions in a quiet room in their school. Instructions were the same for each session. Children were told that words would appear on a computer screen, that only some of them would be real words in English, and that others would be words from a different alien language. They were told that it was their job to determine whether the word was a real word in English, and they were encouraged to make their best guess if they remained uncertain. They were also instructed to press the *Z* key with their left index finger to indicate that ''No, the word is not a real word in English'' or the *?* key with their index finger to indicate that ''Yes, the word is a real word in English.'' The appropriate keys were modeled by the experimenter before each session. The children were told that if they responded correctly, a smiley face would appear on the screen, but that if they responded incorrectly, a frowning face would appear instead. They were instructed to earn as many smiley faces as possible. Then children were encouraged to make their decision as quickly as possible and, in order to stress this point, were told that it was better for them to make a few mistakes than to wait for several seconds before pressing a key. They were told to keep their fingers over the correct keys to help them go as fast as possible.

The same set of instructions was given to the college-aged young adults, except that no alien language was mentioned. They were told they would be seeing strings of letters on the screen, and their task was to determine whether the strings of letters made a real word in English. They were instructed to use the *Z* and *?* keys to make a response, to respond as quickly and accurately as possible, and to rest their fingers on the keys to facilitate this. They were also told that they would receive corrective feedback in the form of smiley and frowning faces.

For both age groups, the experiment began with 16 practice strings appearing one at a time in the upper left-hand corner of the computer. The word remained on the screen until a response was given. Then the screen cleared for 50 ms and either a smiley face (if the response was correct) or a frowning face (if incorrect) appeared on the screen for 500 ms. Then the screen cleared once more for 50 ms before the next word appeared. For the children, this practice was followed by eight blocks, each consisting of 44 strings (22 words and 22 nonwords). Each session lasted approximately 20 min. For the college-aged adults, who were given the strings from both sessions at once, practice was followed by 16 blocks, each consisting of 44 strings (22 words and 22 nonwords), and the experiment took approximately 40 min to complete. Between each block, all participants were encouraged to take a brief rest break if desired.

Results

As in Experiment 1, both children and college-aged adults performed with high accuracy. In the easy condition, with nonwords chosen from random letter strings, children responded correctly to 91% of words and 93% of nonwords. Even in the more difficult condition, where nonwords were pronounceable, children responded correctly to 91% of words and 86% of nonwords (compared to college-aged adults' 95% accuracy for words and 94% for nonwords). Given our concern that third graders might find the nature of the lexical decision task to be prohibitively difficult, their high accuracy is particularly reassuring.

Accuracy and RTs. Accuracy and mean RT data for Experiment 2 is presented in Table 2. As described in the Procedure section, 13 children participated in one difficulty condition (either easy or hard), whereas the other 13 children participated in both difficulty conditions (the easy condition for Session 1 and the hard condition for Session 2). Because all 13 of the students who contributed data to both difficulty conditions began in the easy condition and ended in the hard condition (i.e., difficulty was not counterbalanced across session order), we performed a post hoc analysis on their data to determine how much they improved with practice. Paired *t* tests revealed no significant

Table 2
Accuracy and Mean Reaction Time (RT) for Experiment 2

	Word stimuli			Nonword stimuli		
Participants and condition	Accuracy	Correct mean RT	Error mean RT	Accuracy	Correct mean RT	Error mean RT
Children, pronounceable nonwords	0.919	856.3	971.1	0.863	1,001.4	964.8
Children, random letter string nonwords	0.911	743.9	705.4	0.926	783.4	681.6
College age, pronounceable nonwords	0.945	585.1	637.6	0.942	658.4	637.0
College age, random letter string nonwords	0.950	531.9	489.8	0.966	549.0	463.2

improvement in performance from the first session to the second for either RT, $t(12) = 0.89$, $p = .38$, or accuracy, $t(12) = 1.3$, $p = .22$. Because we expected any practice effects to be, if anything, even smaller for the participants who switched conditions after the first session, we decided to collapse data across sessions, and to combine the data from same-condition and different-condition participants for the analysis. Although this means that some of the children are contributing twice as much data as others for a particular difficulty condition, the large amount of data collected in each session still allows us to produce individual diffusion model fits for all children.

A 2×2 ANOVA was run on accuracy values and mean RTs for words and nonwords, with age (third grade, college) and condition (easy, difficult) as between factors. For words, the main effect of age on accuracy values was significant, $F(1, 76) = 21.88$, $MSE = 0.0196$, $p < .001$, with college-aged adults responding more accurately than third graders. The main effect of condition was not significant, $F(1, 76) = 1.31$, $MSE = 1.18 \times 10^{-3}$, $p = .256$, and neither was the interaction between age and condition, $F(1, 76) < 1.0$, $MSE = 8.13 \times 10^{-4}$, $p = .345$.

There was, however, a significant main effect of condition on accuracy values for nonwords, $F(1, 76) = 6.08$, $MSE = 0.0225$, $p < .05$, with accuracy lower for the pronounceable nonwords. The main effect of age was also significant for nonwords, $F(1, 76) = 26.40$, $MSE = 0.0978$, $p < .001$. The interaction between age and condition does not reach significance, $F(1, 76) = 2.44$, $MSE = 9.04 \times 10^{-3}$, $p = .122$. Thus, traditional analyses found only the effect of condition difficulty in participants' responses to nonwords.

For words, the main effect of age on mean RT was significant, $F(1, 76) = 140.4$, $MSE = 1.10 \times 10^{6}$, $p < .001$, with responses faster for college-aged adults. The main effect of condition, however, did not quite reach significance, $F(1, 76) = 3.18$, $MSE = 2.49 \times 10^{4}$, $p = .079$. There was no interaction between age and condition, $F(1, 76) = 1.31$, $MSE = 1.03 \times 10^{4}$, $p = .256$.

For nonwords, the pattern of RT data is different. The main effects of both age and condition were significant, $F(1, 76) = 138.54$, $MSE = 1.67 \times 10^{6}$, $p < .001$ for age, and $F(1, 76) = 19.18$, $MSE = 2.32 \times 10^{5}$, $p < .001$ for condition. Mean RTs for nonwords were faster for college-aged adults than for third graders, and faster for random letter strings (easy condition) than for pronounceable nonwords (hard condition). The interaction between age and condition barely missed significance at the .05 level, $F(1, 76) = 3.79$, $MSE = 4.58 \times 10^{4}$, $p = .056$.

Diffusion model parameters. As for Experiment 1, we will only discuss the diffusion model parameter values most responsible for RT and accuracy differences. Means and standard deviations for diffusion model parameters across participants in Experiment 2 are displayed in Table 1.

Drift-rate parameters for college-aged adults were higher than for the third graders. A 2×2 ANOVA was run on the drift-rate parameters, with age (third grade, college) and condition (easy, hard) as between subjects factors, and found a significant effect of age for both words, $F(1, 76) = 71.09$, $MSE = 0.317$, $p < .001$, and, in a separate analysis, for nonwords, $F(1, 76) = 55.48$, $MSE = 0.270$, $p < .001$. Drift-rate parameters for individual participants are plotted in Figure 4. As in Experiment 1, the children extracted lower quality information from the match between the stimulus and memory than did the adults. There was also a significant main effect of condition on the drift-rate parameter for words, $F(1, 76) = 10.03$, $MSE = 0.0447$, $p < .01$, and, in a separate analysis, for nonwords, $F(1, 76) = 4.65$, $MSE = 0.0226$, $p < .05$. No significant interaction between age and condition was found for either words or nonwords (all Fs < 2), suggesting that our difficulty manipulation equally affected both third graders' and college-aged adults' ability to extract meaningful information from the letter strings.

A 2×2 ANOVA was also run on the other parameters of interest, with age and condition as between factors. There were significant main effects of both age, $F(1, 76) = 11.87$, $MSE = 8.10 \times 10^{-3}$, $p < .001$, and condition, $F(1, 76) = 22.41$, $MSE = 0.0153$, $p < .001$, on boundary separation, a, with children generally adopting more conservative (wider) boundaries than adults, and both children and adults adopting wider boundaries in the easy condition than in the hard condition. There was no significant interaction between age and condition, $F(1, 76) = 2.83$, $MSE = 1.93 \times 10^{-3}$, $p = .097$. There was also a significant main effect of age on the nondecision components, T_{er}, $F(1, 76) = 140.34$, $MSE = 0.263$, $p < .001$, with the duration of this component longer for children than adults. The main effect of condition barely missed significance at the .05 level, $F(1, 76) = 3.42$, $MSE = 6.42 \times 10^{-3}$, $p = .068$, as did as the interaction between age and condition, $F(1, 76) = 3.93$, $MSE = 7.37 \times 10^{-3}$, $p = .051$.

Age also had a significant main effect on our three variability parameters of interest: drift-rate variability, η, $F(1, 76) = 16.33$, $MSE = 0.0602$, $p < .001$; variability in the starting point, s_z (a measure of decision criterion), $F(1, 76) = 4.73$, $MSE = 3.98 \times 10^{-3}$, $p < .05$; and variability in the nondecision components, s_t, $F(1, 76) = 26.34$, $MSE = 0.106$, $p < .001$. Children had less variability in drift rate but more variability in starting point and nondecision component, from trial to trial. Condition had a significant main effect only on drift-rate variability, η, $F(1, 76) = 9.59$, $MSE = 0.0353$, $p < .01$, with the value of η higher in the easy condition than the hard condition, and the interaction between age and condition had no significant effects on any of the variability parameters (all Fs < 1.6).

Discussion

In general, these results parallel those from Experiment 1. Third graders were a little less accurate than college-aged adults, and their responses were slower. They responded to our difficulty manipulation (i.e., random letter string nonwords vs. pronounceable nonwords) in a similar manner to adults and became both slower and less accurate on the nonwords, but not the words. In short, traditional data analyses demonstrate systematically poorer performance for third graders than for college students.

The diffusion model analysis also mirrored that from Experiment 1: Third graders were slower at lexical decision than were college-aged adults because they adopted a more conservative decision criterion (i.e., required more evidence to make a decision) and because they were slower at stimulus encoding, memory access, and response output. Critically, they also extracted poorer information from the letter strings, with drift rates about 1.5 times lower (unlike older adults; Ratcliff et al., 2001, 2010).

General Discussion

The results from Experiment 1 showed the rise in accuracy from children to college-aged adults (roughly increasing from 75% for participants in Grades 2–3 to 85% for college-aged participants) and a large decrease in mean RT (roughly falling from 1,000 ms for participants in Grades 2–3 to 600 ms for college-aged participants). Similarly, in the lexical decision task in Experiment 2, accuracy values increased from 90% to 95%, and the decrease in mean RT was from 900 to 550 ms. These results are consistent with many others in the literature.

In applying the diffusion model to children's data for the first time, we had three aims: to see whether the diffusion model could successfully capture children's speeded RT data, to discover which components of the decision process are responsible for performance differences between children and young adults, and to compare these components with those affected during healthy aging. First, our results show that the diffusion model can be extended to data from children in second and third grades: As our fits in Figure 4 and Table 1 demonstrate, the diffusion model was capable of fitting data from young children for both numerosity discrimination (Experiment 1) and lexical decision (Experiment 2) tasks.

Second, drift-rate parameters, as well as nondecision and boundary separation parameters, changed with age. Thus, all aspects of the decision process are responsible for performance differences. To further look at the relative contribution of each to RT differences in Experiment 1, we first set all diffusion model parameters to the values for the college-aged participants and then, one at a time, changed both the boundary separation and drift-rate parameters to the values for the second and third graders. Boundary separation alone changed mean correct RT from 594 to 975 ms, and drift rate changed mean correct RT from 594 to 644 ms. Therefore, most of the slowing for second and third graders versus college students was due to the children adopting more conservative decision criteria (e.g., 380 ms). But there were also contributions from drift rate (50 ms) and nondecision components (70 ms). Because increasing boundary separation increases accuracy, only the drift rate is responsible for the lower accuracy for the second and third graders relative to college students. For Experiment 2, the same analysis finds a 70-ms contribution of boundary separation, a 55-ms contribution of drift rate, and a 100-ms contribution of the nondecision component. Thus, for the two tasks, the size of the effects for the different components of processing differs. Any simple single-process account for the developmental changes in performance cannot accommodate these results (cf. Kail, 1986).

Our findings also pose a problem for theories of speed of processing that claim that the mechanisms responsible for developmental changes in performance and those responsible for individual differences within the same age group are separate (e.g., Anderson, 1992). In a recent set of studies, Ratcliff et al. (2010) tested three different age groups (college age, 60- to 74-year-olds, and 75- to 90-year-olds) with a range of IQ scores (83–146) on three different tasks: the numerosity discrimination and lexical decision tasks described previously and a word recognition task. Ratcliff and colleagues found an effect of age, but no effect of IQ, on the nondecision component and the decision criterion, and an effect of IQ, but no effect of age, on the drift rate. This body of research reinforces the fact that very different components of the response process can be responsible for longer RTs. Importantly, though, it also demonstrates that drift rate corresponds well to individual differences in IQ and developmental differences. Perhaps, this is not fatal for Anderson's (1992) theory, as the drift rate could presumably encompass both information processing and the strategy use that Anderson advocates. Any further decomposition of drift rate, however, would require detailed task-specific models (e.g., models that specify how numerosity judgments are extracted from asterisks, or models of lexical processing) as in Ratcliff (1981) and Smith and Ratcliff (2009).

We can draw definitive conclusions, however, regarding our third aim. Results from Experiments 1 and 2, combined with those from earlier studies by Ratcliff and colleagues, suggest that the developmental increase and subsequent slowdown of RTs recorded at both ends of the life span do not originate from the same source. For the numerosity discrimination task, children did experience a developmental increase in their ability to process meaningful, decision-related information. In contrast, the slowdown older adults experienced on the same task (Ratcliff et al., 2001, 2010) could be attributed to increased decision criteria and nondecision components, but not to a developmental decline in ability to process decision-related information. For lexical decision, the story is similar—although it should be noted that our stimuli were chosen to accommodate the average third grader's vocabulary and do not directly match the words and nonwords used in Ratcliff et al. (2010). In these tasks, we find that the quality of the information extracted from the stimulus—the drift rate—increases from early childhood until young adulthood, but does not decrease with normal aging. This set of results challenges any theory of speed of processing across the life span that relies on a single mechanism or set of processes to explain both development and aging. (For some tasks, though, there is a decrement in drift rate; see Thapar et al., 2003, which also makes any single process account impossible.)

Although the difference in drift rate is an important one, there are also some notable similarities in the response components of children and older adults. First, both groups were slower in the nondecision aspects of responding than were college-aged adults. As visual and motor development and decline are well documented in the literature (e.g., Faubert, 2002; Getchell, 2006; Stelmach, Goggin, & Amrhein, 1988), this is not surprising, although it need not imply a change in the speed of cognitive processing. Second, both children and older adults adopted more conservative decision criteria than do young adults. Unlike the drift-rate parameter (v), which participants are generally unable to change at will, the boundary separation parameter (a) indicates response style, and the specific instructions given (e.g., instructions that stress either speed or accuracy) *can* influence participants' response styles quite drastically (Ratcliff et al., 2001; Thapar et al., 2003). Thus, the adoption of wider boundaries (i.e., a conservative decision criterion) by both young children and older adults should be considered *a preference* rather than an inherent characteristic of the developing information-processing system. Starns and Ratcliff (2010) suggest that older adults are unwilling to make avoidable errors (such as accidentally pressing the wrong key) because they prioritize accuracy. That young children behave similarly, though, suggests that both age groups may be more self-conscious about, or less confident in, their ability to perform a new task well and so play it safe by adopting conservative speed-accuracy decision criteria.

Implications of Diffusion Model Analysis

Our analysis using the diffusion model has several theoretical and methodological implications regarding developmental changes in speeded responses. Theoretically, the diffusion model analysis reveals that the U-shaped decline and rise in RTs over the life span mask a crucial difference between children's and older adults' RTs, as well as an unexpected similarity. The crucial difference is that in these tasks, children showed lower rates of evidence accumulation than college students, whereas drift rates for older and college-aged adults were quite similar. The surprising similarity is that children—like older adults—adopt much more conservative decision criteria than college students. This suggests that it may be possible to greatly reduce the relative difference between children and college-aged adults simply by providing children with instructions that emphasize speed and by providing adults with instructions that emphasize accuracy. Finally, on a methodological note, we believe that the diffusion model is useful for analyzing *all* components of the decision process. In our view, use of the diffusion model represents an exciting direction for future research on child development.

References

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, *99*, 245–264.

Anderson, M. (1992). Intelligence and development: A cognitive theory. Oxford, UK: Blackwell.

Betjemann, R. S., & Keenan, J. M. (2008). Phonological and semantic priming in children with reading disability. *Child Development*, *79*, 1086–1102.

Bisanz, J., Danner, F., & Resnick, L. B. (1979). Changes with age in measures of processing efficiency. *Child Development*, *50*, 132–141.

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: Thomas.

Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life-span. *Acta Psychologica*, *86*, 109–197.

Chi, M. T. H. (1977). Age differences in the speed of processing: A critique. *Developmental Psychology*, *13*, 543–544.

Duñabeitia, J. A., & Vidal-Abarca, E. (2008). Children like dense neighborhoods: Orthographic neighborhood density effects in novel readers. *Spanish Journal of Psychology*, *11*, 26–35.

Espinoza-Varas, B., & Watson, C. (1994). Effects of decision criterion on response latencies of binary decisions. *Perception and Psychophysics*, *55*, 190–203.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology*, *56*, 164–176.

Fisher, D. L., & Glaser, R. A. (1996). Molar and latent models of cognitive slowing: Implications for aging, dementia, depression, development, and intelligence. *Psychonomic Bulletin & Review*, *3*, 458–480.

Fry, E. (2004). *1000 instant words: The most common words for teaching reading, writing, and spelling*. Westminster, CA: Teacher Created Resources.

Geddes, J., Ratcliff, R., Allerhand, M., Childers, R., Frier, B. M., & Deary, I. J. (2010). Modeling the effects of hypoglycemia on a two-choice task. *Neuropsychology*, *24*, 652–660.

Getchell, N. (2006). Age and task-related differences in timing stability, consistency, and natural frequency of children's rhythmic, motor coordination. *Developmental Psychobiology*, *48*, 675–685.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development*, *61*, 653–663.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Oxford, UK: Elsevier.

Kail, R. (1986). The impact of extended practice on rate of mental rotation. *Journal of Experimental Child Psychology*, *42*, 378–391.

Kail, R. (1988). Developmental functions for speeds of cognitive processes. *Journal of Experimental Child Psychology*, *45*, 339–364.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.

Lee, W., & Janke, M. (1964). Categorizing externally distributed by stimulus samples for three continua. *Journal of Experimental Psychology*, *68*, 376–382.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*, 1357–1372.

Manis, F. R., Keating, D. P., & Morrison, F. J. (1980). Developmental differences in the allocation of processing capacity. *Journal of Experimental Child Psychology*, *29*, 156–169.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging*, *7*, 257–270.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308–313.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*, 552–572.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.

Ratcliff, R., Perea, M., Colangelo, A., & Buchanan, L. (2004). A diffusion model account of normal and impaired readers. *Brain and Cognition*, *55*, 374–382.

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin and Review*, *7*, 1–25.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical decision task. *Psychology and Aging*, *19*, 278–289.

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, *16*, 323–341.

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin and Review*, *9*, 438–481.

Ratcliff, R., & Van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin and Review*, *16*, 742–751.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261–300.

Roth, C. (1983). Factors affecting developmental changes in the speed of processing. *Journal of Experimental Psychology*, *35*, 509–528.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, *116*, 283–317.

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*, 377–390.

Stelmach, G. E., Goggin, N. L., & Amrhein, P. C. (1988). Aging and the restructuring of precued movements. *Psychology and Aging*, *3*, 151–157.

Stigler, J. W., Nusbaum, H. C., & Chalip, L. (1988). Developmental changes in speed of processing: Central limiting mechanism or skill transfer? *Child Development*, *59*, 1144–1153.

Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, *18*, 415–429.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory and Cognition*, *32*, 1206–1220.

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*, 140–159.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*, 39–52.