Journal of Experimental Psychology: Applied
2007, Vol. 13, No. 4, 224–248

1076-898X/07/$12.00 DOI: 10.1037/1076-898X.13.4.224

The Model Human Processor and the Older Adult: Parameter Estimation and Validation Within a Mobile Phone Task

Tiffany S. Jastrzembski and Neil Charness
Florida State University

The authors estimate weighted mean values for nine information processing parameters for older adults using the Card, Moran, and Newell (1983) Model Human Processor model. The authors validate a subset of these parameters by modeling two mobile phone tasks using two different phones and comparing model predictions to a sample of younger ($N = 20$; $M_{\text{age}} = 20$) and older ($N = 20$; $M_{\text{age}} = 69$) adults. Older adult models fit keystroke-level performance at the aggregate grain of analysis extremely well ($R = 0.99$) and produced equivalent fits to previously validated younger adult models. Critical path analyses highlighted points of poor design as a function of cognitive workload, hardware/software design, and user characteristics. The findings demonstrate that estimated older adult information processing parameters are valid for modeling purposes, can help designers understand age-related performance using existing interfaces, and may support the development of age-sensitive technologies.

Keywords: age, modeling, GOMS, human engineering, design, mobile phone

A hallmark of good design is represented by the human factors injunction to “know the user” (Nielson, 1994, p. 73) before design specifications are made. It is essential to account for basic human processing requirements in order to optimize performance with respect to time, accuracy, and satisfaction, in order to produce “user-friendly” systems, interfaces, or devices. Research in human factors and human-computer interaction has a legacy of designing products around the needs of the younger user, in part because many designers are young themselves (the largest population density of workers are between the ages of 30–34; National Science Foundation, 2006), and, as a result, older adults often find that they have difficulties interacting with existing technology and report feelings of frustration and confusion when trying to adapt (e.g., Fisk & Rogers, 1997; Rogers, Fisk, Mead, Walker, & Cabrera, 1996; Walker, Philbin, & Fisk, 1997).

As it currently stands, older adult competencies and limitations are not usually accounted for in these earliest stages of design and engineering. Some strides have been made to help accommodate older adults to existing systems through modification of interfaces and systems (e.g., changing acceleration of a mouse, enlarging font size, enhancing contrast), or through modification of the individual (e.g., adapting the user to the product with training and practice); yet a disconnect exists between the fields of cognitive aging and human engineering design. Human engineering design refers to testing systems with cognitive models that simulate human performance, so that designers may predict what design specifications will be most effective. These simulations help bypass costly, time-intensive user studies and may help prevent the construction or adoption of poorly designed products from the start.

Charness and Bosman (1990) attempted to assemble necessary parameters to simulate an older user for human engineering purposes, but due to the sparseness of literature available at the time were unable to provide reliable estimates of older adult abilities in a form that allowed for easy cognitive modeling. Thus, one objective of this research is to formalize the description of the cognitive and psychomotor performance of older adults so that the gap between the cognitive aging and human engineering design fields may be bridged. Our approach to this goal entails estimating information processing parameters from the ever-expanding cognitive aging literature to try to update parameters of the Model Human Processor (Card, Moran, & Newell, 1983) and simulating an older adult’s perceptual, motor, and cognitive functions at an elementary grain of resolution, the keystroke level. Through the attainment of these basic perceptual, motor, and cognitive building blocks, it may be possible to generalize to many types of routine tasks, interfaces, or systems. These parameters require expression in the form of a mean or typical value.

Thus, a second goal of this research is to develop a modified form of meta-analysis, not for effect size, but for generation of means for the parameters of interest originally obtained by Card, Moran, and Newell (1983). No published examples of such work exist to date, nor did Card, Moran, and Newell (1983) methodically gather studies (many estimations relied on only two or three data sets) or statistically estimate parameters using a validated technique (they calculated the median value of performance from limited pools of data and did not weight study sample size). Here we attempt to collect a representative sample of studies for inclusion in each meta-analysis (electronic database searches were

---

Tiffany S. Jastrzembski and Neil Charness, Department of Psychology, Florida State University.

This research was supported in part by the National Institute on Aging, NIA 1 PO1 AG17211-07, Project CREATE and served to fulfill in part the requirements for the doctoral dissertation for Tiffany S. Jastrzembski. A portion of this research was presented at the Cognitive Aging Conference, Atlanta, Georgia, April 2006.

Correspondence concerning this article should be addressed to Tiffany S. Jastrzembski, Department of Psychology, Florida State University, 1107 W Call Street, Tallahassee, FL 32306-4301. E-mail: tiffany.jastrzembski@mesa.afmc.af.mil

defined and essential citations from included works were also assessed) and obtain weighted means based on sample size and sample variance of sample means, in an attempt to capture valid parameters for the typical older adult.

The third goal for this project is to validate the estimated parameters. We will use them to predict both older and younger adult performance in two common mobile phone tasks. The aging literature is inconsistent about definitions of older groups, posing problems for attaining averages across groups in a consistent way. We chose to include only studies with participant means greater than 60 years of age for our meta-analyses of older adult parameters. If the parameter estimates that we derive fit older adult data closely and at least as well as do the earlier parameter estimates for young adults, we will consider this to be a validation of the meta-analyses. Ultimately, the overarching goal of this research is to provide older adult performance parameters to allow designers and engineers to simulate mean older adult performance and make predictions across varying design specifications and tasks. In this way, designs may be optimized for typical older adults during the very early stages of design.

Knowing the Older User

Over the recent past decades, an extensive literature in the field of aging has been amassed. Researchers have quantified age-related declines affecting performance with regard to cognitive and physical abilities (Craik & Salthouse, 2000) and have revealed compensatory strategies older adults may use to circumvent typical age-related slowing in specific tasks (Bosman, 1993; Charness, 1981a, 1981b, 1981c; Charness, Krampe, & Mayr, 1996; Krampe & Ericsson, 1996; Meinz, 2000; Salthouse, 1984). With proper application, this knowledge could inform the design of products, systems, and environments for more effective use by the older population.

As aptly stated by Welford (1958):

> Where age changes do impinge upon performance some relatively trivial factor may often be limiting what can be done, so that comparatively small changes in the task could bring it within the capacities of older people.. . .and would benefit both young and old. . .Attempts to ‘fit the job to the man’ in such ways would seem a far better approach. . .than attempting to move men to other jobs (p. 287).

Indeed, the fields of human factors and gerontechnology (Harrington & Harrington, 2000) have grown considerably. These disciplines attempt to optimize the relationship between technology and the human (Kantowitz & Sorkin, 1983) and aim to enhance “the development and design of products and services to increase the quality of life” (Bouma, 2001). They stress efficiency and ease of use of product design, arguing that the performance abilities of the user (such as age-related change) must be taken into consideration.

The study and practice of aging and human factors/engineering design is becoming evermore critical as the older population in America is expected to burgeon between the years 2010 and 2030, when the “Baby Boom” generation reaches age 65. According to the United States Census Bureau (2004), there will be about 71.5 million older persons by the year 2030, representing 20% of the population and a doubling in size compared to the year 2000.

Current trends of increasing affordability of computing and rapid adoption of technological innovation in developed countries are expected to continue to accelerate and spread (Jastrzembski, Charness, Hollay, & Feddon, 2005). Because there is less information available on general technology use, we now examine Internet use as a proxy for technology use. Recent studies show that older adult cohorts may comprise one of the fastest growing segments of the estimated 80 million Internet navigators in the United States, having jumped 47% from 2000–2004 (Fox, 2004); however, absolute use of computers and the Internet by those cohorts has remained well below that of younger cohorts with approximately one third of those aged 65 + using the Internet compared to about 80% or greater use for those 18–50 years of age (Pew Internet & American Life Project, 2006).

Although future cohorts of older adults, today’s baby boom generation, are likely to be quite familiar with and continue to use the Internet as they reach old age, the pace of introduction and diffusion of new communication technologies appears to be increasing (Charness & Czaja, 2005). Hence, we may expect to find an age-related technology divide for future older adult generations. Fisk, Rogers, Charness, Czaja, and Sharit (2004) have argued that older adults may be more likely to adopt new and existing technologies if interfaces and devices were better designed, so that discomfort and anxiety interacting with the system would be reduced. To encourage this, designers must pay particular attention to several age-related changes that affect older adults. These changes include more limited and more variable motor control and coordination (Benson & Marano, 1998; for a review, see Vercruyssen, 1997), greater likelihood of inhibitory process failure (Hasher & Zacks, 1988; Rabbitt, 1965), and less efficient working memory and decline in fluid intelligence (Salthouse, 1991). Additionally, general slowing translates to an older adult requiring approximately 1.5–2 times the length of time that a younger adult needs to complete a task (Hale & Myerson, 1995; Salthouse, 1996).

Better design of technology for older adults may be supported by bringing the knowledge of age-related performance to human engineering design, meaning that products and devices could be modeled and modified to suit the older user in the very earliest stages of design. Ideally, an older user, with known cognitive, perceptual, and motor capabilities (gleaned from empirical data in the literature), could be simulated so that designers might predict learning or task execution time for the typical older user, what the optimal design would be for the typical older user before prototypes are built, and what design iterations should be made to enhance performance for the typical older user. This type of modeling is deliberately approximate in order to account for most typical users (Card et al., 1983; John & Kieras, 1994) and helps answer many different types of questions a designer may ask, such as whether or not the typical user will achieve “good enough” performance, whether alternative designs would achieve better performance for the typical user, and how widespread performance differences will be based upon level of skill, as defined by parameter bounds estimating poor and better users plus or minus two standard errors from the typical user. These modeling ideals may then help designers or companies target specifically designed products to specific sectors of consumers, so that desirable features may be emphasized and performance may be enhanced for those target audiences (e.g., typical older users).

Card, Moran, and Newell (1983) sought to mesh psychology with the engineering approach to produce a simple model that could generate quantitative predictions for human performance and help designers make low-level design decisions. Their model integrated psychological knowledge of human perception and performance with the design process and translated those findings to a form suitable for human-computer interaction analyses. Specifically, they created a simulated user, the Model Human Processor, possessing three interacting subsystems (cognitive, perceptual, and motor) by extracting key parameters relating to general abilities of humans from existing psychological and human-computer interaction (HCI) literature. These parameters were drawn from literature covering both simple and complex psychological phenomena.

Each system within the Model Human Processor possessed a set of memories [μ - the storage capacity in items, δ – the decay time of an item, κ – the main code type (physical, acoustic, visual, semantic), and a set of processors (τ – the cycle time), which worked according to "principles of operation"; see Cardet al., 1983, for more detail]. These components could be used to approximate and predict human behavior in a specified task scenario (Card et al., 1983, 1986). Highly specific processing times, storage capacities, and rates of decay were computed for each system by gathering data for the most typical values gleaned from representative sources in the literature (to be most useful in most design conditions) and informed by known human cognitive processes (Welford, 1974). As these values could vary according to the task and the skill of the observer, it should be reiterated that this model was an attempt to provide estimates for the modeling community, and thus, values represent a mean value to account for most users. To include some of the variability aforementioned, Card, Moran, and Newell (1983, 1986) provided values for the upper and lower bounds of performance (by finding the worst and best individual performer across all studies) to allow designers to account for different population abilities dependent on the task or level of skill (see Table 1 for a list of parameters). In this way, it would be possible to simulate the performance of an amateur using a given environment or system and compare it to the performance of a more highly skilled user, or account for practice and training, such that the model could predict how long it would take a novice to become proficient user.

Because Card, Moran, and Newell's model was published in 1983, the parameters calculated for inclusion in the model stemmed from psychological research that existed at the time. These data were primarily based on the performance of younger adults, given that most experimental studies used convenience samples of college undergraduates as participants. Charness and Bosman (1990) sought to supplement their model with age-relevant information, but were limited as to what could be accomplished given the paucity of literature in cognitive aging at the time. Charness and Jastrzembski (2001) updated regression equations and constant values similar to Charness and Bosman's (1990) original update of the literature (see Table 2 below), but these estimates are quite different from Card, Moran, and Newell's (1983) Human Model Processor values and would not be as amenable to human engineering design as precise parameter values would be. Thus, this work seeks to refine previous work to create a model that may account for life span cognitive changes.

Table 1
Estimated Parameters Used in Model Human Processor (taken from Card et al., 1986)

Parameters of interest	Typical parameter estimates [lower and upper bounds]
Perceptual system-visual	
Eye movement (travel + fixation)	230 [70–700] ms
Memory decay time (visual)	200 [90–1000] ms
Capacity (visual)	17 [7–17] letters
Cycle time (response time to brief pulse of light)	100 [50–200] ms
Perceptual system-auditory	
Memory decay time (auditory)	1500 [90–0–3500] ms
Capacity (auditory)	5 [4.4–6.2] letters
Cycle time (time needed to differentiate fast clicks)	100 [50–200] ms
Motor system	
Cycle time (time needed to perform a discrete micromovement, as assessed by tapping rates)	70 [30–100] ms
Cognitive system	
Memory decay time (working memory-general)	7 [5–226] seconds
Memory decay time (WM-1 chunk)	73 [73–226] seconds
Memory decay time (WM-3 chunks)	7 [5–34] seconds
Storage capacity (pure working memory)	3 [2.5–4.2] chunks
Storage capacity (effective WM-augmented by LTM)	7 [5–9] chunks
Cycle time (time needed to match item against WM-calculated as typical value over many tasks (e.g. digits, colors, letters, words, shapes)-can be shortened by practice, task pacing, more effort, or reduced accuracy)	70 [25–170] ms

By updating Model Human Processor values to model the older user, several potential benefits could arise. Older adults may not need to perform tasks as quickly as possible due to no longer being in the workforce, so producing optimal designs with regard to error-free performance may be more appropriate than producing the most time-efficient designs for the older population. These designs may be more streamlined for all users by reducing the number of steps in any process to reduce the overall likelihood of making an error, or they may have better error recovery methods so that the user could resume normal processing with minimal interruption. Modeling different designs with a simulated older user might also make it possible to predict and anticipate the types of errors that would likely occur (Lerch, Mantei, & Olson, 1989) so that designers could foresee potential problem areas and devise ways to circumvent them.

In order to update the Model Human Processor to account for the older adult, a modified meta-analysis was performed, such that rather than computing effect sizes for each phenomena of interest parameter values for cognitive, perceptual, and motor processors were extracted. Methodologically speaking, the only precedent for performing this type of analysis stems from the original work of Card, Moran, and Newell (1983). Through correspondence with the first author of that book and through discussions with behavioral research statisticians, agreement was reached that a reasonable and statistically plausible way to update this model was to collect a representative sample of studies (making use of existing meta-analyses when available and any additional studies not included in those meta-analyses) comparable to the types of tasks

Table 2
Estimates of Performance by Age and Task (taken from Charness & Jastrzembski, 2001)

Model component	Value	Source
A. Perceptual system		
Eye fixation duration	Young = 207 ms; Old = 254 ms	Scialfa & Joffe (1997)
Visuospatial processing speed	RT Old = 2.27* RT Young	Zheng, et al. (2000)
Perceptual processor cycle time	Young = 100 ms; Old = 190 ms	Card, Moran, & Newell (1983); Meyer, et al. (2001)
B. Motor system		
Motor processor cycle time	Young = 70 ms; Old = 139 ms	Card, Moran, & Newell (1983);
Acquisition time	Young = 862 ms; Old = 1060 ms	Liao, Jagacinski & Greenberg (1997)
Movement time	Young = 641 ms; Old = 747 ms	
C. Cognitive system		
Verbal working memory span	Y = 7.4–.016* age (.11)	Jenkins, et al. (1999)
Cognitive processor cycle time	Young = 70 ms; Old = 139 ms	Card, Moran, & Newell (1983); Meyer, et al. (2001)
False alarm rate for recognition	Young = .05 (.07); Old = .09 (.11)	Titov & Knight (1997)
Choice RT-2	Young = 353 (46); Old = 692 (196)	Hertzog, Cooper & Fisk (1996)
Choice RT-4	Young = 375 (47); Old = 733 (196)	
Choice RT-8	Young = 430 (76); Old = 884 (209)	
Digits forward	Old = 7.2 (.86)	Rochan, Waters, Caplan (2000)
Digits backward	Old = 5.5 (1.2)	
D. General slowing effect		
Decay constant	OldRT = −.34 + 1.88 YoungRT (.38)	Verhaeghen & DeMeersman (1998); Sliwinski & Hall (1998)
Power law of practice constant	Young = .61 (.46); Old = .42 (.12)	Jenkins & Hoyer (2000)

Note. Values enclosed in parentheses indicate standard errors.

used in the original model. In this manner, typical values, defined as weighted mean values of meta-analytic sample means, for the older population were obtained. As was also provided in the original model, upper and lower bounds were estimated, defined as plus or minus two standard deviations of means, in order to encompass the performance of approximately 95% of older adults. These bounds were chosen because this type of modeling is meant to encompass the majority of typical users and to extend boundaries any further would certainly go outside the range of where typical users would likely perform. Furthermore, these bounds provide designers with a reasonable way to test design efficacy across slower and faster users, and allows them to decide whether or not design characteristics lead to performance that would achieve good enough performance for 95% of the target population. As such, these bounds can be seen as covering variation due to user experience and specific task features. In the next section of the paper, we detail specific age-related changes that affect each perceptual, motor, and cognitive parameter that we estimate for our model and provide a description of what each parameter is used for.

Model Parameter Estimation for the Older Adult

The Perceptual System

Aging is associated with a gradual decline in perceptual-motor functioning that affects performance on most visual tasks, largely accounted for by changes in the ocular media of the eye and other neural factors (Kline & Scialfa, 1996). Indeed, older adults tend to exhibit two forms of ocular slowing compared to younger adults, such that they lag further behind intended targets and they make smaller, slower movements overall (e.g., Jagacinski, Greenberg, Liao, & Wang, 1993; Welford, 1958, 1973) translating to prolonged saccadic fixation durations approximately 60 ms longer than those of their younger counterparts (Abel, Troost, & Dell'Osso, 1983; Huaman & Sharpe, 1993; Pirozzolo & Hansch, 1981; Scialfa & Joffe, 1997) and requiring more saccades onto the same target to extract the necessary information (Ho, Scialfa, Grau, & Caird, 2001; Scialfa & Joffe, 1997).

Literature searches to attain all of the following parameters were conducted for studies using the PsycINFO and Web of Science electronic databases across all years and by checking references found in the articles thus retrieved. All parameter searches were performed between April 2005 and April 2006. General criteria for inclusion in these meta-analyses required a participant sample of healthy, older adults (free from known psychological or physiological pathology as indicated by self-report), having normal or corrected to normal vision. Institutionalized individuals, those with dementia, and those with physical disabilities affecting motor abilities of the hand, did not meet inclusion criteria for these analyses. In order to provide estimates for the typical, older user, level of education was not used as an inclusion or exclusion criterion. For all searches, each keyword was used in conjunction with the Boolean search term, "AND," *older adults*, and each keyword was again combined in a separate search with the Boolean search term, "AND," *aging*. In the case of meta-analyses being identified in these searches, individual studies also retrieved in the literature search that overlapped in the meta-analysis were removed from the current investigation, to avoid being weighed multiple times. Selection of studies was then made by reading abstracts and ensuring that the dependent variable was appropriate. Irrelevant studies were eliminated at this point and relevant studies were read in their entirety. If the published data included results at the parameter-level needed, they were utilized in estimations just as they were. If the published data did not include results at the parameter-level needed, but did include the necessary data defined by the equations provided in the following sections, those studies

Table 3
Literature Comprising Meta-Analysis for Fixation Duration Parameter Estimation

Study	*n*	Mean age	Ages sampled	Mean fixation duration (ms)	*SE* (ms)
Bono, Oliveri, Zappia, Aguglia, Puccio, & Quattrone (1996)	10	66	59–72	265	28
Chan, Armstrong, Pari, Riopelle, & Munoz (2005)	18	67	58–81	276	30
Gottlob (2006)	14	69	64–73	260	30
Kemper, Crow, & Kemtes (2004)	34	75	71–79	260	50
Kline (1994)	8	65	58–71	265	27
Kosnik, Kline, Fikre, & Sekuler (1987)	12	69	65–74	264	32
Kramer, Boot, McCarley, Peterson, Colcombe, & Scialfa (2006)	10	70	62–80	281	12
Maltz & Shinar (1999)	5	66	62–80	270	29
Munoz, Broughton, Goldring & Armstrong (1998)	22	65	60–79	282	35
Olincy, Ross, Young & Freedman (1997)	24	66	52–80	271	28
Scialfa, Jenkins, Hamaluk, & Skaloud (2000)	12	68	66–70	265	15
Scialfa & Joffe (1997)	12	67	59–73	265	30
Underwood, Phelps, Wright, van Loon, & Galpin (2005)	12	68	61–76	246	33
Weighted mean parameter estimate	193	67	52–81	267	248–286*

Note. n = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

were utilized by making the required calculations. In the final case, some studies appeared to have collected relevant data (as gleaned by procedural method and variables collected), but those data were not used in study analyses and were therefore not provided in the article. In those cases, study authors were contacted for access to raw data.

Duration of saccadic eye movements. An example task used to estimate this parameter entails the participant moving the eyes from a central fixation point to a target that appears on the screen. In healthy younger adults, the average fixation duration between successive saccades is estimated to be approximately 225 ms (Salthouse & Ellis, 1980; Westheimer, 1954). For typical older adult parameter estimation, a literature search was conducted using keywords *saccadic eye movement, saccadic duration, saccade duration, visual search,* and *fixation duration.* These electronic database searches led to the identification of the following studies, from which cited references were also included and study authors were contacted for appropriate breakdown of data and means if not presented in publication (for this parameter estimate as well as all others). Table 3 presents a listing of independent groups for the latency analysis.

A modified meta-analysis was performed to estimate mean fixation duration across studies. Rather than using measurements of effect size, independent study means were weighted with respect to sample size, using calculations derived from the Hunter and Schmidt (1990) method. Hunter and Schmidt advocate their method as a random-effects model based on the belief that this technique is appropriate for the type of inferences behavioral scientists wish to make. A random effects model is more realistic than a fixed effects model when a researcher wishes to make general conclusions about the research domain as a whole, rather than restricting findings to only those studies included in the meta-analysis. As such, central tendency is measured using the average correlation coefficient in which untransformed correlations are weighted by the sample size on which they are based. Equations stemming from this method are used to calculate all parameters in the current research, as this method allows greater flexibility to generalize beyond included studies and estimate parameters for the typical older adult.

The following parameter estimate was attained through utilization of the aforementioned method. The mean combined fixation duration across all studies produced an estimate of 267 ms for older adults, a standard deviation of means of 9.7 ms (calculated as the square root of the sample variance of sample means), and lower and upper bounds of means (defined as plus or minus two standard deviations of means) of 248 ms and 286 ms. By comparison, Card, Moran, and Newell's (1983) typical estimate for fixation duration of a young adult was 230 ms, and ranging between 70 and 700 ms. Of note, Card, Moran, and Newell (1983) bounded range of performance by absolute lowest and highest values attained for individual participants within all included studies, whereas the current analysis bounded ranges by using a measure of plus or minus two standard deviations of means for studies.

Decay half-life of visual image store. When light strikes the eyes, the results of processing linger briefly in the perceptual system, such that a person may recall any part of the visual array in minute detail in the initial tenths of a second (Sperling, 1960). This kind of sensory memory is termed iconic memory and traces fade quickly to allow for new sensory input to be continually updated and integrated with earlier representations. Research shows that iconic memory of older adults remains quite stable over the life span, is more resistant to age-related decline than other types of memory, and age-related loss seems to spare memory for visual attributes (Sekuler & Sekuler, 2000). An example task used to estimate this parameter involves computing a least-squares fit to estimate the half-life of letters in excess of the memory span that participants are able to report. A literature search for keywords *visual memory, pictorial memory, iconic memory,* and *visual image store* led to the identification of the following studies to estimate the half-life of images stored in memory before decay (see Table 4).

Clearly, this search exposes a gap in the cognitive aging literature, evidenced by the limited number of studies that could be found by electronic database search. More work remains to be

Table 4
Literature Comprising Meta-Analysis for Visual Image Decay Parameter Estimation

Study	*n*	Mean age	Ages sampled	Mean visual image decay (ms)	*SE* (ms)
Lu, Neuse, Madigan & Dosher (2005)	16	81	65–99	170	38
Madden et al. (2002)	12	62	61–70	145	25
Weighted mean parameter estimate	28	72	61–70	159	135–193*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

done to extract this parameter using adequate sample sizes. Mean visual image decay half-life for older adults was found to be 159 ms, with a standard deviation of means of 12 ms and lower and upper bounds of means ranging from 135 ms to 183 ms. By comparison, Card, Moran, and Newell (1983) estimated the decay half-life of the visual store for a younger adult to be 200 ms and ranging between 90 ms and 1000 ms.

Cycle time of the perceptual processor. This parameter is defined as the amount of time that passes between the onset presentation of a stimulus and the time at which the information becomes available in working memory. According to the Variable Perceptual Processor Rate Principle (Card et al., 1983), the perceptual processor cycle time varies inversely with stimulus intensity. Card, Moran, and Newell (1983) originally estimated this parameter with unit impulse responses (the time response of the visual system to a brief pulse of light; Ganz, 1975; Harter, 1967), and reports of how many clicks were perceived when presented at rates from 10 to 30 clicks per second (Cheatham & White, 1954). A literature search for keywords *visual processing, visual memory, echoic processing, echoic memory, sensory processing, sensory memory, perceptual processing, perceptual memory, flicker,* and *tracking* produced the following studies for inclusion in this parameter estimation (see Table 5). Mean perceptual processor cycle time was calculated to be 178 ms, with a standard deviation of means of 8.6 ms and upper and lower bounds ranging from 161 ms to 195 ms. By comparison, Card, Moran, and Newell (1983) estimated the usual cycle time for a younger adult to be 100 ms, and ranging between 50 ms and 200 ms.

The Motor System

Aging is associated with multiple deficits in planning, control, and execution of movements; and loss of muscular strength, endurance, and tone are some of the most obvious age-related motor limitations (e.g., Mortimer, Pirozzolo, & Maletta, 1982; Spirduso & MacRae, 1990). As such, older adults tend to make slower movements to maintain accuracy (Welford, 1977). Furthermore, older adults take longer to initiate a movement, require more time to accelerate and reach peak velocity during the movement, are less able to efficiently decelerate and terminate the movement, and are less able to calibrate appropriate levels of force (Haaland, Harrington, & Grice, 1993). The maximum speed of movement may decline by as much as 90% from the 20s to the 60s (Spirduso, 1995).

Motor processor cycle time. Motor processor cycle time is defined as the time required to initiate a movement response (activating patterns of voluntary muscles) following a visual, auditory, or other sensory signal. This parameter thereby reflects the speed of transmission from the central nervous system to the appropriate muscle groups for action (Stelmach & Goggin, 1988).

Table 5
Literature Comprising Meta-Analysis for Perceptual Processor Cycle Time Parameter Estimation

Study	*n*	Mean age	Ages sampled	Perceptual processor cycle time (ms)	*SE* (ms)
Dennis, Scialfa, & Ho (2004)	12	65	60–70	179	27
Fisk, et al., (1995)	104	72	65–81	175	30
Glass et al. (2000)	10	65	60–70	205	35
Ho & Scialfa (2002)	10	66	60–72	182	40
Johnson, Reeder, Raye, & Mitchell (2002)	27	74	67–84	185	22
Madden, et. al., (2005)	24	67	62–82	157	29
Marshall, et al., (1983)	53	70	60–82	183	32
McFarland, Warren, & Karis (1958)	20	70	65–75	176	14
Nielsen-Bohlman & Knight (1999)	10	71	65–76	192	26
Picton, et al., (1984)	36	70	65–79	167	25
Picton, et al., (1986)	10	69	65–75	180	18
Podlesny & Dustman (1982)	16	71	65–75	180	30
Salthouse & Meinz (1995)	85	70	60–79	184	34
Sliwinski, et al., (1994)	51	77	69–85	176	20
Sliwinski (1997)	73	66	51–86	185	30
Tremblay, Piskosz, & Souza (2002)	10	68	61–79	180	22
Verhaeghen (2002)	35	71	66–77	175	28
Verhaeghen, Vandenbroucke, & Dierckx (1998)	13	64	59–69	150	17
Zheng, Myerson, & Hale (2000)	40	71	65–77	171	29
Weighted mean parameter estimate	639	69	51–86	178	161–195*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

Table 6
Literature Comprising Meta-Analysis for Motor Processor Cycle Time Parameter Estimation

Study	*n*	Mean age	Ages sampled	Motor processor cycle time (ms)	*SE* (ms)
Basak & Verhaeghen (2003)	29	74	69–79	123	28
Batsakes & Fisk (2000)	24	72	64–82	119	24
Glass et al. (2000)	10	65	60–70	135	25
Jagacinski, Greenberg, Liao, & Wang (1993)	31	65	60–69	141	24
Jagacinski, Liao, & Fayyad (1995)	16	65	60–69	142	30
Kemper, Herman, & Lian (2003)	75	73	70–80	139	19
Ketcham, Seidler, Van Gemmert & Stelmanch (2002)	15	68	62–74	129	20
Leonard (1952)	25	66	60–72	140	15
Miles (1931)	259	65	55–85	145	35
Podlesny & Dustman (1982)	16	71	65–75	160	22
Romero, Van Gemmert, Adler, Bekkering, & Stelmach (2003)	15	68	53–80	133	29
Smith, Umberger, et al. (1999)	385	75	60–90	154	37
Welford (1958)	86	65	61–69	138	31
Weighted mean parameter estimate	986	64	53–90	146	128–164*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

Studies that reported initiation of movement times were included in this analysis, by searching the literature for keywords *Fitts' law, motor processing, movement response time, motor processing speed, movement processing speed, processing speed, tapping rates, motor response time,* and *movement response time*, which led to the identification of the following studies to be included in this parameter estimation (see Table 6). The mean value for motor processing cycle time in older adults was computed to be 146 ms, with a standard deviation of means of 9 ms and lower and upper bounds of 128 ms and 164 ms. This parameter was estimated to be 70 ms for younger adults, with bounds of 30 and 100 ms (Cardet al., 1983).

Fitts' law. One common approach to assessing movement slowing is to manipulate task difficulty (information to be processed) in a parametric fashion. Fitts′ law, a well-studied law in motor control research (Fitts, 1954), states that, as the difficulty of the movement increases, the speed of the movement decreases. Research reveals that older adults tend to move slower than young adults at all levels of difficulty and are differentially slower at higher levels of difficulty (Bashore, Osman, & Heffley, 1989; Brogmus, 1991; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994; Goggin & Meeuwsen, 1992; Hines, 1979; Ketcham, Seidler, Vann Gemmert, & Stelmach, 2002; Pohl, Winstein, & Fisher, 1996; Salthouse, 1988; Walker et al., 1997).

Fitts (1954) defined an index of difficulty (ID) for a movement as the logarithm (to base 2) of the ratio of twice the target's distance to the target's width. Movement time was then fit as a linear function of ID, with the slope of that function estimated for a given task and population. The Fitts' Law slope is given in ms/bit. Welford (1968) modified Fitts' equation to improve fit between predictions and observations. The Welford (1968) modified equation has been applied to the following estimates.

A literature search for keywords *Fitts' Law, motor response, motor response time, movement response, movement response time,* and *movement time* yielded the following studies for inclusion in this analysis (see Table 7). Estimates were attained for peak velocity speeds for movements using a mouse. The weighted mean for Fitts' law slope estimation was calculated to be 175 ms/bit for older adults, with a 12 ms/bit standard deviation of means and lower and upper bounds between 163 ms/bit to 199 ms/bit. By

Table 7
Literature Comprising Meta-Analysis for Fitts' Law Slope Constant Estimation

Study	*n*	Mean age	Ages sampled	Fitts' law slope constant (ms/bit)	*SE* (ms/bit)
Chaparro, Bohan, Fernandez, Choi & Kattel (1999)	10	70	65–75	177	24
Ketcham, Seidler, Van Gemmert & Stelmach (2002)	15	68	62–74	190	22
Liao, Jagacinski, & Greenberg (1997)	40	65	60–69	163	38
McCrea & Eng (2005)	20	61	55–67	198	33
Rogers, Fisk, McLaughlin, & Pak (2005)	40	59	51–65	159	32
Romero, Van Gemmert, Adler, Bekkering, & Stelmach (2003)	15	68	53–80	179	34
Walker, Philbin, & Fisk (1997)	16	70	65–75	178	32
Welford (1958)	30	70	65–75	179	24
Worden, Walker, Bharat & Hudson (1997)	16	70	65–75	181	30
Weighted mean parameter estimate	202	67	51–80	175	163–199*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

comparison, Card, Moran, and Newell (1983) estimated the rate of movement for younger adults to be 100 ms/bit, and ranging from 70 – 120 ms/bit.

Power law of practice. It is a well-known finding that as people practice a task, performance speeds up and they become more proficient as training progresses. The Time T_n to perform a task on the *n*th trial approximates a power law. To estimate the rate of improvement in basic information processing tasks for older adults, a literature review searching for keywords *practice, practice effect, block, block effect, skill acquisition, improvement rate,* and *power law of practice* was conducted. Included studies are listed in Table 8.

As older adults generally begin with poorer performance than younger adults, they have more to gain from practice over time (Charness, Holley, Feddon, & Jastrzembski, 2004). Thus, calculations for the power law slope constant resulted in a value of 0.49 for older adults, with a standard deviation of means of 0.05 and lower and upper bounds of 0.39 and 0.59. Younger adults, by comparison, were originally estimated to learn at a power law slope constant rate of 0.40 and ranging from 0.20 to 0.60 (Card et al., 1983).

The Cognitive System

A large body of research, including both cross-sectional studies and longitudinal studies, has investigated changes in cognitive function with aging, and findings reveal that working memory, information processing, selective attention, and problem-solving ability show differential decline with aging (e.g., Salthouse, 1996). Multiple studies have revealed that a significant proportion of age-related decline in cognitive performance can be explained by deterioration of working memory (Raz, 2000; Salthouse, 1994).

Storage capacity. Many tasks have been developed that attempt to measure Working Memory (WM) capacity and are defined by involving both a processing and a storage component. Good examples include the alphabet span task (Craik, 1986), requiring participants to repeat back a series of words after arranging them in alphabetical order, and the backward digit span task (Botwinick & Storandt, 1974), requiring participants to repeat a series of digits in reverse order. Card, Moran, and Newell (1983) estimated effective WM capacity to be in the 7 ± 2 range based on performance of visual presentation of stimuli for letter span, digit span, backward digit span, and location span. Therefore, a literature search using keywords *working memory capacity, effective capacity, working memory, letter span, digit span, location span,* and *backward digit span* was conducted to attain the following studies for inclusion in the parameter estimation (see Table 9). Effective WM storage capacity for older adults was computed to be 5.4 items, with a standard deviation of means of 0.25 items and lower and upper bounds of 4.9 and 5.9 items.

In addition to the estimate of effective capacity of working memory, Card, Moran, and Newell (1983) also estimated the pure capacity of working memory. The distinction between the two rests on whether short-term memory is able to be augmented by long-term memory. In the first case, effective capacity of working memory is more conducive to rehearsal or chunking strategies, thereby enhancing performance. In the latter case, pure capacity refers to items in short-term memory that do not need to be rehearsed or deeply encoded to be maintained.

Waugh and Norman (1965) devised an equation to quantify pure capacity based on the probability of successful recall at any given point in a serial recall task and Card, Moran, and Newell (1983) utilized this method to estimate pure working memory capacity for the Model Human Processor. A literature search for this parameter estimate was conducted using keywords *pure capacity, primary capacity, primary memory, short-term memory, short term memory, episodic memory, memory capacity, serial position,* and *serial recall,* and for calculations for pure WM capacity for older adults were estimated using the Waugh and Norman (1965) method. Studies identified are shown in Table 10.

Pure (or primary) WM storage capacity for older adults was computed to be 2.3 items, with a standard deviation of means of 0.11 items and lower and upper bounds of 1.9 and 2.5. This estimate is very close to Card, Moran, and Newell's (1983) original Waugh and Norman (1965) method estimate of a 2.5 item

Table 8
Literature Comprising Meta-Analysis for the Power Law of Practice Exponent Parameter Estimation

Study	*n*	Mean age	Ages sampled	Power law constant	*SE*
Bherer, Kramer, Peterson, Colcombe, Erickson, & Becic (2005)	36	70	62–77	0.48	0.07
Charness & Campbell (squaring task, 1988)	16	67	59–75	0.43	0.11
Fisk, Cooper, Hertzog, Anderson-Garlach & Lee (1995)	104	72	65–81	0.55	0.09
Glass et al. (2000)	10	65	60–70	0.42	0.07
Hertzog, Cooper, & Fisk (1996)	104	72	65–81	0.48	0.1
Hoyer, Cerella, & Onyper (2003)	28	72	60–75	0.45	0.14
Jamieson & Rogers (2000)	40	69	60–80	0.45	0.12
Jenkins, Myerson, Joerding, & Hale (2000)	16	71	62–77	0.56	0.14
Ratcliff, Thapar, & McKoon (2004)	41	70	60–75	0.51	0.11
Rodrigue, Kennedy, & Raz (2005)	7	70	62–82	0.46	0.11
Strayer & Kramer (1994)	20	68	65–75	0.42	0.13
Touron & Hertzog (2004)	40	67	60–75	0.43	0.09
Touron, Hoyer, & Cerella (2001)	24	71	64–75	0.46	0.08
Welford (1958)	12	70	65–75	0.45	0.09
Weighted mean parameter estimate	498	70	59–82	0.49	0.39–0.59*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

Table 9
Literature Comprising Meta-Analysis for Effective Working Memory Storage Capacity Parameter Estimation

Study	*n*	Mean age	Ages sampled	Effective working memory storage capacity (items)	*SE* (items)
Bopp & Verhaeghen (2005)	855	70	61–78	5.3	0.9
Gregoire & Van der Linden (1997)	426	68	60–79	5.3	0.8
Haarmann, Ashling, Davelaar, & Usher (2005)	36	75	66–85	4.2	1.1
Hale, Myerson, Faust, & Fristoe (1995)	23	70	65–75	5.6	0.6
Hofer, Berg, & Era (2003)	1006	72	68–76	5.6	1.0
Jenkins, Myerson, Joerding, & Hale (2000)	16	71	62–77	4.9	0.2
Kemper, Crow, & Kemtes (2004)	34	75	71–79	4.7	0.9
Light & Anderson (1985)	18	70	64–76	4.9	0.5
Myerson, Hale, Rhee, & Jenkins (1999)	20	70	65–75	5.5	0.7
Rochon, Waters, & Caplan (2000)	15	71	65–78	5.5	1.2
Stine & Wingfield (1987)	24	69	59–81	4.2	0.6
Waters & Caplan (2003)	83	74	60–80	5.5	1.3
Weighted mean parameter estimate	2556	71	59–85	5.4	4.9–5.9*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

capacity in primary WM for younger adults, and ranging between 2 and 4.1 items. This is not surprising as the literature encompassing this parameter estimate reported no significant age group differences for primary memory capacity. Rather, age differences become apparent when secondary memory processing (e.g., rehearsal, encoding into chunks) is required.

Cognitive processor cycle time. The recognize-act cycle is the basis for cognitive processing, such that contents of WM associatively link with actions in Long-Term Memory (LTM; “recognize”), which in turn modify the contents of WM (“act”), so that the cycle may begin anew (Cardet al., 1983). Procedures, plans, decision-making, reasoning, and problem-solving are all built up out of organized sets of recognize-act cycles, and thus, many types of tasks may be included to estimate the cycle time required to complete each recognize-act sequence. These include memory scanning rates, enumeration tasks, perceptual judgments, choice reaction times, lexical decision tasks, and counting rates. It should be noted that according to the Variable Cognitive Processor Rate Principle, cognitive processor cycle times decrease as effort induced by task demands increase, and cycle times also diminish with increased amounts of practice, expertise, and automaticity. A literature search for keywords *choice reaction, choice reaction time, lexical decision, enumeration, memory scanning, perceptual judgment, cognitive processor cycle, cognitive processor, cognitive processing,* and *cognitive processing rate* yielded the following studies for inclusion in this analysis (see Table 11). Mean Cognitive Processor cycle time calculated across studies was revealed to be 118 ms for older adults, with a standard deviation of means of 7.9 ms and lower and upper bounds ranging from 102 ms to 134 ms. This estimate is 1.67 times that of the young adult estimate of 70 ms produced by Card, Moran, and Newell (1983), having lower and upper bounds of 25 and 170 ms.

Summary of Parameter Estimations for Younger and Older Adults

The following table is composed of original Card, Moran, and Newell (1983) Model Human Processor components based on performance of younger adults, corresponding older adult parameter estimations attained from the current research, and ratios of parameters across age (see Table 12). It is worth noting that estimates for cognitive, motor, and perceptual processor cycle

Table 10
Literature Comprising Meta-Analysis for the Pure Working Memory Storage Capacity Parameter Estimation

Study	*n*	Mean age	Ages sampled	Pure working memory storage capacity (items)	*SE* (items)
Allen & Coyne (1989)	18	69	60–75	2.5	0.7
Allen & Crozier (1992)	20	70	60–78	2.6	0.6
Bäckman & Small (1998)	130	82	75–90	2.1	1.1
Castel, Benjamin, Craik, & Watkins (2002)	18	72	65–75	2.2	0.6
Floden, Stuss, & Craik (2000)	24	71	65–77	2.3	0.7
Kahana, Howard, Zaromb, & Wingfield (2002)	28	75	69–84	2.2	0.8
Maylor, Vousden, & Brown (1999)	99	65	56–73	2.2	0.4
Murphy, Craik, Li, & Schneider (2000)	15	71	65–79	2.5	0.7
Nilsson (2003)	331	72	65–80	2.3	0.7
Weighted mean parameter estimate	683	72	56–90	2.3	1.86–2.52*

Note. *n* = number of participants in age range of interest per study.
* Values represent plus or minus two standard deviations of means.

Table 11
Literature Comprising Meta-Analysis for the Cognitive Processor Cycle Time Parameter Estimation

Study	*n*	Mean age	Ages sampled	Cognitive processor cycle time (ms)	*SE* (ms)
Allen, et al., (2004)	97	71	60–87	102	22
Anderson (1999)	24	70	59–81	131	18
Hertzog, Cooper, & Fisk (1996)	104	72	65–81	132	27
Ratcliff, Thapar, & McKoon (2001)	40	68	60–75	111	17
Ratcliff, Thapar, Gomez, & McKoon (2004)	44	68	60–75	108	20
Ratcliff, Thapar, & McKoon (2004)	41	70	60–75	104	17
Thapar, Ratcliff, & McKoon (2003)	38	69	60–75	107	26
Verhaeghen, Marcoen, & Goossens (met a-analysis, 1992)	1539	69	Means over 60	125	24
Verhaeghen, Steitz, Sliwinski, & Cerella (met a-analysis, 2003)	1165	Over 60	Means over 60	108	30
Welford (1958)	12	70	65–75	125	25
Weighted mean parameter estimate	3104	70	59–87	118	102–134[*]

Note. *n* = number of participants in age range of interest per study.
[*] Values represent plus or minus two standard deviations of means.

times are nearly twice as long in older than younger adults, and some fairly straightforward predictions can be made by decomposing tasks into their processing components.

In GOMS (Goals, Operators, Methods, Selection Rules) modeling, cycle times are iterated quite often when a task is modeled; hence, they tend to dominate in predicting response times. Previous research in the cognitive aging field has not formally estimated processing cycle times for older adults using task analyses and, therefore, has been able to provide only a simple rule of thumb to designers, suggesting that when using the same strategy for performing a task, older adults will typically take 1.5 to 2 times as long as a younger adult (Fisk et al., 2004). However, task components will skew the estimates for response time, so if tasks rely heavily on motor responses, old-young differences are predicted to be on the order of 2.1:1, but when eye movements play the weightier role for instance, old-young differences are predicted to be considerably smaller given the 1.2: 1 ratio observed for saccade duration. Furthermore, it is fairly clear that when novel tasks are kept simple, by minimizing the number of steps that must be maintained in working memory, young-old differences in error rates should be minimized, given the relatively small differences in capacity of working memory and in memory decay rate, although older adults would be expected to perform these tasks more slowly. We now apply a subset of the older adult information processing parameter estimates to a model validation study using a mobile phone task.

Application and Validation of Age-Related Parameters to Mobile Phone Tasks

The GOMS modeling technique. We have chosen the venue of mobile phone technology to validate model parameters for older adults by comparing goodness-of-fit to previously validated models of younger adults using the GOMS modeling technique in both a simple dialing task, and a more complex text messaging task. GOMS represents a widely used simplified cognitive architecture that has provided accurate predictions of routine performance in a wide range of procedural tasks (e.g., Card et al., 1983; see Olson & Olson, 1990, for a review).

GOMS methodology estimates how a simulated user would perform in a simulated system using a specified set of procedures. Procedures are decomposed to the keystroke-level of analysis to make performance predictions for any specified strategy and to help determine what the most effective (in terms of task completion time or accuracy) strategy would be (Card et al., 1980, 1983), while accounting for constraints of perceptual, motor, and cognitive workload and processing time. The GOMS-level approach has the ability to demonstrate how greatly milliseconds matter, and how even very small increments of time at the fine-grained, low-level microstrategy may balloon into significant differences by the end of a task. This cascade of operators, in a poorly designed system, may compound to make a system feel sluggish

Table 12
Summary of Parameter Estimates for Younger and Older Adults with Ratio of Differences Across Age

Parameter of interest	Younger adult estimate (Card, Moran, & Newell, 1983)	Older adult estimate (Jastrzembski, 2006)	Ratio of old to young
Duration of saccadic eye movements	230 ms (70–700)	267 ms (248–286)	1.16:1
Decay half-life of visual image store	200 ms (90–1000)	159 ms (135–183)	0.8:1
Cycle time of the perceptual processor	100 ms (50–200)	178 ms (161–195)	1.78:1
Cycle time of the motor processor	70 ms (30–100)	146 ms (128–164)	2.1:1
Power law of practice constant	0.4 (0.2–0.6)	0.49 (0.39–0.59)	1.2:1
Fitts' law slope constant	100 ms/bit (70–120)	175 ms/bit (163–199)	1.75:1
Effective capacity of working memory	7 items (5–9)	5.4 items (4.9–5.9)	0.77:1
Pure capacity of working memory	2.5 items (2.0–4.1)	2.3 items (1.86–2.52)	0.92:1
Cycle time of cognitive processor	70 ms (25–170)	118 ms (102–134)	1.69:1

and nonresponsive and may result in higher workload or increased likelihood of error by the user (Gray & Boehm-Davis, 2000).

Mobile phone technology and usability. Over two billion people worldwide currently use cellular telephones, and that figure continues to climb as it spreads to new demographics and countries (International Telecommunications Union, 2006). According to research conducted in the United Kingdom, of mobile phone users, 70% of people aged 55–64 years old and 53% people aged 65–74 own a mobile phone (although the ownership drops to 24% for people 75 years old and above; Office for National Statistics, 2003). Yet despite the widespread use of mobile technology, 30% of users claim that the mobile telephone is their most hated invention, and also the one they are most unable to live without (Lemelson-MIT Program, 2004). Indeed, older adults tend to avoid using features above and beyond the simple function of placing and receiving phone calls because they argue that displays that are too small and too difficult to read, buttons and characters are too small to see or press (causing wrong numbers to be pushed), too many functions are packed into a single key, and finding help for how to use a function is oftentimes difficult to achieve (Chattratichart & Brodie, 2003; NTT DoCoMo, 2000).

Text entry and mobile telephones. An example of a mobile phone design feature that addresses critical inadequacies of the mobile phone is the case of text entry, which is required for both simple and complex mobile phone tasks. In the simple case, text entry is necessary for creating shortcuts in a phonebook by entering name and phone number information for contacts that a user wishes to save in his or her phone directory. In more complex cases, text entry can be extremely intensive for tasks like creating appointment reminders, text messaging other contacts, or even sending emails or browsing the web.

Mobile phones are not naturally suited to text input, and unique challenges exist for entering text on mobile phones compared to through standard keyboards. This is due to size and space constraints in design that require alphabetic characters (e.g., in English, the letters A-Z), 10 numbers (0–9), and numerous punctuation marks to be confined to a minimal set of keys. The 12-key keypad is the most typical design found in contemporary mobile phones (see Figure 1, Panel A), although drastic differences are apparent in key size, key shape, and even the way each key must be pressed across telephone designs. An additional challenge is that smaller keys (e.g., mobile phone keypads) are harder to press successfully than larger keys (e.g., a computer keyboard digit keypad), leading to higher error rates (Fitts, 1954), and this problem may be compounded if small keys are placed too closely together, increasing the probability that larger or less coordinated fingers may slide off an intended button accidentally.

The multipress method is the most common type of text input for mobile phones, requiring users to specify the input of a character by the number of presses on a single key. This approach carries the problem of segmentation, such that when a character is placed on the same key as the previously entered character (termed a key-based digraph), the system must determine if the new key press is part of the previous character, or if it represents the beginning of a new character. These designs are generally criticized as being slow and cumbersome, and workload levels vary not so much by the number of letters in a word, but rather the placement of those letters with regard to individual key mapping sequences (James & Reischel, 2001).

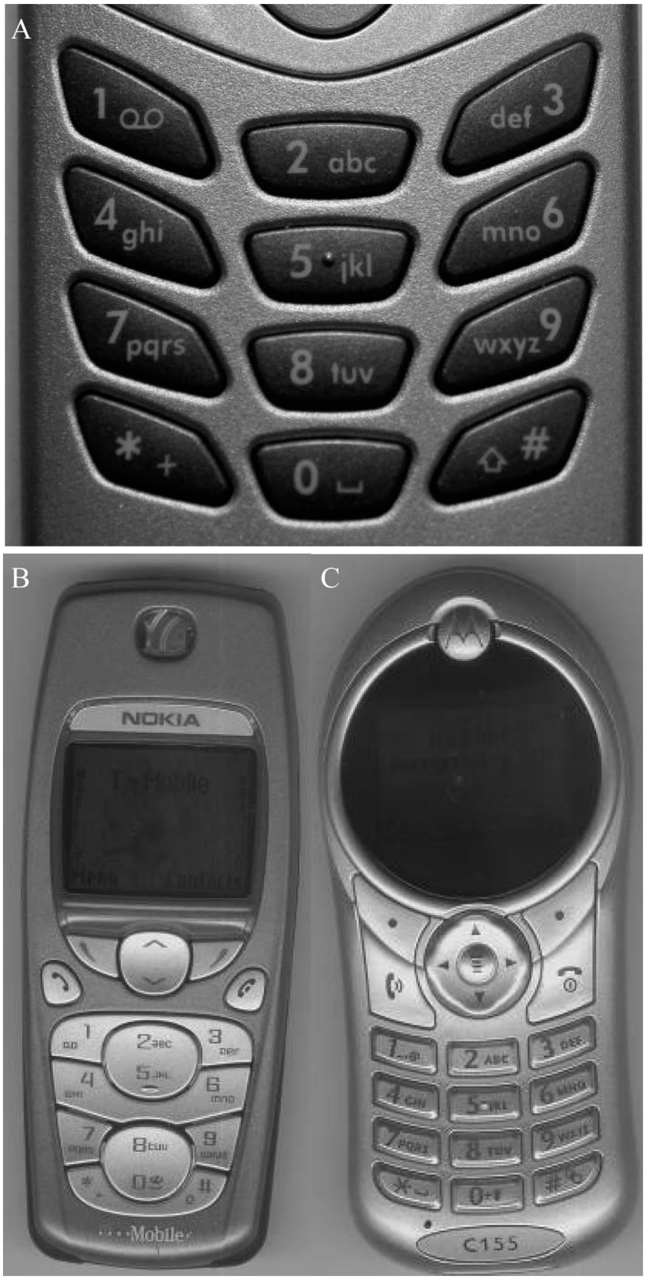

Figure 1. Panel A. The 12-key keypad. Panel B. Nokia 3595. Panel C. Motorola C155.

Furthermore, designers often specify a timeout period, usually between 0.5 and 1.5 seconds, so that the system automatically assigns a key press to a new character after the timeout ends. These timeout periods may be quite problematic for users trying to enter information quickly enough (e.g., older users or novices may scroll too slowly so that the device accepts a character before the intended character is reached) or may create the problem of feeling sluggish for faster users waiting for the timeout period to end (e.g., younger users or experts may wait too short a time between subsequent characters on the same key). As a result, this multipress method poses specific, predictable text-entry challenges for both younger and older users.

Mobile telephones and menu navigation. Mobile phones typically represent electronic information with hierarchical menu structures, but standards for arrangement of functions and keys to ensure best usability have not been defined (Bay & Ziefle, 2005) and are not consistent across mobile phone brands. Further, menus are displayed on miniaturized screens that allow only a limited number of items to be seen at one time. Thus, the user may navigate through a menu whose complexity, extension, and spatial structure is not readily apparent, and is hidden from sight. This isolation of information across screens has consequences for memory load and cognitive processing, and if the user is unable to successfully internalize the menu hierarchy, the user may lose orientation, feel lost, or be unable to complete targeted actions (Ziefle & Bay, 2005). To support this claim, Ziefle and Bay (2005) found that older adults strayed from the critical path of specified tasks more frequently than younger adults and exited the system more often due to loss of orientation within the menu hierarchy.

Methodological basis for current validation study. St. Amant, Horton, and Ritter (2004) conducted a model-based evaluation of a representative mobile phone to assess usability and test model-based predictions of behavior. Cell phone tones produced by each key press were transmitted through the earphone jack of the cell phone, connected to the microphone jack of a computer, and recorded using the GoldWave shareware program. The onset of each key press was detectable by the unique frequency from the touch-tone keypad so that evaluators could decipher which button was pressed based upon the frequency of the waveform. Thus, this method of recording allowed for precise measurement of button presses to be attained in an ecologically valid way, as users were able to interact with the mobile phone device, rather than a computer simulation or mock-up.

Hypotheses. For the current study, we seek to validate Model Human Processor parameters for the older adult using GOMS analyses for specific tasks on two different mobile phone devices. As such, both younger and older adults will attain proficiency on ecologically realistic tasks using actual mobile phones (e.g., dialing and text messaging), and GOMS simulations implementing Model Human Processor parameters for younger and older adults were generated a priori and compared against the performance of human users. It is hypothesized that goodness-of-fit for the older adult model compared to the mean older adult human will be comparable to that of the younger adult model compared to the younger adult human, for both the dialing and text messaging tasks, and specific predictions are detailed below.

In terms of mobile phone usability, we hypothesize that older adults will perform all mobile phone tasks more slowly than younger adults as a result of general slowing. More specifically, we expect that older adults will take approximately 1.4–1.5 times longer than younger adults (averaged across all tasks and devices) to complete tasks with each device, as based on younger and older GOMS models. This prediction results from expected longer processing times for older adults on basic cognitive, perceptual, and motor building blocks shown in the Tables above.

In terms of device, we hypothesize that task completion times will be shorter using the smaller, more complicated, Motorola C155 mobile phone (Phone B) than the larger, less complex, Nokia 3595 model (Phone A, described in the Materials and Equipment section below), for the simple dial-a-number task. This is primarily expected because of Fitts' law analyses that predict faster performance on the basis of button size and button distance and because critical pathways are otherwise identical between phones for this simple task. Response time differences are also predicted to be minimized between younger and older adults using Phone B as compared to Phone A, on the basis of Fitts' law predictions. However, greater speed associated with the smaller Phone B comes at the potential cost of more error-prone performance. Because Phone B buttons are smaller, the user may be more likely to encounter situations when their fingers slip, miss, or hit multiple keys, compared to the larger, grooved buttons on Phone A, and thus, we hypothesize that in the simple dialing task, the smaller Phone B model will incur significantly higher error rates for both younger and older adults.

In the more complex text messaging task, the current study predicts that both human and model performance will be most time-efficient and least error-prone in menu navigation/text entry tasks using the mobile phone of least complex design (Ziefle, 2002; Ziefle & Bay, 2005). As will be detailed in the procedure section, we predict that the mobile phone with the larger buttons, larger display, and more environmentally supportive (Craik, 1986) interface will help users complete complex tasks with greater success. As such, we also hypothesize that the less demanding device specifications of the Phone A device (as detailed in the hardware and software description in the subsequent section) and the fewer perceptual and cognitive operations required to complete the task (as explicated by and predicted by GOMS critical pathway analyses) will reveal faster and more accurate performance for users both young and old during text messaging. We hypothesize that Phone B will lead to more errors, due to greater taxing of working memory, the requirement to integrate more information across screens, and the need for more fixations to extract necessary information among distracters (as revealed by GOMS analyses). Similarly, we predict that response time differences between younger and older adults will be greater for the Phone B device than the Phone A device, due to the iterative nature of having to engage perceptual and cognitive processors more frequently.

Method

Participants

The sample for these analyses consisted of 20 younger adults (university students obtained from the subject pool, $M_{\text{age}} = 18.7$, $SD = 1.3$) and 20 older adults, $M_{\text{age}} = 69.1$, $SD = 3.5$. recruited from a newspaper advertisement and paid a small stipend. (Because this study was embedded within a concurrent stress and technology investigation, complete data for a total of 28 younger and 28 older adults were eventually collected, due to a need for greater power to examine physiological differences across dependent variables. Because the current study was an independent line of research from the stress and technology study and required only 20 participants per age group a priori, only data collected from the first 20 participants in each age group were analyzed). All participants had normal or corrected-to-normal vision (determined by a prescreening measure through self-report), normal or corrected-to-normal hearing (assessed using an audiometer at the time of experimental testing), and were free of any disorder than could make it difficult to complete fine motor movements (e.g., severe arthritis, joint diseases; determined by a prescreening measure through self-report).

Materials and Equipment

Mobile phone designs. Experimental data were recorded using the method of St. Amant, Horton, and Ritter (2004). Technicians from the Florida State University Psychology department designed unique wires to transmit button presses from each mobile phone to the computer in order to retrieve and record precise data. Tones from each mobile phone were transmitted through the earphone jack of the device and connected to a microphone jack on the computer and were recorded with the GoldWave shareware program using millisecond timing; keys pressed and error information were identified by means of tone frequency. Response time and accuracy information were later analyzed for each button press and compared to model performance in a step-by-step fashion.

Mobile phone models chosen for analysis included the 12-key keypad, multipress Nokia 3595 mobile phone (Phone A: see Figure 1, Panel B), and the Motorola C155 mobile phone (Phone B: see Figure 1, Panel C). Phone A has a 2-directional navigational key with large selection keys directly below menu choices on the display, a 4096 color display with resolution of 96x65 pixels, a display size of up to four text lines using 12-point font and two information bars, a grooved keypad design making buttons somewhat larger (buttons range from 1–2 cm in width and height) and providing a homing tactile raise on the “5” key to help orient users, and possesses dimensions of 4.64” x 1.95” x 0.87”. This phone is hypothesized to be the more ergonomically efficient mobile phone of the two that this study compares.

Phone B has a 4-directional navigational key with the Menu key (shown as a circle with three horizontal bars, approximately 0.4 cm in diameter) centered within it (this is not the OK button and users must remember to press this when they wish to see menu options). It also possesses a 4096 color display with resolution of 96x65 pixels, a display size of up to two text lines using 10-point font and two information bars, and a fairly standard keypad design (using buttons of equal sizes of approximately 0.9 cm wide and 0.5 cm high) with a raised notch on the “5”-key to aid homing. Its dimensions are 4.09” x 1.93” x 0.94”.

As mentioned above, Phone A is hypothesized to aid performance in terms of both task completion time and accuracy. The 2-directional navigational key is twice as large as Phone B’s directional keys, meaning that intentional presses do not require as much precision and thereby are less likely to produce an error. The display screen is approximately 25% larger than Phone B, allowing more lines of text to be displayed on a single screen at once and further reducing the need for users to integrate information across several screens. This helps reduce working memory load so that users may complete tasks with less cognitive workload and recover from errors with greater ease given the extra environmental support.

Text is also presented in a larger font size (Phone A: menu options are 14 pt and text entry is 12 pt; Phone B: menu options are 12 pt and text entry is 10 pt), making information easier to extract on a perceptual level. Buttons on the actual keypad are also larger and designed to cradle the finger, allowing the intended key to be pressed with less room for error. Lastly, all menu options are displayed on the screen at all points in time through use of right and left selection keys. This screen display technique provides environmental support to help the user make decisions at any juncture (e.g., Morrow, 2003), without relying on memory to complete unwritten sets of sequences. This technique differs from Phone B, which requires users to press a specific “Menu” button, labeled only by three horizontal bars, to access functions that are not currently listed as options on the screen. In summary, these models were sufficiently different in terms of hardware (e.g., button size, type of button press requirements) and software (e.g., different critical paths in the menu hierarchy to complete each task), to allow for human factors analyses to help determine which design specifications are most usable for a set of representative phone tasks for both younger and older adults and to highlight the benefits of GOMS modeling for predicting what type of device will be most usable.

Instructions. Specific graphic and oral instructions regarding functionality for each device were given to participants during learning, so that users were familiarized with button functions and navigation of the menu hierarchy. Separate instructions were provided before each task and for each device, so that critical pathways from one task would not interfere with the current task of interest. At the beginning of the session, general introductions to the specific phone that would be used that day were provided. The experimenter gave the participant an enlarged picture of the device on a sheet of paper, with labels denoting what and where keys were located, while scripted information was read regarding key functionality. The experimenter pointed to each item on the paper as they were read in the script. Specifically pointed out were the location of the selection keys for choosing options displayed on screen, the “Call/Send” button, the space key for text entry, the arrow navigational keys to scroll through a menu hierarchy, the “Clear” button for deleting mistakes and moving backward one step at a time, and ending calls or escaping the system altogether when lost by using the “End” button. Participants were able to ask any questions about the device at this time.

The experimenter then went on to read scripted instructions for the practice session of the first task, a simple dial-a-number task. The experimenter demonstrated this task by dialing a 10-digit number and pressing the “Call/Send” button to place the call and end the trial. Participants were then told they were to dial the same 10-digit number until performance had stabilized (see the following procedure section for additional details), signifying that practice had come to an end. Because GOMS-level analyses and model validations require tasks to be routine and for users to be proficient, it was necessary to insure that individual performance had stabilized by the end of each practice session (measured by an asymptote in performance defined as three consecutive trials performed within one second of each other and determined to be adequate by initial pilot study analyses), and, as such, the length of practice sessions and the number of trials required to reach stabilization was expected to vary across participants. Upon completion of practice, the experimenter informed participants that practice had ended, that no help could be offered for the next trials, and that the experimental blocks for the dialing task would begin.

After the dialing task, the experimenter provided instructions for the next task, a text-messaging task. Scripted instructions were specifically tailored for each device, so that the vastly different pathways required to fulfill each trial were explicated sufficiently and in adequate detail (e.g., different menu hierarchies, different key functions, and different number of steps). The experimenter read scripted details regarding how letters were entered on the 12-key keypad, that a timeout period determined character spacing, and that in order to type two letters located on the same key one must wait for the timeout period to end. The experimenter also demonstrated these details by pressing the number "2" key three times to enter the letter "c," and then showing participants they must wait for the cursor to appear (indicating the timeout period had ended) to then press the number "2" once more to enter the letter "a." The space key was also specifically pointed out on each device, as it was located on the "0" key on the Phone A device and on the "*" key of the Phone B device. The participant was then handed the phone and asked to type "c" and then "a," to insure that the premise of text-messaging was understood, and any questions were answered at that time.

After participants successfully entered the practice text and indicated that they understood text messaging, experimenters read detailed, scripted task-specific instructions tailored to the device's critical pathway. The experimenter first demonstrated a specific text-messaging trial to the participant, reading the list of steps to perform (e.g., press the menu key and then press the down navigational arrow two times to get to the "Text Message" option), showing the participant how to complete each step of the procedure, and then allowing the participant to perform the task with the device in hand while being told how to proceed in a step-by-step fashion. Participants were given the opportunity to ask for help or explicit guidance if assistance was required to learn the entire pathway, and adequate practice was provided so that participants could memorize the precise critical path they were to learn and execute. As with the dialing task, participants practiced the same text-messaging trial over and over, until performance had become stable, defined as three consecutive trials completed within one second of each other, without any assistance from the experimenter. Again, participants were allowed to ask any questions they may have during this time, and upon successful completion of practice participants were notified that practice had ended and that experimental blocks would begin.

Procedure

Participants completed a total of three sessions over the course of this study and were administered informed-consent forms at the beginning of each day to explain the procedure for each session (and were debriefed at the end of the study).

Session one – cognitive battery. The first session consisted of a group-administered cognitive battery to assess individual differences and was collected as part of a larger study on cognitive performance and aging (see Czaja et al., 2006, for details on the content of the battery). This session lasted approximately three hours for older adults and approximately two hours for younger adults. Background, demographic, and technological experience information were collected via questionnaire at the beginning of the first session.

Sessions two and three – mobile phone experiment. Participants then made appointments to return on two consecutive days to complete the mobile phone portion of the study, which also included measures of physiological stress as part of a concurrent study (Dijkstra, Charness, & Yordon, 2007). Participants completed a questionnaire detailing mobile phone experience at the end of these final two sessions. Sessions lasted approximately two to two-and-a-half hours for older adults, and one to one-and-a-half hours for younger adults. Session time differences were a function of both practice length and actual experimental completion times to successfully complete each task.

Participants worked with only one mobile phone model per session, to reduce potential transfer errors that would have been made if both mobile phones were presented at the same time. To minimize effects of order, mobile phone presentation order was counterbalanced across participants. Task presentation was administered identically for each phone over the two experimental sessions, meaning that order was consistent upon return for the second day of testing with the other mobile phone.

The experimenter familiarized participants with the characteristics of each phone using scripted text and a diagram extracted from the manual accompanying each phone. Participants were allowed to ask questions and practice using the phone during this acclimation period. They were next administered a practice session to familiarize themselves with the exact dynamics of the task as described in the previous section.

Experimental tasks. The tasks of interest varied in complexity and included:

1) the simple dialing of a number (e.g., on Phone A, the user would Type 8506442040 "Send") and

2) typing a text message and sending it to a person in the contact list (on Phone A, this requires navigating the menu to the Message menu, moving to Text Messages, selecting Create, typing the message, selecting send, then options, then find, scrolling to the proper name in the contact list, selecting the proper number from the user name, and finally selecting OK to send the message).

Text entry portions of the text messaging task were carefully chosen, to ensure that equal number of digraphs appeared in each trial (e.g., the trials "Call Will" and "Meet Bill" each possess three digraphs requiring users to wait for the system to timeout before moving on to the next character) and no common shortcuts or acronyms could be typed in place of any word (e.g., users often shorten the word "You" to "U," "See" to "C," or "later" to "l8r").

Upon completion of practice, participants were administered five unique trials of the learned task per block. These trials consisted of different messages to type and different people to send them to from the phone directory. Four blocks were administered consecutively, presenting the same five trials in the same order for each block, and a short break was given between blocks. Trials were presented on individual sheets of paper and included only the data necessary for entry (information was centered and presented in 24 pt font). No instruction or help concerning the pathway was presented. The experimenter presented one trial at a time, one sheet at a time, so that the participant was not required to put down and pick up the mobile phone device between trials. Participants were given a break between trials and blocks if necessary, and each block was recorded and labeled by participant identification number. A total of 20 trials were tested for each of the three tasks (five unique trials tested over the course of four blocks), and each task followed the same practice and performance procedure as described above.

Analysis

A priori response time predictions were made using GOMS models defined separately for younger and older adults. An example is shown in the Appendix for the dial-a-number task. We chose to analyze the first trial, averaged across all four experimental blocks. Although later trials may provide important information, we limited our analysis to first trial data for ease of analysis, and we also limited our consideration to GOMS models with error-free performance. Thus, errors in human data were removed for initial goodness-of-fit analyses. The amount of time passing from the first incorrect button press to the resumption of correct behavior was subtracted from total task completion time so that commission of errors did not interfere with normal task analysis of correct button presses. Studies have shown that people tend to slow down after the commission of an error and tend to speed back up until another error occurs (e.g., Rabbitt & Rodgers, 1977; Rabbitt & Vyas, 1970), suggesting that this trimming procedure could lead to performance times that are slower after error commission than before the error occurred. Of further note, studies reveal that large amounts of total task time are spent performing error recovery operations after an error is committed (Card et al., 1983), so it may be the case that slowing is especially evident during these recovery procedures, and that people speed back up once they realize they are back on track.

Because the study design repeated trials over four blocks, a total of four identical trials were used to compute the mean performance time of each button press, across each participant. As the Model Human Processor has been validated many times over for younger adults (e.g., mouse-driven text editing, Card & Moran, 1988; interactive computer systems, Card, Moran, & Newell, 1980; telephone operator workstation, Gray, John, & Atwood, 1993; expert performance on dynamic computer tasks requiring auditory and visual processing, John, 1990), construction of the critical path for a younger adult was used to help validate correct implementation of the model for extension to older adults. Pilot studies tested whether younger and older adult pathways were equivalent, and it was determined that although older adults tended to verify actions much more often than younger adults during the beginning stages of learning (as observed and counted by the experimenter), they eventually reduced the number of fixations to match those of younger adults. Specifically, older adults made an average of 5 fixations in the first practice trial of the dialing task, but reduced that number to an average of 2 fixations by the time practice had ended. Younger adults remained constant at 2 fixations over the course of all trials. Thus, equivalent GOMS critical pathways were constructed for the young and old models. Response time and accuracy measures were extracted through the GoldWave program on a step-by-step basis.

Model and human data were compared at each observable data point. Root mean squared deviations (*RMSD*) and correlation coefficients (*R*) were then calculated across the entire task routine to determine overall goodness of fit. This method of testing model fits computes the mean of the squared deviation between each model prediction and the corresponding observed data point and then summarizes the overall fit with a correlation measure.

Results

Analyses for the current study can be broadly categorized into human factors analyses for performance on mobile phones as a function of age and goodness-of-fit between GOMS models and human performance data. Each of these components of the data analysis will be considered in turn. Unless otherwise noted, the alpha level for all statistical tests reported was set to $p < .05$.

Human Factors Analyses

Usability for each task was analyzed in terms of both human performance completion times and error rates. These data were analyzed in a 2 (Age: younger, older) x 2 (Device: Phones A & B) mixed measures analysis of variance (ANOVA). Results will be presented separately for practice and experimental sessions described below.

Practice Sessions. An interaction of Age (younger, older) × Task (dial-a-number, text messaging) × Device (Phones A & B) was observed for length of practice session, $F(1, 38) = 26.18$, $MSe = 1560.31$, $f = 0.75$, such that contrasts of performance time differences across age were greatest in the text messaging condition and considerably longer when using Phone B, $F(1, 38) = 69.16$, $MSe = 3182.39$, $f = 1.78$ (see Figure 2).

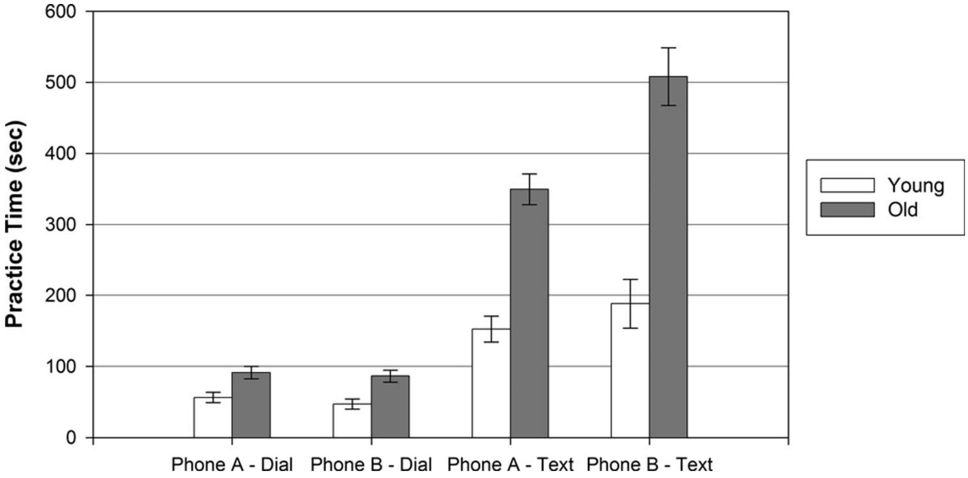

Figure 2. Age by task by device interaction for length of practice session with 95% confidence interval bars.

Dial-a-number experimental task. As predicted, results revealed an interaction of Age x Device, such that task completion time differences were minimized between older and younger adults when using Phone B compared to Phone A, $F(1, 38) = 7.18$, $MSe = 69097.8$, $f = 0.40$ (see Figure 3, Panel A). This effect is predicted by Fitts' law calculations using parameters from Table 12, which reveal smaller mean differences across age using Phone B compared to the Phone A ($M_{\text{Phone B}} = 2.78$ sec; $M_{\text{Phone A}} = 3.32$ sec), as a function of shorter distances between buttons on the smaller Phone B. It was not possible to conduct an ANOVA on accuracy because the only errors that occurred were by older adults on Phone B (2% error rate).

Text messaging experimental task. In accord with hypotheses, results revealed an interaction of Age × Device, such that task completion time differences were minimized between older and younger adults when using Phone A compared to the Phone B, $F(1, 38) = 547.52$, $MSe = 71556.74$, $f = 3.44$ (see Figure 3, Panel B). These results show an opposite performance advantage to the simple dialing task due to greater task complexity in general and a simpler critical pathway inherent to the Phone A hardware and software. A main effect of age was revealed for accuracy in the text messaging task, $F(1, 38) = 4.94$, $MSe = 1.45$, $f = 0.32$, such that older adults made significantly more errors per trial than younger adults ($M_{\text{old}} = 2.63$; $M_{\text{young}} = 2.07$).

Goodness-of-Fit Between Models and Data

Dial-a-number task. The simulated trial for this analysis required users to dial the number "268–413-0734." GOMS task analyses for the dial-a-number task were identical across phones, except for Fitts' law calculations based on button size and button distance as a function of device hardware. As such, a total 11

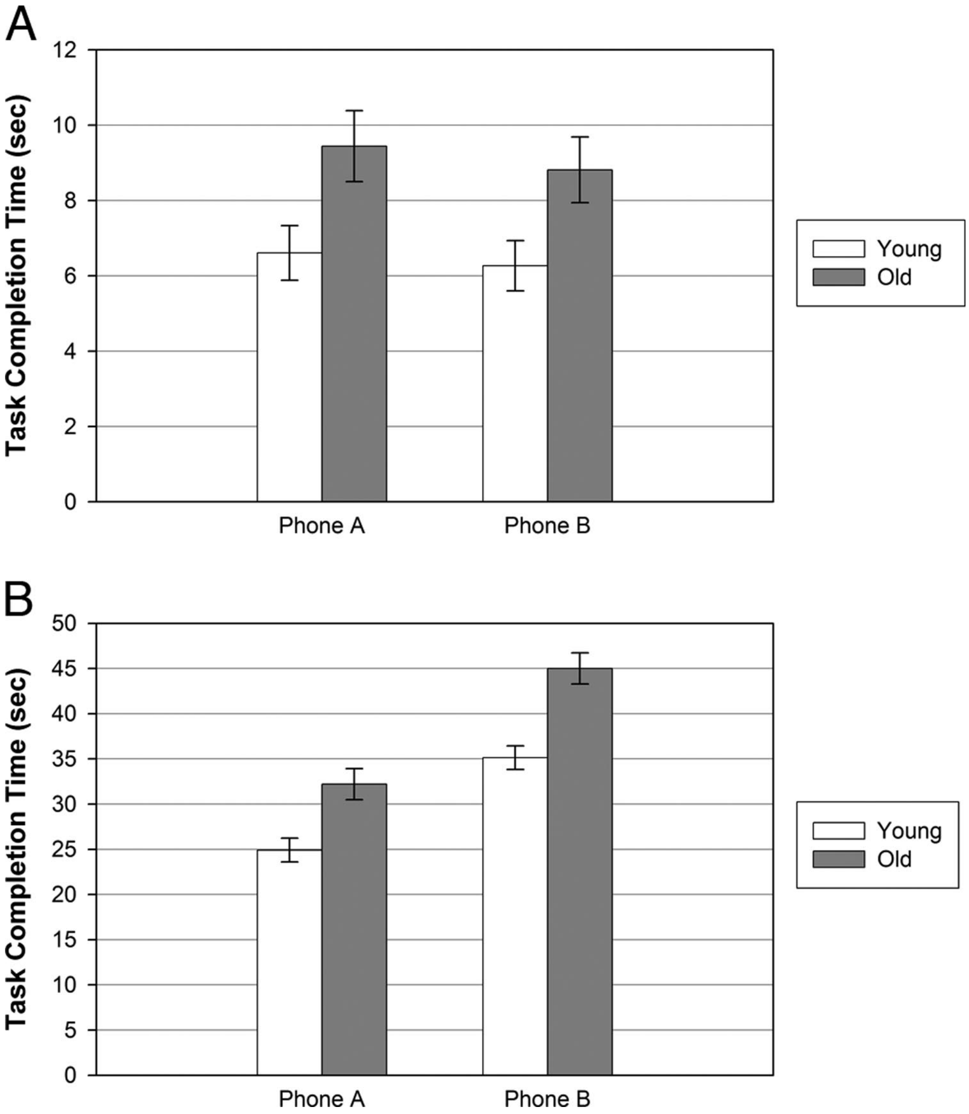

Figure 3. Panel A. Age by device interaction for dial-a-number task with 95% confidence interval bars. Panel B. Age by device interaction for text messaging task with 95% confidence interval bars.

motor movements were required to press buttons (dial the 10 digit number and press send), 17 fixations were required to extract digit information from the trial sheet and verify correct or incorrect actions on the display screen, and 14 cognitive processing cycles were required to make the necessary decisions for how to proceed (see the Appendix, Table A1).

RMSD and *R* (average of point-by-point raw correlations, *r,* calculated for each button press) analyses were calculated to assess the goodness-of-fit between the older adult model parameters and older adult humans and between younger adult model parameters and younger adult humans, for each button press of the critical pathway. As such, 11 data points are analyzed in this dial-a-number task and correspond with the 10 digit phone number required for keyed entry and the final sending of the call by pressing the "send/call" button. Results are displayed in Figures 4 and 5 below.

Further *RMSD* analyses by age reveal that model fits for older adults are as good as model fits for younger adults for both Phone A [$t_{\text{young}}(19) = 0.148$, $t_{\text{old}}(19) = 0.004$] and Phone B [$t_{\text{young}}(19) = 0.148$, $t_{\text{old}}(19) = 0.051$]. To examine the issue of subject variability, separate computations using the lower and upper bounds of parameters, estimated from meta-analytic study standard deviations of means, were conducted. These analyses successfully encompassed individual older adult performance, such that older adult model computations for the Phone A device ranged from 8.6 to 10.1 seconds, while older adult performance ranged from 8.8 to 9.9 seconds; and older adult model computations for the Phone B ranged from 8.1 to 9.6 seconds, while older adult performance ranged from 8.4 to 9 seconds. This suggests that Model Human Processor parameters for older adults are valid simulations of routine older adult human performance in this dialing task.

Additionally, GOMS analyses add credence to the human factors analyses previously discussed, such that task completion times were expected to be minimized between age groups when using Phone B, that performance completion times for the Phone B would be faster than for the Phone A device, and that older adults would perform the task on either device more slowly than younger adults, due to general slowing. These results help illustrate how effective simulations could be for comparing different devices or interfaces even before they had been built.

Text messaging task. The simulated critical pathway for Phone A was composed of 115 steps, made up of 60 eye fixations, 84 cognitive processing cycles, and 36 motor button presses. Phone B was simulated with a critical pathway of 121 steps, made up of 74 eye fixations, 106 cognitive processing cycles, and 38 motor button presses.

RMSD and R^2 analyses were calculated to assess the goodness-of-fit between the older adult model parameters and older adult humans and between younger adult model parameters and younger adult humans, for each button press of the critical pathway. As such, 36 data points are analyzed in this text messaging task for Phone A, and 38 data points are analyzed in this text messaging task for Phone B. Differences in critical pathways across device are due to differences in interface design and hierarchical menu structures, which affect the amount of effort a user must exert to complete the given task. Each data point corresponds with a single button press, such that users must navigate a menu using directional and selection keys and must enter text using the 12-key keypad. Results are displayed in Figures 6 and 7 below.

Further analyses of *RMSD* by age reveal that model fits for older adults are as good as model fits for younger adults for both Phone A [$t_{\text{young}}(19) = 0.015$, $t_{\text{old}}(19) = 0.004$], and Phone B [$t_{\text{young}}(19) = 0.105$, $t_{\text{old}}(19) = 0.051$]. Separate computations using the lower and upper bounds of model parameters successfully encompassed individual older adult performance in this text messaging task, such that older adult model computations for the Phone A device ranged from 31.9 to 38.3 seconds, while older adult performance ranged from 34.5 to 35.6 seconds; and older adult model computations for the Phone B ranged from 42.7 to 48.6 seconds, while older adult performance ranged from 44.6 to 45.3 seconds. This suggests that Model Human Processor parameters for this particular sample of older adults who are healthy and well-

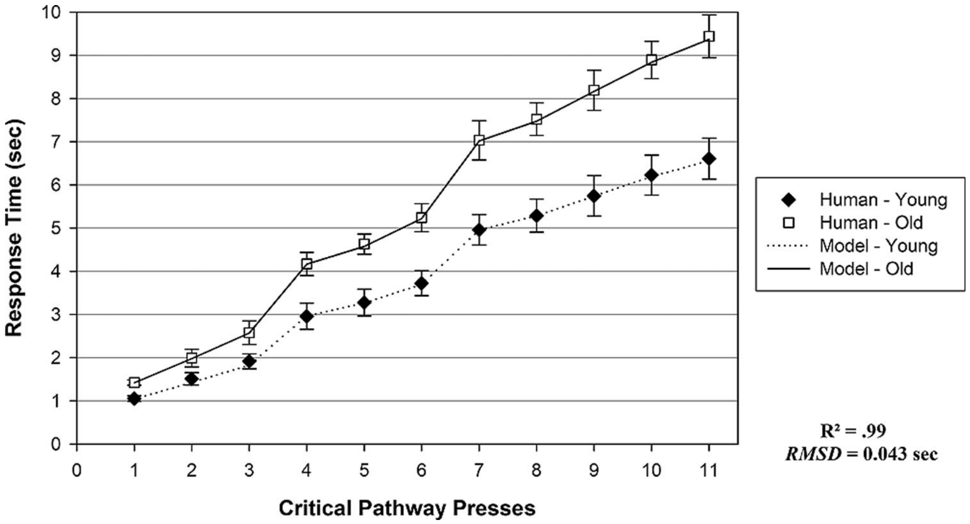

Figure 4. GOMS model fits for Phone A in the dial-a-number task with 95% confidence interval bars around human performance data.

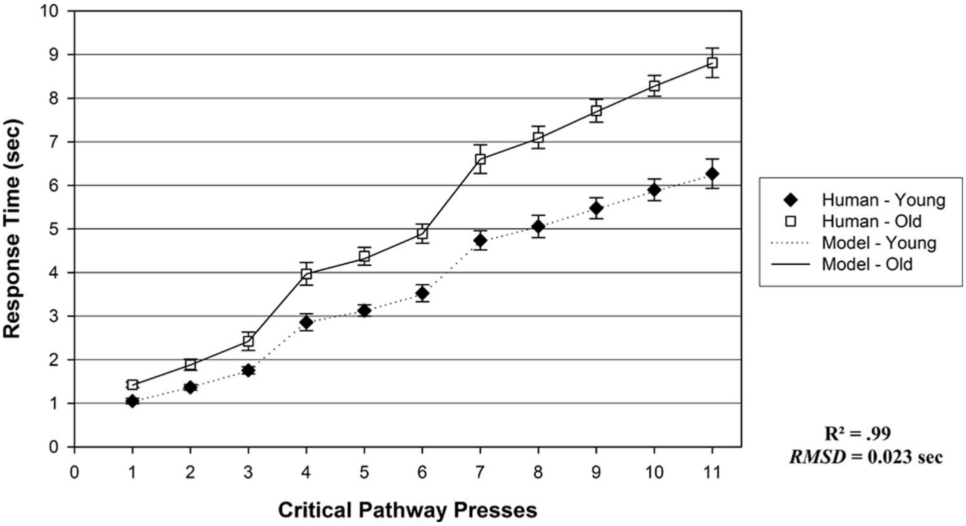

Figure 5. GOMS model fits for Phone B in the dial-a-number task with 95% confidence interval bars around human performance data.

practiced on the task are valid simulations of routine older adult human performance in this more complex text messaging task. In sum, these GOMS analyses add considerable precision to the general human factors predictions and analyses discussed above, such that task completion times were expected to be longer using Phone B due to greater critical path complexity, older adults were expected to perform more slowly than younger adults overall, and performance completion time differences across age were predicted to be minimized when using the simpler Phone A device.

Discussion

A primary goal of this research was to extend human performance modeling techniques to the case of older adult performance. The current research extends Card, Moran, and Newell's (1983) seminal work by using meta-analysis to estimate Model Human Processor parameters for older adults. Additionally, the successful simulation of mobile phone performance using these parameters provides a strong rationale for engineers to utilize GOMS modeling when designing technologies to suit the older user.

From a human factors perspective, this research sheds light on usability concerns as a function of age in the domain of mobile phone design. For simple dialing, the miniaturized keypad for Phone B actually reduced task completion time performance differences across age, compared to the larger keypad of the Phone A device, though at the cost of slightly increased error rates among older adults. This low two percent error rate for simple dialing on Phone B could translate to a potential cost to phone users and to bandwidth providers that may be substantial when considering a multiplier factor of millions of older consumers; however, for consideration of individual efficiency and personal acceptance on

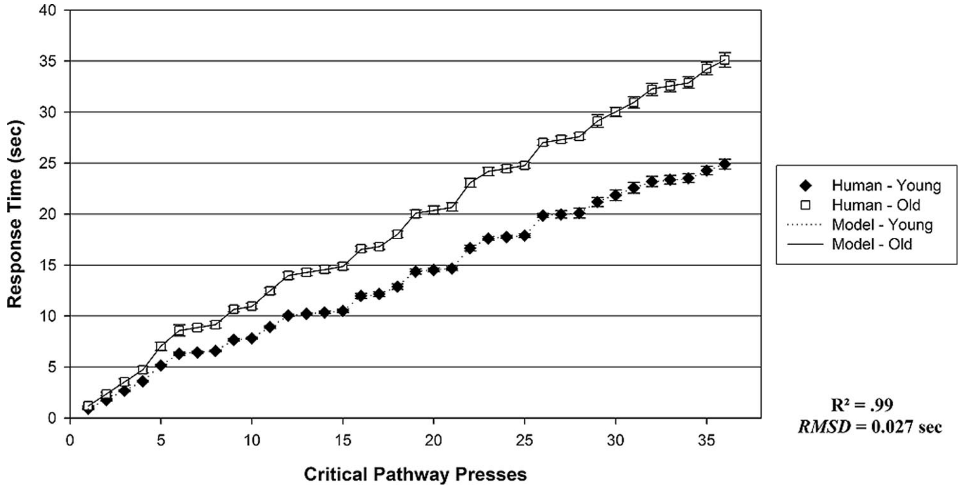

Figure 6. GOMS model fits for Phone A in the text messaging task with 95% confidence interval bars around human performance data.

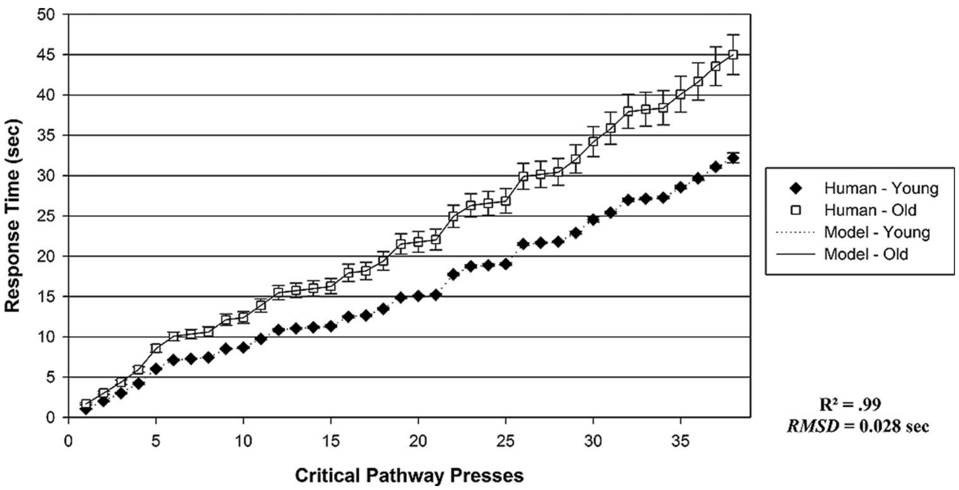

Figure 7. GOMS model fits for Phone B in the text messaging task with 95% confidence interval bars around human performance data.

a daily basis, this very low error rate may be perfectly satisfactory for the typical older adult.

For more complicated mobile phone tasks like text messaging, older adults may be less likely to engage this specific feature in everyday life than for simple and straightforward dialing tasks; however, given that text entry is required for entering contact information into a phone directory, some aspects of the investigated task may translate to problems using more common features on mobile phones. As such, this study revealed significantly smaller performance differences across age when using the least complex device, Phone A, for both time ($f = 18.2$; see Figure 3, Panels A and B) and accuracy ($f = 0.21$). Phone A possessed larger, easier buttons to press and navigate menus with and possessed a menu hierarchy requiring less cognitive workload than that of Phone B.

This research also validated a subset of the older adult parameter estimates by comparing goodness-of-fit for performance predictions to previously validated younger adult parameters (Cardet al., 1983). A limitation of this validation is that we simulated just two phone models using two representative phone tasks and the simulations did not include all the parameter values, so there is more work to be done on validation. Younger adult performance models served as a check that critical pathway calculations had been correctly composed, and goodness-of-fit from model to human data demonstrated that the novel meta-analytic method used to estimate weighted means for older adult Model Human Processor parameters was indeed appropriate. Meta-analyses for older adult task components indicated general slowing of perceptual, cognitive, and motor processing cycles in the range of 1.5–2 times that of younger adults (e.g., as suggested in Fisk et al., 2004); however, individual task components may skew estimates for response time, such that, when eye movements play a weighty role for instance, old-young differences may be considerably reduced given the 1.2:1 ratio observed for saccade duration. As such, designers may be able to design critical paths for new products that minimize such old-young performance differences by careful consideration of such parameters.

In terms of model predictions, this research revealed that older adult models produced excellent fits to older adult data across mobile phone tasks of increasing complexity and comparable fits to previously validated younger adult models. Goodness-of-fit across analysis of each button press in each task and device correlated with human data at a level of approximately $R = 0.99$, for both young and old, showing that not only is the older adult model as good as the younger adult model, but that it also predicts practiced, routine human performance at a very precise level. Close examination of task analyses for younger and older users also illuminated portions of critical pathways that made performance more sluggish or more error-prone. This type of examination, made available by this type of modeling, may enable designers to pinpoint problematic features of designs easily and assist them in producing better future designs.

Menu structure in the text messaging task also was shown have on impact on the performance of older adults. Because the Phone B screen allowed room for one less line of text than Phone A's screen and because the menu structure required deeper navigation, older adults encountered greater difficulty integrating information across screens and had to sift through more extraneous information unrelated to the goal of the task. The more streamlined menu structure of Phone A, with less information to process on each screen, proved to aid older adults in menu navigation, as revealed by faster completion times and higher levels of accuracy.

Other problems revealed for older adults, more so than for younger adults, included buttons that were too small or required very fine motor control to press in the intended direction (e.g., four-way navigational buttons on the Phone B). Modelers may use Fitts' law calculations based on older adult movement times to decide how large buttons should be, as well as how far apart buttons should be, as small, closely placed buttons increase the index of difficulty and require slower movements for the target to be correctly pressed. This implies that modeling with respect to age is a practical way to test usability of different interfaces, devices, and systems and also provides a rationale to simulate older adult performance with prototypical devices in the very earliest stages of design.

In sum, using predictive models to simulate trained users is a pragmatic approach to testing design specifications without building prototypes or training and testing users. Models, at their best, are valuable, informative tools for designers to make use of and compare methods or strategies that may be employed to complete a task (Bellman & MacKenzie, 1998; MacKenzie & Zhang, 1999; Silfverberg, MacKenzie, & Korhonen, 2000). Although the current research investigated usability of two existing products, future extensions of this work could address whether simulations along these lines provide a valid means of testing hypothetical prototypes with hypothetical critical paths for the older adult population. If extensions prove robust, this technique holds the promise that design flaws could be removed before production, and the resulting effective interfaces could allow for more user-friendly technology interactions. Application of these parameters could indeed successfully bridge the fields of cognitive aging and human engineering design.

Caveats

The conceptualization of cognitive processes that dominated in the early 1980s, at the time of publication of the original model human processor parameters, has undergone considerable change. Nonetheless, present day cognitive modeling using this “first approximation” characterization of the human information processing system has proven to be quite robust (e.g., via the EPIC architecture: Meyer & Kieras, 1999; or via modifications to Anderson’s (1996) ACT-R: John, Prevas, Salvucci & Koedinger, 2004). With regard to the parameters estimated here for older adults, we note that estimations were often based on very small numbers of studies with very small sample sizes; thus, additional investigations should proceed to assess robustness of validity and generality. Also, typical older adult samples are biased ones in the sense that only relatively healthy, cognitively intact older adults typically volunteer for lab-based research (Camp, West, & Poon, 1989; Todd, Davis, & Cafferty, 1983). Hence, these estimates are likely to be biased in a positive direction. However, a similar bias may exist for the young adults (e.g., college undergraduates compared to those with high school only education levels). Also, as discussed earlier, the Model Human Processor parameters given for young adults were not estimated as systematically as we have done for older adults. Thus, similar meta-analytic techniques could be applied to generate more precise estimates for younger adults.

Furthermore, the type of modeling utilized in this research is deliberately approximate in order to account for the “typical” older user, and includes the level of detail necessary only to predict performance for precise tasks with exact device specifications (Card et al., 1983). Thus, generalizations to other populations, tasks, or devices would not be possible without additional analyses. Similarly, because GOMS task analyses are designed to handle routine performance, this work required adequate practice sessions to ensure that participants achieved consistent performance (trials were within one second of each other). Therefore, results from this work could not easily be extended to make predictions for novel, unpracticed tasks.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abel, L. A., Troost, B. T., & Dell’Osso, L. F. (1983). The effects of age on normal saccadic characteristics and their variability. *Vision Research, 23*(1), 33–37.

*Allen, P. A., & Coyne, A. C. (1989). Are there age differences in chunking? *Journals of Gerontology, 44*(6), P181–P183.

*Allen, P. A., & Crozier, L. C. (1992). Age and ideal chunk size. *Journals of Gerontology, 47*(1), P47–P51.

*Allen, P. A., Murphy, M. D., Kaufman, M., Groth, K. E., & Begovic, A. (2004). Age differences in central (semantic) and peripheral processing: The importance of considering both response times and errors. *Journal of Gerontology: Psychological Sciences and Social Sciences,* 59B, P210–P219.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51,* 355–365.

*Anderson, N. D. (1999). The attentional demands from encoding and retrieval in younger and older adults: 2. Evidence from secondary task reaction time distributions. *Psychology and Aging, 14*(4), 645–655.

*Backman, L., & Small, B. J. (1998). Influences of cognitive support on episodic remembering: Tracing the process of loss from normal aging to Alzheimer’s disease. *Psychology and Aging, 13,* 267–276.

*Basak, C., & Verhaeghen, P. (2003). Subitizing speed, subitizing range, counting speed, the Stroop effect, and aging: Capacity differences, speed equivalence. *Psychology and Aging, 18,* 240–249.

Bashore, T. R., Osman, A., & Heffley, E. F. (1989). Mental slowing in elderly persons: A cognitive psychophysiological analysis. *Psychology and Aging, 4*(2), 235–244.

*Batsakes, P. J., & Fisk, A. D. (2000). Age and dual-task performance: Are performance gains retained? *Journal of Gerontology: Psychological Sciences,* 55B, P332–P342.

Bay, S., & Ziefle, M. (2005). Children using cell phones: The effects of shortcomings in user interface design. *Human Factors, 47,* 158–168.

Bellman, T., & MacKenzie, I. S. (1998). A probabilistic character layout strategy for mobile text entry. In *Proceedings of Graphics Interface ’98* (pp. 168–176). Toronto: Canadian Information Processing Society.

Benson, V., & Marano, M. A. (1998). Current estimates from the National Health Interview Survey, 1995. National Center for Health Statistics. Retrieved September 6, 2006, from http://www.cdc.gov/nchs/data/series/sr_10/sr10_199acc.pdf

*Bherer, L., Kramer, A. F., Peterson, M., Colcombe, S. J., Erickson, K. I., & Becic, E. (2005). Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging, 20,* 695–709.

*Bono, F., Oliveri, R. L., Zappia, M., Aguglia, U., Puccio, G., & Quattrone, A. (1996). Computerized analysis of eye movements as a function of age. *Archives of Gerontology and Geriatrics, 22*(3), 261–169.

*Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60,* P223–P233.

Bosman, E. A. (1993). Age-related differences in the motoric aspects of transcription typing skill. *Psychology and Aging, 1,* 87–102.

Botwinick, H. R., & Storandt, M. (1974). *Memory, relation function and age.* Springfield, IL: Charles C Thomas.

Bouma, H. (2001). Creating adaptive technological environments. *Gerontology, 1,* 1–3.

Brogmus, G. E. (1991). Effects of age and sex on speed and accuracy of hand movements and the refinements they suggest for Fitts’ law. Paper presented at the Human Factors Society 35th Annual Meeting, San Francisco.

Camp, C. J., West, R. L., & Poon, L. W. (1989). Recruitment practices for psychological research in gerontology. In M. P. Lawton & A. R. Herzog

(Eds.), *Special research methods for gerontology* (pp. 163–189). Amityville, NY: Baywood.

Card, S. K., & Moran, T. P. (1988). User technology: From pointing to pondering. In A. Goldberg (Ed.), *A history of personal workstations* (pp. 183–198). New York: ACM Press.

Card, S. K., Moran, T. P., & Newell, A. (1980). The Keystroke-Level model for user performance time with interactive systems. *Communications of the ACM, 23*(7), 396–410.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

Card, S. K., Moran, T. P., & Newell, A. (1986). The model human processor: An engineering model of human performance. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance* (pp. 1–35). New York: Wiley and Sons.

*Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30,* 1078–1085.

*Chan, F., Armstrong, I. T., Pari, G., Riopelle, R. J., & Munoz, D. P. (2005). Deficits in saccadic eye-movement control in Parkinson's disease. *Neuropsychologia, 43,* 784–796.

*Chaparro, A., Bohan, M., Fernandez, J. E., Choi, S. D., & Kattel, B. (1999). The impact of age on computer input device use: Psychophysical and physiological measures. *International Journal of Industrial Ergonomics, 24,* 503–513.

Charness, N. (1981a). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 467–476.

Charness, N. (1981b). Visual short-term memory and aging in chess players. *Journals of Gerontology, 36,* 615–619.

Charness, N. (1981c). Aging and skilled problem solving. *Journal of Experimental Psychology: General, 110,* 21–38.

Charness, N., & Bosman, E. A. (1990). Expertise and aging: Life in the lab. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 343–385). Amsterdam: Elsevier

*Charness, N., & Campbell, J. I. D. (1988). Acquiring skill at mental calculation in adulthood: A task decomposition. *Journal of Experimental Psychology: General, 117,* 115–129.

Charness, N., & Czaja, S. J. (2005). Adaptation to new technologies (7.13). In M. L. Johnson (Ed.), *Cambridge handbook on age and ageing* (pp. 662–669). Cambridge: UK.

Charness, N., Holley, P., Feddon, J., & Jastrzembski, T. (2004). Light pen versus mouse for a menu selection task: Age, hand, and practice effects. *Human Factors, 46,* 373–384.

Charness, N., & Jastrzembski, T. (2001). Cognition and aging: A human engineering model. *International Workshop on Gerontology,* Tsukuba City, Japan, March 2001.

Charness, N., Krampe, R., & Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.), *The road to excellence* (pp. 51–80). Mahwah, NJ: Erlbaum.

Chattratichart, J., & Brodie, J. (2003). The age factor in the design equation of cell phones. *Proceedings of the 12th Annual Usability Professionals' Association Conference,* June 23–27, Arizona.

*Cheatham, P. G., & White, C. T. (1954). Temporal numerosity: III. Auditory perception of number. *Journal of Experimental Psychology, 47,* 425–428.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms and performance* (pp. 409–422). Amsterdam: Elsevier.

Craik, F. I. M., & Salthouse, T. A. (2000). *The handbook of aging and cognition.* Mahwah, NJ: Erlbaum.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., et al (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21,* 333–352.

*Dennis, W., Scialfa, C. T., & Ho, G. (2004). Age differences in feature selection in triple conjunction search. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 59,* P191–P198.

Dijkstra, K., Charness, N., & Yordon, R. (2007). Changes in physiological and self-reported stress in younger and older adults as a result of relaxation and stress induction. Manuscript in preparation.

*Fisk, A. D., Cooper, B. P., Hertzog, C., Anderson-Garlach, M. M., & Lee, M. D. (1995). Understanding performance and learning in consistent memory search: An age-related perspective. *Psychology and Aging, 10*(2), 255–268.

Fisk, A. D., & Rogers, W. A. (Eds.). (1997). *Handbook of human factors and the older adult.* San Diego, CA: Academic Press.

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2004). *Designing for older adults: Principles and creative human factors approaches.* Boca Raton: CRC Press.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, xlvii,* 381–391.

*Floden, D., Stuss, D. T., & Craik, F. I. M. (2000). Age differences in performance on two versions of the Brown-Peterson task. *Aging, Neuropsychology, and Cognition, 7,* 245–259.

Fox, S. (2004, March 25). Older Americans and the internet. *Pew Internet and American Life Project*Retrieved September 6, 2006, from http://www.pewinternet.org/pdfs/PIP_Seniors_Online_2004.pdf

Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age differences and changes in reaction time: The Baltimore Longitudinal Study of Aging. *Journal of Gerontology, 49*(4), P179–P189.

Ganz, L. (1975). Temporal factors in visual perception. In E. C. Carterette and M. P. Friedman (Eds.), *Handbook of perception: Volume V, Seeing* (pp. 169–230). New York: Academic Press.

*Glass, J. M., Schumacher, E. H., Lauber, E. J., Zurbriggen, E. L., Gmeindl, L., Kieras, D. E., et al. (2000). Aging and the psychological refractory period: Task-coordination strategies in young and old adults. *Psychology and Aging, 15*(4), 571–595.

Goggin, N. L., & Meeuwsen, H. J. (1992). Age-related differences in the control of spatial aiming movements. *Research Quarterly for Exercise and Sport, 63*(4), 366–372.

*Gottlob, L. R. (2006). Aging and comparative search for feature differences. *Aging, Neuropsychology, and Cognition, 13,* 435–457.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied, 6*(4), 322–335.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: A validation of GOMS for prediction and explanation of real-world task performance. *Human-Computer Interaction, 8*(3), 237–309.

*Gregoire, J., & Van der Linden, M. (1997). Effects of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4*(2), 140–149.

Haaland, K. Y., Harrington, D. L., & Grice, J. W. (1993). Effects of aging on planning and implementing arm movements. *Psychology and Aging, 5,* 617–632.

*Haarmann, H. J., Ashling, G. E., Davelaar, E. J., & Usher, M. (2005). Age-related declines in context maintenance and semantic short-term memory. *The Quarterly Journal of Experimental Psychology,* 58A, 34–53.

Hale, S., & Myerson, J. (1995). Fifty years older, fifty percent slower? Meta-analytic regression models and semantic context effects. *Aging and Cognition, 2,* 132–145.

*Hale, S., Myerson, J., Faust, M., & Fristoe, N. (1995). Converging evidence for domain-specific slowing from multiple nonlexical tasks and multiple analytic methods. *Journals of Gerontology: Psychological Sciences,* 50B, P202–P211.

Harrington, T. L., & Harrington, M. K. (2000). *Gerontechnology: Why and how*. Maastricht, The Netherlands: Shaker.

Harter, M. R. (1967). Excitability and cortical scanning: A review of two hypotheses of central intermittency in perception. *Psychological Bulletin, 68,* 47–58.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 193–225). San Diego, CA: Academic Press, Inc.

*Hertzog, C., Cooper, B. P., & Fisk, A. D. (1996). Aging and individual differences in the development of skilled memory search performance. *Psychology and Aging, 11,* 497–520.

Hines, T. (1979). Information feedback, reaction time and error rates in young and old subjects. *Experimental Aging Research, 5*(3), 207–215.

*Ho, G., & Scialfa, C. (2002). Age, skill transfer, and conjunction search. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57,* P277–P287.

Ho, G., Scialfa, C., Graw, T., & Caird, J. (2001). Conspicuity of traffic signs: The effect of clutter, luminance, and aging. *Human Factors, 43,* 194–207.

*Hofer, S. M., Berg, S., & Era, P. (2003). Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychology and Aging, 18,* 285–305.

*Hoyer, W. J., Cerella, J., & Onyper, S. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition, 31,* 1260–1270.

Huaman, A. G., & Sharpe, J. A. (1993). Vertical saccades in senescence. *Investigative Ophthalmology & Visual Science, 34,* 2588–2595.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: correcting error and bias in research findings.* Beverly Hills, CA: Sage.

International Telecommunications Union. (2006). *Telecommunications/ICT Statistics.* Retrieved January 29, 2007 from http://www.itu.int/ITU-D/ict/statistics

*Jagacinski, R. J., Greenberg, N., Liao, M., & Wang, J. (1993). Manual performance of a repeated pattern by older and younger adults with supplementary auditory cues. *Psychology and Aging, 8,* 429–439.

*Jagacinski, R. J., Liao, M., & Fayyad, E. A. (1995). Generalized slowing in sinusoidal tracking by older adults. *Psychology and Aging, 10*(1), 8–19.

*James, C. L., & Reischel, K. M. (2001). Text input for mobile devices: Comparing model prediction to actual performance. *Proceedings of CHI 2001,* March 31–April 5, 365–371.

*Jamieson, B. A., & Rogers, W. A. (2000). Age-related effects of blocked and random practice schedules on learning a new technology. *Journals of Gerontology: Psychological Sciences,* 55B, P343–P353.

Jastrzembski, T. S., Charness, N. H., Holley, P., & Feddon, J. (2005). Input devices for web browsing: Age and hand effects. *Universal Access in the Information Society, 4*(1), 39–45.

*Jenkins, L., Myerson, J., Joerding, J. A., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging, 15,* 157–175.

John, B. E. (1990). Extensions of GOMS analyses to expert performance requiring perception of dynamic visual and auditory information. *Proceedings of the Human Factors in Computing Systems, CHI '90 Conference.* New York: ACM Press.

John, B. E., & Kieras, D. E. (1994). The GOMS family of analysis techniques: Tools for design and evaluation. *Carnegie Mellon University School of Computer Science Tech* (No. CMU-CS-94–181).

John, B. E., Prevas, K., Salvucci, D. D., & Koedinger, K. (2004). Predictive human performance modeling made easy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 455–462). New York: ACM Press.

*Johnson, M. K., Reeder, J. A., Raye, C. L., & Mitchell, K. J. (2002). Second thoughts versus second looks: An age-related deficit in selectively refreshing just-active information. *Psychological Science, 13,* 64–67.

*Kahana, M. J., Howard, M. W., Zaromb, F. M., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 530–540.

Kantowitz, B. H., & Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships.* New York: Wiley.

*Kemper, S., Crow, A., & Kemtes, K. (2004). Eye fixation patterns of high and low span young and older adults: Down the garden path and back again. *Psychology and Aging, 19,* 157–170.

*Kemper, S., Herman, R. E., & Lian, C. H. T. (2003). The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech or noise. *Psychology and Aging, 18,* 181–192.

*Ketcham, C. J., Seidler, R. D., Van Gemmert, A. W. A., & Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target-size, and movement amplitude. *Journals of Gerontology: Psychological Sciences and Social Sciences,* 57B, P54–P64.

*Kline, D. W. (1994). Optimizing the visibility of displays for older observers. *Experimental Aging Research, 20*(1), 11–13.

Kline, D. W., & Scialfa, C. T. (1996). Visual and auditory aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 181–203) San Diego: Academic Press.

*Kosnik, W., Kline, D. W., Fikre, J., & Sekuler, R. (1987). Ocular fixation control as a function of age and exposure duration. *Psychology and Aging, 2,* 302–305.

*Kramer, A. F., Boot, W. R., McCarley, J. S., Peterson, M. S., Colcombe, A. F., & Scialfa, C. T. (2006). Aging, memory, and visual search. *Acta Psychologica, 122,* 288–304.

Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General, 125,* 331–359.

Lemelson-MIT Program (2004). *Cell phone edges alarm clock as most hated invention, yet one we cannot live without*. Retrieved March 23, 2007, from http://web.mit.edu/invent/n-pressreleases/n-press-04index.html

*Leonard, J. A. (1952). *Some experiments on the temporal relation between information and action.* Unpublished doctoral thesis. Cambridge University.

Lerch, F. J., Mantei, M. M., & Olson, J. R. (1989). Skilled financial planning: The cost of translating ideas into action. In *Proceedings of the ACM CHI 1989 conference on human factors in computing systems* (pp. 121–126). New York: ACM Press.

*Liao, M., Jagacinski, R. J., & Greenberg, N. (1997). Quantifying the performance limitations of older and younger adults in a target acquisition task. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1644–1664.

*Light, L. L., & Anderson, P. A. (1985). Working memory capacity, age, and memory for discourse. *Journal of Gerontology, 40,* 737–747.

*Lu, Z-L., Neuse, J., Madigan, S., & Dosher, B. (2005). Decay of iconic memory in observers with mild cognitive impairment. *Proceedings of the National Academy of Science, 102,* 1797–1802.

MacKenzie, I. S., & Zhang, S. (1999). The design and evaluation of a high-performance soft keyboard. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '99* (pp. 25–31). New York: ACM.

*Madden, D. J., Turkington, T. G., Provenzale, J. M., Denny, L. L., Langley, L. K., Hawk, T. C., et al. (2002). Aging and attentional guidance during visual search: Functional neuroanatomy by position emission tomography. *Psychology and Aging, 17*(1), 24–43.

*Madden, D. J., Whiting, W. L., Spaniol, J., & Bucur, B. (2005). Adult age differences in the implicit and explicit components of top-down attentional guidance during visual search. *Psychology and Aging, 20*(2), 317–329.

*Maltz, M., & Shinar, D. (1999). Eye movements of younger and older drivers. *Human Factors, 41*(1), 15–25.

*Marshall, N. K., Bashore, T. R., Miller, G. A., Coles, M. G. H., & Donchin, E. (1983). ERP consistency in young-adult and elderly subjects. *Psychophysiology, 20*(4), 422.

*Maylor, E. A., Vousden, J., & Brown, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology and Aging, 14,* 572–594.

*McCrea, P. H., & Eng, J. J. (2005). Consequences of increased neuromotor noise on reaching movements in persons with stroke. *Experimental Brain Research, 162,* 70–77.

*McFarland, R., Warren, B., & Karis, C. (1958). Alterations in critical flicker frequency as a function of age and light:dark ratio. *Journal of Experimental Psychology, 56*(6), 529–538.

Meinz, E. J. (2000). Experience-based attenuation of age-related differences in music cognition tasks. *Psychology and Aging, 15,* 297–312.

*Meyer, D. E., Glass, J. M., Mueller, S. T., Seymour, T. L., & Kieras, D. E. (2001). Executive-process interactive control: A unified computational theory for answering twenty questions (and more) about cognitive aging. In U. Mayr, D. Spieler, & R. Kliegl (Eds.), *Aging and executive control* (pp. 123–164). Hove, UK: Psychology Press.

Meyer, D. E., & Kieras, D. E. (1999). Precis to a practical unified theory of cognition and action: Some lessons from computational modeling of human multiple-task performance. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 17–88). Cambridge, MA: MIT Press.

*Miles, W. R. (1931). Measures of certain human abilities throughout the life span. *Proceedings of the National Academies of Science, xvii,* 627–633.

Morrow, D. G. (2003). Technology as environmental support for older adults' daily activities. In N. Charness & K. W. Schaie (Eds.), The impact of technology on successful aging (pp. 290–305). New York: Springer.

Mortimer, J. A., Pirozzolo, F. J., & Maletta, G. J. (Eds.) (1982). *The aging motor system.* New York: Praeger.

*Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research, 121,* 391–400.

*Murphy, D. R., Craik, F. I., M., & Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging, 15,* 323–334.

*Myerson, J., Hale, S., Rhee, S. H., & Jenkins, L. (1999). Selective interference with verbal and spatial working memory in young and older adults. *Journals of Gerontology: Series B: Psychological Sciences & Social Sciences, 54,* P161–P164.

National Science Foundation, Division of Science Resources Statistics. (2006). *Science and Engineering Indicators 2006.* Retrieved March 21, 2007, from http://www.nsf.gov/statistics/seind06/toc.htm

*Nielsen-Bohlman, L., & Knight, R. T. (1999). Prefrontal cortical involvement in visual working memory. *Cognitive Brain Research, 8*(3), 299–310.

Nielson, J. (1994). *Usability engineering.* San Francisco: Morgan Kaufmann.

*Nilsson, L. G. (2003). Memory function in normal aging. *Acta Neurologica Scandinavica, 107,* 7–13.

NTT DoCoMo (2000). Mobile phones increasingly popular among the elderly. *Press Release 34(11).* Retrieved October 11, 2007, from: http://www.nttdocomo.com/files/presscenter/34_No11_Doc.pdf

Office for National Statistics (2003). Adult mobile phone ownership or use by age, 2001 and 2003. *Social Trends, 34.* Retrieved on October 11, 2007, from http://www.statistics.gov.uk/STATBASE/ssdataset.asp?vlnk=7202

*Olincy, A., Ross, R. G., Young, D. A., & Freedman, R. (1997). Age diminishes performance on an antisaccade eye movement task. *Neurobiology of Aging, 18,* 483–489.

Olson, J. R., & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Human-Computer Interaction, 5,* 221–265.

Pew Internet & American Life Project. (2006). Demographics of internet use. Retrieved March 21, 2007, from http://www.pewinternet.org/trends.asp

*Picton, T. W., Champagne, S. C., Cerri, A. M., Stuss, D. T., & Nelson, R. F. (1986). The effects of age and task difficulty on the late positive component of the auditory evoked potentials. *Electroenceph Clin Neurophysiol, 38,* 132–133.

*Picton, T. W., Stuss, D. T., Champagne, S. C., & Nelson, R. F. (1984). The effects of age on human event-related potentials. *Psychophysiology, 21,* 312–325.

Pirolli, P. (1999). Cognitive architectures and cognitive engineering models in human-computer interaction. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of applied cognition* (pp. 441–447). New York: Wiley.

Pirozzolo, F. J., & Hansch, E. C. (1981). Oculomotor reaction time in dementia reflects degree of cerebral dysfunction. *Science, 214,* 349–351.

*Podlesny, J. A., & Dustman, R. E. (1982). Age effects on heart rate, sustained potential, and P3 responses during reaction-time tasks. *Neurobiology of Aging, 3,* 1–9.

Pohl, P. S., Winstein, C. J., & Fisher, B. E. (1996). The locus of age-related movement slowing: Sensory processing in continuous goal-directed aiming. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 51*(2), P94–P102.

Rabbitt, P. (1965). An age-decrement in the ability to ignore irrelevant information. *Journal of Gerontology, 20,* 233–238.

Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Quarterly Journal of Experimental Psychology, 29,* 727–743.

Rabbitt, P., & Vyas, S. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica, 33,* 56–76.

*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. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language, 50,* 408–424.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik and T. A. Salthouse (Eds.), *Handbook of aging and cognition – II* (pp. 1–90). Mahwah, NJ: Erlbaum.

*Rochon, E., Waters, G. S., & Caplan, D. (2000). Working memory and sentence comprehension in patients with Alzheimer's disease: Evidence from multiple measures of working memory capacity and sentence comprehension. *Journal of Speech, Language and Hearing Research, 43*(2), 395–413.

*Rodrigue, K. M., Kennedy, K. M., & Raz, N. (2005). Aging and longitudinal change in perceptual-motor skill acquisition in healthy adults. *Journal of Gerontology: Psychological Sciences,* 60B, P174–P181.

*Rogers, W. A., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a screen or turn a knob: Choosing the best device for the job. *Human Factors, 47*(2), 271–288.

Rogers, W. A., Fisk, A. D., Mead, S., Walker, N., & Cabrera, E. F. (1996). Training older adults to use automatic teller machines. *Human Factors, 38,* 425–433.

*Romero, D. H., Van Gemmert, A. W. A., Adler, Ch. H., Bekkering, H., & Stelmach, G. E. (2003). Altered aiming movements in Parkinson's disease patients and elderly adults as a function of delays in movement onset. *Experimental Brain Research, 151*(2), 249–261.

Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General, 13,* 345–371.

*Salthouse, T. A. (1988). Cognitive aspects of motor functioning. *Annals of the New York Academy of Sciences, 515,* 33–41.

Salthouse, T. A. (1991). *Theoretical perspectives on cognition and aging.* Hillsdale, N J: Erlbaum Press.

Salthouse, T. A. (1994). Age-related differences in basic cognitive processes: Implications for work. *Experimental Aging Research, 20,* 249–255.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye fixation duration. *American Journal of Psychology, 93,* 207–234.

*Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory and speed. *Journals of Gerontology: Psychological Sciences,* 50B, P297–P306.

*Scialfa, C. T., Jenkins, L., Hamaluk, E., & Skaloud, P. (2000). Aging and the development of automaticity in conjunction search. *Journal of Gerontology: Psychological Sciences,* 55B, P27–P46.

Scialfa, C. T., & Joffe, K. M. (1997). Age differences in feature and conjunction search: Implications for theories of visual search and generalized slowing. *Aging, Neuropsychology, & Cognition, 4,* 227–246.

Sekuler, R., & Sekuler, A. B. (2000). Vision and aging. In A. Kazdin (Ed.), *Encyclopedia of psychology, Vol. 8* (pp. 180–183). American Psychological Association and Oxford University Press.

Silfverberg, M., MacKenzie, I. S., & Korhonen, P. (2000). Predicting text entry speed on mobile phones. *Proceedings of CHI, 2000,* 9–16.

*Sliwinski, M. (1997). Aging and counting speed. Evidence for process-specific slowing. *Psychology and Aging, 12,* 38–49.

*Sliwinski, M., Buschke, H., Kuslansky, G., Senior, G., & Scarisbrick, D. (1994). Proportional slowing and addition speed in old and young adults. *Psychology and Aging, 9,* 72–80.

*Smith, C. D., Umberger, G. H., Manning, E. L., Slevin, J. T., Wekstein, D. R., Schmitt, F. A., et al. (1999). A critical decline in fine motor hand movements in human aging quantified by an automated activities panel. *Neurology, 53*(7), 1391–1396.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74,* 1–29.

Spirduso, W. W. (1995). *Physical dimensions of aging.* Champaign, IL: Human Kinetics.

Spirduso, W. W., & MacRae, P. G. (1990). Motor performance and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 183–200) Orlando, FL: Academic Press.

St. Amant, R., Horton, T. E., & Ritter, F. E. (2004). Model-based evaluation of cell phone menu interaction. *Proceedings of the ACM Conference on Human Factors in Computing Systems,* Vienna, Austria.

Stelmach, G. E., & Goggin, N. L. (1988). Psychomotor decline with age. *Physical Activity and Aging, 22,* 6–17.

*Stine, E. L., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging, 2*(3), 272–279.

*Strayer, D., & Kramer, A. (1994). Aging and skill acquisition: Learning-performance distinctions. *Psychology and Aging, 9,* 589–605.

*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.

Todd, M., Davis, K. E., & Cafferty, T. P. (1983). Who volunteers for adult development research?: Research findings and practical steps to reach low volunteering groups. *International Journal of Aging and Human Development, 18,* 177–184.

*Touron, D. R., & Hertzog, C. (2004). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging, 19,* 452–466.

*Touron, D. R., Hoyer, W. J., & Cerella, J. (2001). Cognitive skill acquisition and transfer in younger and older adults. *Psychology and Aging, 16,* 555–563.

*Tremblay, K. L., Piskosz, M., & Souza, P. (2002). Aging alters the neural representation of speech-cues. *NeuroReport, 13,* 1865–1870.

*Underwood, G. Phelps, N., Wright, C., van Loon, E., & Galpin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic & Physiological Optics, 25*(4), 346–356.

United States Census Bureau. (2004). *U.S. Projections by age, sex, race, and origin*. Retrieved April 20, 2005, from http://www.census.gov

Vercruyssen, M. (1997). Movement control and speed of behavior. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 55–86). San Diego, CA: Academic Press.

*Verhaeghen, P. (2002). Age differences in efficiency and effectiveness of encoding for visual search and memory search: A time-accuracy study. *Aging, Neuropsychology, and Cognition, 9,* 114–126.

*Verhaeghen, P., & De Meersman, L. (1998). Aging and negative priming: A meta-analysis. *Psychology and Aging, 13,* 435–444.

*Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journals of Gerontology: Psychological Sciences and Social Sciences, 48,* P157–P171.

*Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging, 18,* 443–460.

*Verhaeghen, P., Vandenbroucke, A., & Dierckx, V. (1998). Growing slower and less accurate: The effects of age on time-accuracy functions for recall from episodic memory. *Experimental Aging Research, 24,* 3–19.

*Walker, N., Philbin, D. A., & Fisk, A. D. (1997). Age-related differences in movement control: Adjusting submovement structure to optimize performance. *Journal of Gerontology: Psychological Sciences,* 52B, P40–P52.

*Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers, 35*(4), 550–564.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*(2), 89–104.

Welford, A. T. (1958). *Ageing and human skill.* London: Oxford University Press.

Welford, A. T. (1968). *Fundamentals of skill.* London: Methuen.

Welford, A. T. (1973). *Ageing and human skill.* Westport, CT: Greenwood Press. (Original work published 1958).

Welford, A. T. (1974). Introductory lecture to session IV: On the sequencing of action7. *Brain Research, 74,* 381–392.

Welford, A. T. (1977). Motor performance. In J. E. Birren and K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 450–496). Van Nostrand Reinhold, New York.

Westheimer, G. (1954). Mechanism of saccadic eye movements. *Archives of Opthalmology, 52,* 710–724.

*Worden, A., Walker, N., Bharat, K., & Hudson, S. (1997). Making computers easier for older adults to use: Area cursors and sticky icons. *Proceedings of the CHI '97 Conference on Human Factors in Computing Systems* (pp. 266–271). New York: ACM.

Ziefle, M. (2002). The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behaviour & Information Technology, 1,* 303–311.

Ziefle, M., & Bay, S. (2005). How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behaviour & Information Technology, 24,* 375–389.

*Zheng, Y., Myerson, J., & Hale, S. (2000). Age and individual differences in visuospatial processing speed: Testing the magnification hypothesis. *Psychonomic Bulletin & Review, 7*(1), 113–120.

Received September 29 2006
Revision received July 14, 2007
Accepted August 30, 2007 ■

Appendix

Table A1

A GOMS Task Analysis for Phone A in the Dial-a-Number Task

Task analysis pathway	Operator	Young parameter time (ms)	Old parameter time (ms)	Net time (ms) young	Net time (ms) old
Step 1: Fixate 1st chunk of numbers on paper (1st 3 numbers)	F	230	267	230	267
Step 2: Encode first 3 digits	3 C	3 (70)	3 (118)	440	621
Step 3: Fixate keypad	F	230	267	670	888
Step 4: Decode first chunk	C	70	118	740	1006
Step 5: Fixate first digit	F	230	267	970	1273
Step 6: Dial 1st digit	M	70	146	1040	1419
Step 7: Fixate second digit	F	230	267	1270	1686
Step 8: Dial 2nd digit	M	70	146	1424.8	1980
	Fitts	84.8	148		
Step 9: Fixate third digit	F	230	267	1654.8	2247
Step 10: Dial 3rd digit	M	70	146	1824.8	2568
	Fitts	100	175		
Step 11: Fixate 2nd chunk of numbers on paper (2nd 3 numbers)	F	230	267	2054.8	2835
Step 12: Encode second 3 digits	3 C	3 (70)	3 (118)	2264.8	3189
Step 13: Fixate keypad	F	230	267	2494.8	3456
Step 14: Decode second chunk	C	70	118	2564.8	3574
Step 15: Fixate first digit	F	230	267	2794.8	3841
Step 16: Dial 1st digit	M	70	146	2964.8	4162
	Fitts	100	175		
Step 17: Fixate second digit	F	230	267	3194.8	4429
Step 18: Dial 2nd digit	M	70	146	3264.8	4575
	Fitts	0	0		
Step 19: Fixate 3rd digit	F	230	267	3494.8	4842
Step 20: Dial 3rd digit	M	70	146	3697	5219
	Fitts	132.2	231		
Step 21: Fixate last chunk of numbers on paper (last 4 numbers)	F	230	267	3927	5486
Step 22: Encode last 4 digits	4 C	4 (70)	4 (118)	4207	5958
Step 23: Fixate keypad	F	230	267	4437	6225
Step 24: Decode last chunk	C	70	118	4507	6343
Step 25: Fixate first digit	F	230	267	4737	6610
Step 26: Dial 1st digit	M	70	146	4955.5	7016
	Fitts	148.5	260		
Step 27: Fixate second digit	F	230	267	5185.5	7283
Step 28: Dial 2nd digit	M	70	146	5281.3	7475
	Fitts	26.3	46		
Step 29: Fixate third digit	F	230	267	5511.8	7742
Step 30: Dial 3rd digit	M	70	146	5740.3	8165
	Fitts	158.5	277		
Step 31: Fixate fourth digit	F	230	267	5970.3	8432
Step 32: Dial 4th digit	M	70	146	6188.8	8838
	Fitts	148.5	260		
Step 33: Return with goal accomplished	C	70	118	6258.8	8956
Step 34: Fixate green send button	F	230	267	6488.8	9223
Step 35: Press green send button	M	70	146	6558.8	9369
Total time				6.59 sec	9.37 sec

Note. In the Operator column, F refers to an eye fixation, M refers to a motor processing cycle, Fitts refers to the amount of time added to the motor processing cycle to press a button with a given width and distance from the starting position, and C refers to a cognitive processing cycle.