INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION, *14*(3&4), 423–446

Effect of Auditory Waiting Cues on Time Estimation in Speech Recognition Telephony Applications

Melanie D. Polkosky
James R. Lewis
IBM Voice Systems

Previous empirical research in subjective time estimation and applied work in auditory interface design imply that designers can use auditory stimuli during system processing to manipulate users' perception of its duration. Two experiments investigate the effect of system response time (SRT) duration and rate of change of an auditory waiting cue on participants' subjective time estimates and perceived affect. The results showed that perceived SRT duration and ratings of perceived anxiety, stress, and impatience increased as ticking rate increased. However, with a slow rate (2-sec ticking), participants underestimated the duration of SRT, but indicated a significant increase in negative affect as compared with silent conditions. These results suggest that interface designers may reduce the subjective duration and negative affective states of SRT through carefully chosen, slow tempo system processing tones. The results of this research also stress the importance of thoughtful, informed interface design that makes contact with the empirical literature of the cognitive sciences.

1. INTRODUCTION

The empirical literature has demonstrated that system response time (SRT) is a component of human–computer interfaces that can dramatically affect user acceptance of an application (Shneiderman, 1984). SRT is the time required for a computer to receive a user's input, process the response, and send a reply back to the user (Thadhani, 1981). During SRT, the user waits for system processing to finish while monitoring the system's task.

The bulk of research on SRT dating from the 1960s (Nickerson, 1969) addresses the response delays of mainframe and desktop computers (Jacko, Sears, & Borella, 2000). More recently, as newer technologies have emerged, studies investigating system processing delays associated with the World Wide Web (Jacko, Sears, & Borella, 2000; Ramsay, Barbesi, & Preece, 1998), networks (Roast, 1998), and virtual reality applications (Watson, Walker, Ribarsky, & Spaulding, 1998) continue to

Requests for reprints should sent to: Melanie D. Polkosky, 5081 Congress Avenue, Suite 2207, Boca Raton, FL 33487. E-mail: polkosky@us.ibm.com

demonstrate that extended SRTs disrupt users' experiences with myriad applications. Limiting the duration of SRT is no less relevant for user acceptance of interactive voice response interfaces.

The use of interactive voice response (IVR) interfaces, especially in telecommunication applications, has grown significantly over the past decade (Spiegal & Streeter, 1997). Although these interfaces allow the user to interact in the more familiar and natural mode of spoken communication, the problem of system processing delay continues to be a matter of debate. Kamm and Helander (1997) contended that

> with continuing advances in processor speed and in the efficiency of recognition search and language processing algorithms, near real-time system response is becoming feasible even for complex speech understanding tasks, so SRT may cease to be a significant interface issue. (p. 1048)

L. Miller and Thomas (1977) made a similarly optimistic prediction nearly 30 years ago, suggesting that technological advances would make SRT concerns obsolete for desktop applications. However, the mountainous volume of literature on SRT and its effect on users is a clear reminder that SRT is problematic for desktop computer users even today. Researchers and designers of the most recent technologies recognize SRT as a by-product of the interaction among complex processing tasks such as speech recognition and multimedia retrieval, hardware capabilities, and the ever-increasing demands of multiple simultaneous users (Balentine & Morgan, 1999; Jacko et al., 2000). If hindsight is any guide, SRT will continue to be a significant and vexing issue for interface designers, even as IVR systems mature.

Telephony applications are particularly susceptible to SRT delays. They employ verbal interaction (by both the application and user) via a telephone for information retrieval and transaction services, including "remote banking, travel reservations, information inquiry, stock, mutual fund, and other financial transactions, international calling, [and] credit card verification" (Balentine & Morgan, 1999, p. 1). In a typical telephony interaction, a remote server hosts the application used for the transaction task, and the SRT delay includes a series of complex processing tasks: recognizing the users' spoken request, connecting to the server, sending data over wireless or other networks, retrieving the requested information, returning data through the network, and generating a synthetic spoken message that is finally presented to the user. This delay may be prolonged even more with heavy network traffic, large data files, and telephones, networks, or servers that are not state-of-the-art in processing capacity. Therefore, the simple user task of checking a bank balance over the telephone can combine all of the most complex and processor-intensive demands possible with the current limitations of technology.

Telephony applications also represent the most constrained and difficult IVR design environment. Visual cues and feedback to the user are significantly limited by the telephone itself, which was originally intended for conversation between humans but adapted to complex information retrieval tasks previously performed with benefit of a visual display. Increasing demand for ubiquitous computing suggests that small, portable telephones will continue to be the preferred design, precluding extensive visual cuing or feedback. Telephone portability also increases the likelihood that users will call for urgently needed information in a wide variety of noisy, dynamic, distracting, and stress-producing environments, diminishing their tolerance for waiting on the phone. For users, speech is a natural, "intelligent" interface, and the telephone is a familiar device. This powerful combination of technological sophistication and familiarity is likely to further elevate users' expectations and make them even less tolerant of long delays. Balentine and Morgan (1999) confirmed that minimal design flaws in other applications become "insurmountable" with "no margin of error" in telephony interfaces (p. 2). Finally, because the technology is emerging, there are few guidelines and little applied research to guide the user-centered development process.

Given that system response delays will be a reality of telephony applications into the foreseeable future, the interface design community must continue to find methods of managing this aspect of the user experience. In two studies, we begin to consider how designers can alter user perception of SRT duration using auditory processing tones in speech recognition telephony applications. We review the literature on SRT effects on users, subjective time estimation, and auditory interface design as a basis for functional interventions during SRT in the unique and emerging context of telephone-based interaction.

1.1. SRT Effects on Users

SRT duration affects the user in a variety of ways, although the previous literature presents an inconsistent picture about the relation between SRT and individual outcomes. Most studies have shown these effects are a result of SRT magnitude, with long SRTs causing the most dramatic user consequences. As SRT magnitude increases, user response time also increases (Barber & Lucas, 1983; Butler, 1984; Thadhani, 1981). However, other studies dispute this finding (Dannenbring, 1983; Goodman & Spence, 1982; Kuhmann, Boucsein, Schaefer, & Alexander, 1987). Not only the magnitude of SRT is problematic for users: Highly variable SRT durations are disruptive because they prevent the user from effectively dividing attention between SRT monitoring and a competing task (Galloway, 1981; L. Miller & Thomas, 1977; Murray & Abrahamson, 1983). Work quality and productivity also are affected by SRT, apparently decreasing as SRT duration increases (Dannenbring, 1983; Kuhmann et al., 1987), but the relation may not be a simple one (Barber & Lucas, 1983; Martin & Corl, 1986; Weiss, Boggs, Lehto, Shodja, & Martin, 1982), nor occur for all tasks (Butler, 1984). Kohlisch and Kuhmann (1997) showed that short SRTs result in poor performance and increased cardiovascular activity in users, which the researchers suggested was an indication of inadequate task readiness. However, they found that long SRTs created boredom (Kohlisch & Kuhmann, 1997). SRTs can also induce stress or other negative emotional states in users, particularly frustration and irritation (Guynes, 1988; Schaefer, 1990; Schleifer & Amick, 1989). Two studies suggest that SRTs also produce anxiety (A. Eisler & Eisler, 1994; Guynes, 1988). Finally, several studies provide evidence of somatic complaints and physiological changes that occur due to SRT (Kolisch & Kuhmann, 1997; Kuhmann et al., 1987; Thum, Boucsein, Kuhmann, & Ray, 1995).

In addition to negative personal outcomes, it seems reasonable that SRTs can impact users' overall perception of an interface. Indeed, the literature has shown that user acceptability varies based on a number of factors beyond the magnitude and variability of SRT (Shneiderman, 1984). Galloway (1981) suggested that several factors influence user acceptance, including whether (a) repeated SRTs interrupt the pace of work (by requiring users to switch attention between the computer and a primary work task), (b) SRTs occur at major breaks in work, (c) information must be retained in memory throughout the SRT, or (d) the user perceives the SRT as appropriate to the task performed by the system. Studies of decreasing user satisfaction and acceptability with SRT (Barber & Lucas, 1983) have led some researchers to recommend maximum SRT for specific applications (Johansson & Aronsson, 1984). More recently, Jacko et al. (2000) found a statistically significant interaction between network delay (short, medium, and long) and document type (text only or text and graphics) affected perceived usability of internet Web sites, as measured by perceived quality of information at the site, information organization, and likelihood of recommending the site to others. They concisely interpreted their findings as showing

> a trade off exists between the type of media used and the delays users experience. When the delays are short enough, users prefer documents that include graphics. However, as delays increase, graphics are viewed as contributing to the delay and simpler text-only documents are preferred. (p. 438)

The research suggests that SRT causes both broad and primarily negative user outcomes, often interacting with user and other system variables. These negative user outcomes may also influence the user's impression of the interface as a whole. To avoid these consequences, it is essential that application designers implement interfaces that effectively minimize users' perception of SRT duration.

1.2 User Perception of Time

The acceptability of SRT, and possibly an entire application, relates at least partially to a user's perception of time. A significant cognitive-psychological literature exists on the topic of subjective time estimation and perception (for a review, see Fraisse, 1984). This literature indicates that the subjective experience of time is a power function of actual time duration multiplied by a constant (subjective time estimate = (actual time)$^{0.9}$*C); H. Eisler, 1976). If, as H. Eisler suggested, the exponent of the subjective time function is 0.9, users' perception of SRT increases approximately linearly as actual SRT duration increases (Meyer, Shinar, & Leiser, 1990). However, a number of independent variables can affect the function's exponent, as argued by H. Eisler.

Researchers have suggested that several individual variables influence subjective time estimation, including age (Block, Zakay, & Hancock, 1998; Craik & Hay, 1999), intelligence (Fink & Neubauer, 2001), personality (Carrasco, Guillem, & Redolat, 2000; Rammsayer & Rammstedt, 2000; Zakay, Lomranz, & Kaziniz, 1984), gender (A. Eisler & Eisler, 1994), attention (Zakay, 1989), and memory (Fraisse, 1984). More recently, studies of individuals with various speech or cognitive impairments provide additional support for the influence of underlying cognitive mechanisms on subjective time estimation (Ezrati-Vinacour & Levin, 2001;

Hellstrom & Almkvist, 1997; Hellstrom, Lang, Portin, & Rinne, 1997; Lange, Tucha, Steup, Gsell, & Naumann, 1995; Nichelli, Venneri, Molinari, Tavani, & Grafman, 1993; Riesen & Schnider, 2001). Researchers implicate memory (Ornstein, 1969; Poynter & Homa, 1983) or attention (Zakay, 1989) as the central cognitive process of subjective time estimation, although empirical support for specific theoretical models has been inconsistent and controversial (Meyer, Shinar, Bitan, & Leiser, 1996; Zakay, 1993).

Perhaps of more direct relevance to application designers, research demonstrates that external variables, especially sensory stimuli, can alter the subjective experience of time. Zakay, Nitzan, and Glicksohn (1983) examined the effect of sensory stimulus and task difficulty on subjective time estimation. Ninety-six participants estimated the length of empty intervals or intervals filled with a fast or slow tempo (manipulated with a flickering light bulb or electronic buzzer). They found that a fast tempo resulted in the longest time estimates, whereas a slow tempo resulted in the shortest time estimates (no tempo produced intermediate time estimates). Yoblick and Salvendy (1970) found that when participants reproduced filled time intervals (auditory tones, visual flicker, or tactile vibrations), they overestimated the duration of lower frequencies significantly more often than higher frequencies only with auditory stimuli. When the time intervals were filled with visual or tactile stimuli, participants estimated high and low frequencies similarly.[1] Glicksohn (1992) studied the effect of an altered sensory environment on subjective time estimation in a 3 × 4 × 4 factorial design using 96 participants (8 participants per cell). He exposed the participants to a combination of visual stimulation (visual overload, visual deprivation, or reading with normal room lighting) and auditory stimulation (dichotic music presentation, different music presented to each ear simultaneously; stereophonic music presentation, the same music presented to both ears simultaneously; white noise; or no auditory stimulation) for 20 min prior to estimating intervals of 4, 8, 16, and 32 sec. Participants' estimates without prior sensory stimulation were used as covariates. Results indicated only a significant interaction for four conditions involving visual deprivation plus dichotic listening or white noise and visual overload plus dichotic listening or white noise. Therefore, dichotic music–visual overload and visual deprivation–white noise led to inflated time estimates but visual deprivation–dichotic music and visual overload–white noise led to depressed time estimates.

Although these laboratory studies suggest that the auditory environment can influence subjective time estimation, the results are only indirectly applicable to interface design. Zakay et al. (1983) included task difficulty as an independent variable, adding a laboratory-based, contrived task that would not be typical of users of telephone-based voice interfaces. In addition, the experimental designs only hint at the idea that, by controlling the sensory environment, interface designers can alter a users' perception of time. A stronger design in an applied setting is needed to validate this tantalizing proposal.

---

[1] The finding of no effect in this study may have been related to the difference in experimental method, also a topic of debate in this literature. Yoblick and Salvendy (1970) required participants to reproduce a time duration they considered equivalent to the experimental period, then measured the duration of that reproduction. Zakay, Nitzan, & Glicksohn (1983) required participants to recall the experimental duration and provide a verbal estimate of its duration.

1.3. Previous Research in Human–Computer Interaction (HCI)

Parallel research in HCI has explored similar issues as the cognitive psychological literature; however, the primary thrust in the applied setting has been to determine how SRT magnitude and duration affect users' perceptions of an interface. A second line of research has explored the design of visual displays that minimize the perceived duration of time (Block, 1990; Levin & Zakay, 1989; Meyer, Bitan, & Shinar, 1995). Few of these studies offer clues to adapting the findings to other sensory modalities, such as auditory presentation.

A notable exception is the work of Meyer, Shinar, and Leiser (1990), who examined the effect of wait messages on participants' estimates of 3- to 16-sec SRTs. The wait messages were static (blank screen, printed "please wait," six-word printed epigram) or dynamic (increasing line of printed X letters, round clock drawing, blinking printed "please wait"), with dynamic messages presented at three rates of change (changing every .33, .50, or .67 sec). Results showed no difference in time estimates for the three static displays and the blinking please wait message. The dynamic displays that changed over time (line of Xs and clock) resulted in longer time estimates when rates of change were faster. This result provides additional, applied support for the Zakay et al. (1983) finding that an external tempo influences subjective time estimation.

A third, related area of HCI research has addressed auditory waiting cues. These cues play during SRT and identify that system processing is occurring, while simultaneously informing the user that the system has not disconnected (Balentine & Morgan, 1999). Several researchers describe the use of waiting tones during SRT without determining their effect on user time perception or SRT acceptability. Beaudouin-Lafon and Conversy (1996) explained their use of Sheppard–Risset tones (sounds that appear to go up and down indefinitely) as audio progress bars. Albers and his colleagues (Albers, 1996; Albers & Bergman, 1995) used a ticking sound for relative transfer time in a World Wide Web browser and in a satellite–ground control application (Albers, 1995). Albers (1995) also used "pops and clicks" to indicate data transfer. Similarly, Balentine & Morgan (1999) advocated the use of "low level ticking" or "pitched wait tones" to indicate the user's need to wait for system processing in telephony applications.

Buxton (1989) identified an analogous monitoring function of auditory cues in interface design, which few other researchers have explored empirically. Rauterberg (1998) used six machine sounds to assist operators in monitoring a simulated plant. The results showed that auditory cues significantly improved the operators' productivity scores, number of status reports, self-assurance, and social acceptance as compared with a condition without these cues. This study provides some limited empirical evidence that auditory cues may have performance and psychological benefits when individuals are monitoring system processes.

1.4. Auditory SRT Cues in Telephony Applications

Telephone-based voice interaction presents a unique design environment. Although the past studies of SRT have used primarily visual applications (desktop computer), telephony applications demand an auditory cue because they do not typically provide a visual display (Schumacher, Hardzinski, & Schwartz, 1995). The use of auditory stimuli to signal SRT and its completion may have several additional cognitive-psychological advantages for users. Complex sounds draw attention, especially when they are changing; and designers can use them to shift listener attention (Moore, 1989). Gaver (1997) noted that sound is generally effective at conveying information about processes. If a system provides a processing tone for users, they may be able to divide attention to continue with an ongoing task and monitor the SRT, thereby limiting work interruption. Similarly, the end of the tone may shift the user's primary attention back to system interaction. The particular sound itself may also alter a listener's mood or emotional state (Gaver, 1997). Perhaps the appropriate use of sound in interface design can counteract or reduce the negative outcomes associated with SRT by the previous literature.

In two experiments, we expand on previously applied research in HCI and empirical work on subjective time estimation. Because previous studies involving SRT have primarily addressed visual wait signals (Meyer et al., 1996; Meyer et al., 1990), we investigated the use of three rates of auditory processing tone on users' subjective time estimation. Yoblick and Salvendy (1970) also investigated the effect of auditory tones on subjective time estimation, but their tones were unchanging over time (36 sinusoidal waves ranging from 80 Hz–14,000 Hz across 36 experimental conditions) and did not represent SRTs. Zakay et al. (1983) provided auditory stimuli using a buzzer that "flickered" with durations of 0.5 sec or 2.0 sec during 14-sec verbal tasks. Therefore, participants engaged in a competing task in this study (as opposed to simply waiting during SRT), and the length of the tasks remained consistent. They used a between-subject design (six participant groups of fast visual, fast auditory, slow visual, slow auditory, and control), which did not allow comparison of a single participant's time estimates with different rates of external stimuli. In addition, the researchers reported post hoc comparisons only between the mean for the three verbal tasks and a condition with no verbal task (three mean comparisons within slow external tempo), leaving the differences between the slow groups and control unclear.

In contrast, this study investigated three rates of dynamic auditory stimuli occurring during realistic SRT durations. The repeated measures design also allows us to compare participants' time estimates among conditions of the independent variables.

2. EXPERIMENT 1

Our initial question was deceptively simple: Can we manipulate user perception of SRT duration in speech recognition telephony applications using auditory stimuli?

Consistent with earlier findings (Meyer et al., 1996; Meyer et al., 1990; Zakay et al., 1983), we hypothesized that more rapid rates of auditory tones would result in greater overestimation of actual SRT durations. We used a ticking tone in this experiment, which is the most common processing tone identified in previous literature (Albers, 1995, 1996; Albers & Bergman, 1995; Balentine & Morgan, 1999). We also included a control condition of silence during the SRT, an "empty" time interval, as in Zakay et al. (1983).

In addition, we explored the effect of processing tones on users' negative affect. Because there is no empirical evidence that users prefer the ticking tone but the literature indicates SRT results in negative affect, we wanted to determine if the ticking rates affected user anxiety, stress, and impatience. We hypothesized that processing tones with faster rates of change would increase users' perceived negative affect.

Finally, consistent with previous findings of gender and age effects (Block et al., 1998; Craik & Hay, 1999; A. Eisler & Eisler, 1994), we included both gender and age as independent variables in this study.

2.1. Method

Participants. Sixteen IBM employees volunteered to complete this study. The participant sample included equal numbers of men and women, with equal numbers of each gender group above and below the age of 40. All participants, except 1 man and 1 women (each over 40 years), described themselves as experienced with speech recognition telephony applications. All participants reported normal hearing.

Stimuli. Participants heard three waiting tones and silence, counterbalanced across participants to reduce order effects. The waiting tones consisted of a ticking sound, edited so the rate of ticking doubled with each successive tone (a tick every 0.5, 0.25, and 0.125 msec, respectively). Each tone and silence played during actual SRT durations of 3 sec, 8 sec, 13 sec, and 18 sec, creating 16 conditions of the independent variables (four auditory stimuli and four SRT durations).

The SRT durations were consistent with the range of times used in the previous literature (Galloway, 1981; Guynes, 1988; Kuhmann et al., 1987; Meyer et al., 1990). In addition, Fraisse (1984) suggested that duration may influence depth of cognitive processing. He suggested that people perceive durations of 100 msec to 5 sec as being in the present, but involve memory for those over 5 sec in duration.

To simulate use of the tones in a speech recognition telephony application, the auditory stimuli occurred between two spoken prompts. The initial prompt was a statement announcing the computer's initiation of processing (e.g., "Please hold while we process your request"), and the second prompt indicated the end of the system processing time (e.g., "Thank you for waiting"). Both prompts were spoken by a woman and recorded (16 bit; 44,100 Hz) using Sound Forge 4.5d™ (Sonic Foundry Inc.), then edited to include the auditory stimuli and SRT durations.

Procedure. The study used a digram-balanced Latin square design to prevent participants from hearing the same tones and SRT durations sequentially. This scheme not only results in standard Latin square counterbalancing of order of appearance in rows and columns of the design, but also controls immediate sequential effects (Bradley, 1958; Lewis, 1993).

Each participant read a brief description of the task and four questions eliciting time estimates and ratings of perceived anxiety, stress, and impatience on three bipolar 7-point rating scales. Each participant received verbal clarification and additional explanation as needed. The study used a prospective paradigm (also used in previous research) in which participants knew that they would be estimating time intervals but they were not permitted to use a watch or other timing device. They listened to a prompt–auditory stimulus combination played over an Andrea CTI ANC-200™ (Andrea Electronics) handset attached to an IBM ThinkPad® (IBM Corp.) computer, which simulated actual listening conditions in a telephony–speech recognition application. Participants then completed the four questionnaire items: (a) "How long was the waiting period (in seconds)," (b) "How anxious did you feel during the waiting period," (c) "How impatient did you feel during the waiting period," and (d) "How stressed did you feel during the waiting period?" Participants repeated this procedure for the remaining audio stimuli and SRT durations.

2.2. Results and Discussion

Subjective time estimation. A 2 × 2 × 4 × 4 mixed model analysis of variance (ANOVA) with two within-subjects variables (auditory stimulus, SRT duration) and two between-subject variables (gender, age) indicated a main effect of auditory stimulus, $F(3, 36) = 4.58$, $MSE = 24.42$, $p = .008$; and SRT duration, $F(3, 36) = 77.70$, $MSE = 39.78$, $p < .0001$. Two interactions, auditory stimulus–SRT duration, $F(9, 108) = 1.97$, $MSE = 6.85$, $p = .05$; and stimulus–duration–age–gender, $F(9, 108) = 2.51$, $MSE = 6.85$, $p = .012$, were also statistically significant. No other main effects and interactions were significant ($p > .17$).

The effect of primary interest, that of auditory stimulus on subjective time estimation, indicated that as the rate of ticking increased, participants' estimates of the mean actual SRT (10.5 sec) also increased (see Table 1). Participants underestimated the mean actual SRT only in the silence condition. Post hoc t tests on the mean difference scores indicated significantly higher time estimates occurred for the 0.25 ticking condition compared to silence, $t(15) = -2.98$, $p = .009$; and the 0.125 ticking condition compared with silence, $t(15) = -3.17$, $p = .011$. All other mean comparisons failed to be significant ($p > .07$, using the Bonferroni correction with $\alpha = 0.016$). This result confirmed our initial hypothesis that individuals would overestimate SRT when they heard a rapid rate of ticking. However, the lack of significance

Table 1: Subjective SRT Estimates for Each Actual SRT Duration and Auditory Stimulus (Experiment 1)

Auditory Stimulus	*Mean Estimate of Mean SRT (Sec)* [a]	*SD (Sec)*	*Actual Length of SRT*	*Mean Subjective Estimate of SRT (Sec)*	*SD (Sec)*
Silence	9.92	5.43	3 sec	3.75	1.78
0.5-sec ticking	11.45	9.64	8 sec	8.67	3.65
0.25-sec ticking	12.02	5.44	13 sec	14.20	6.49
0.125-sec ticking	13.09	8.89	18 sec	19.84	9.11

Note. SRT = system response time.
[a]The mean actual SRT duration was 10.5 sec (mean of 3, 8, 13, and 18 sec).

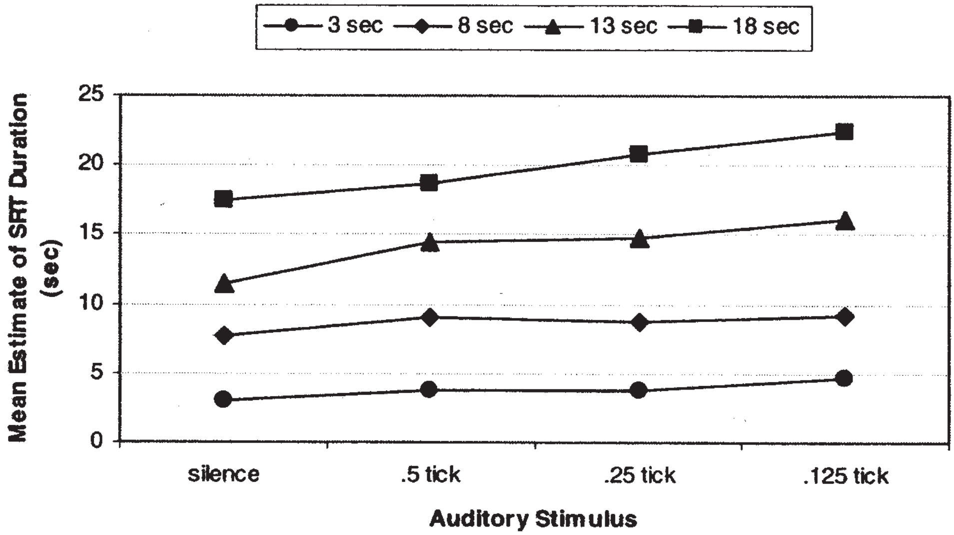

FIGURE 1 Auditory stimulus—system response time (SRT) duration interaction.

among paired comparisons of the three ticking conditions suggests that the time estimates did not necessarily increase with the ticking rate but were overestimates only as compared with the silence condition.

The interaction between SRT duration and auditory stimulus appears in Figure 1. As shown, the estimated duration increased slightly across the four auditory stimulus conditions for 3, 8, 13, and 18 sec SRTs. However, post hoc *t* tests indicated that participants significantly overestimated the 3-sec SRT when they heard 0.125-sec ticking as compared with silence, $t(15) = -3.62$, $p = .003$. Relative to the silence condition, the overestimate of the 13-sec SRT with 0.125-sec ticking approached statistical significance with the Bonferroni correction, $t(15) = -3.198$, $p = .006$. All other within-SRT duration comparisons failed to be significant ($p > .01$, using the Bonferroni correction with $\alpha = 0.004$).

Additional effects of the independent variables, although of interest in the context of previous literature, were not specified in our hypotheses. An expected effect of SRT duration indicated that as the actual duration of the SRT increased, participants' estimates of the duration also increased (see Table 1). Post hoc *t* tests on the difference scores revealed all mean differences were highly significant ($p < .00001$).

Analysis of the time estimate data also indicated a significant four-way interaction (time–rate–age–gender), shown in Figure 2. Women and men over 40 years old similarly estimated SRTs, regardless of the auditory stimulus. However, men under 40 years underestimated the SRTs in all conditions, except when the actual duration of the SRT was 3 sec. Although this interaction is of some theoretical significance and provides additional support for gender and age effects in subjective time estimation (Block et al., 1998; Craik & Hay, 1999; A. Eisler & Eisler, 1994), designers must select a "best" interface for use in a single application. Because this effect has little practical applicability for general design principles for a broad user population, we did not analyze the interaction in more detail.

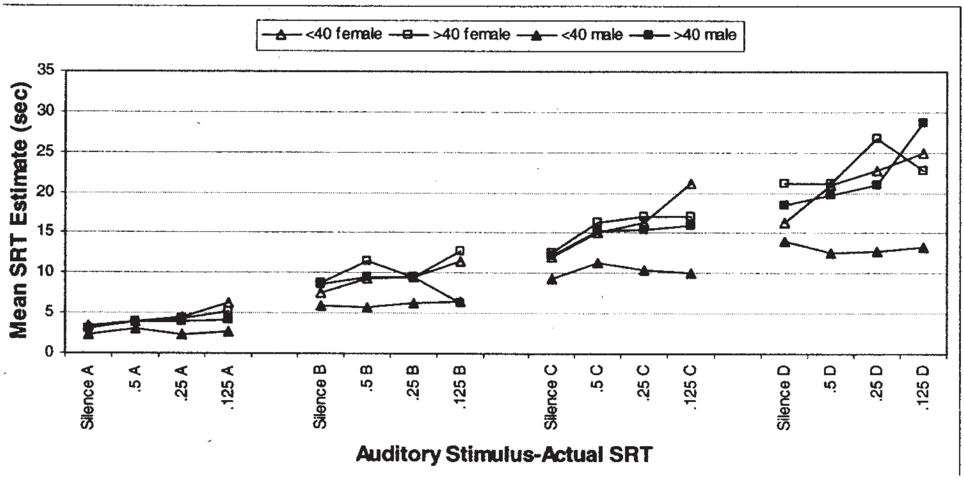

FIGURE 2 Auditory stimulus—actual system response time (SRT) interaction by gender and age (A = 3 sec, B = 8 sec, C = 13 sec, D = 18 sec).

In general, the main effect of auditory stimulus confirmed our initial hypothesis that individuals overestimate SRT when they hear more rapid rates of ticking. Although the slowest rate (0.5-sec ticking) was similar to silence, the two faster rates did result in significant overestimates of SRT duration as compared to silence. The data also suggest that some auditory stimuli rates may interact with the duration of SRT, although these effects did not appear to be systematic.

Perceived negative affect. A mixed model ANOVA demonstrated a main effect of SRT duration, $F(3, 36) = 27.69$, $MSE = 3.02$, $p < .0001$; auditory stimulus, $F(3, 36) = 11.94$, $MSE = 9.55$, $p < .0001$; and affective rating, $F(2, 24) = 5.07$, $MSE = 2.21$, $p = .015$. A significant interaction occurred between SRT duration and affective rating, $F(6, 72) = 11.54$, $MSE = 0.41$, $p < .0001$. No other main effects or interactions were significant ($p > .08$).

The main effects indicated that participants' negative affect (a combined rating consisting of anxiety, stress, and impatience) increased as the SRT duration increased and as the rate of auditory stimuli increased. Figure 3 shows higher negative affect ratings with increased rates of auditory stimuli. Post hoc t tests indicated significantly higher ratings of negative affect among all paired conditions ($p < .01$), except silence and 0.5-sec ticking. Therefore, 0.5-sec ticking resulted in similar negative affect as silence during SRTs, but more rapid rates of ticking increased participants' perceived negative affect. This provides empirical support for our initial hypothesis that, although SRT itself results in negative affect, the auditory stimulus provided during the SRT can also produce or even increase negative affect.

In general, Experiment 1 provided evidence that we did manipulate participant time perception using auditory stimuli. Unfortunately, these manipulations did not decrease user estimates of SRT durations or negative affect associated with SRT. This

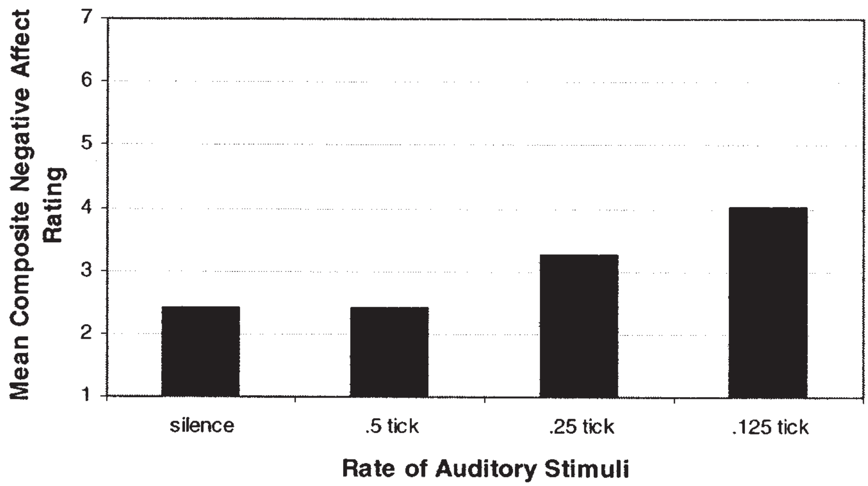

FIGURE 3 Perceived negative affect ratings by auditory stimuli.

fascinating insight into the power of auditory interface design led us to question how we might use this knowledge to promote more constructive user outcomes.

3. EXPERIMENT 2

In Experiment 2, we wanted to expand on our Experiment 1 results and increase the power of the initial experiment through replication. Because faster rates of ticking led to overestimates of SRT, we hypothesized that slower rates of ticking would result in underestimation of SRT. There is limited empirical support for this hypothesis: Zakay et al. (1983) provided auditory stimuli using a buzzer that flickered at rates of 0.5 sec or 2.0 sec during 14-sec verbal tasks. They found a significant main effect of tempo rate on perceived duration of the buzzer tone, indicating that the slower rate did produce lower subjective time estimates. A similar result occurred in other studies using rhythmic visual stimulation (Planas & Treurniet, 1988), auditory presentation of words at rates of one every 6 or 3 sec (Block, 1974), and rhythmic auditory stimulation using a metronome (Jones & Natale, 1973). Therefore, the purpose of Experiment 2 was to determine if even slower rates of ticking (slower than 0.5 sec per tick) cause users to underestimate the true SRT. We again wanted to determine whether the rate of ticking also influenced listeners' perceived negative affect and whether gender and age effects occurred. Such a finding would be important because the primary goal of this line of research is to discover techniques for reducing the subjective duration of SRT.

3.1. Method

This experiment used the same design, procedure, and participant characteristics as the previous study. However, the auditory stimuli included three rates of ticking in which the rate was halved for successive tones (a tick every 0.5, 1, and 2 sec, respectively) and a control condition of silence. Including the silence and 0.5-sec ticking conditions in the second experiment provided an opportunity to partially replicate the first study and determine if the two experiments yielded comparable results. All other methodological details were identical to Experiment 1.

3.2. Results and Discussion

Subjective time estimation. We analyzed the data using a mixed model ANOVA, as in Experiment 1. The ANOVA indicated a main effect of auditory stimulus, $F(3, 36) = 4.02$, $MSE = 113.80$, $p = .015$);and the expected main effect of SRT duration, $F(3, 36) = 210.39$, $MSE = 1878.24$, $p < .0001$. Two significant interactions with SRT duration also occurred: duration–gender, $F(3, 36) = 4.53$, $MSE = 40.42$, $p = .009$; and duration–age–gender, $F(3, 36) = 3.11$, $MSE = 27.77$, $p = .038$. The main effect of gender, $F(1, 12) = 3.86$, $MSE = 344.73$, $p = .073$); and interaction between age and gender, $F(1, 12) = 4.18$, $MSE = 372.80$, $p = .064$, were marginally significant. All other main effects and interactions were not significant ($p > .51$).

Table 2 presents the main effect of auditory stimulus. Participants most accurately estimated the mean actual SRT when they heard the 0.5-sec ticking but underestimated in the other three conditions. Post hoc t tests indicated statistically significant differences between the following mean pairs: silence–0.5 second ticking, $t(15) = -3.24$, $p = .005$; and 0.5-sec ticking–2-sec ticking, $t(15) = 0.62$, $p = .015$. Other mean comparisons were not significant ($p > .09$, using the Bonferroni correction with $\alpha = 0.016$). This result, a stronger result than observed in the previous study (due to the lack of interaction), demonstrates more significant estimate differences among the ticking tones, as well as in comparisons between the ticking and silence conditions. Our hypothesis that slower rates of ticking would result in underestimates of time duration was confirmed.

The remaining effects are interesting in the context of the previous literature, although their discovery did not drive this study. The main effect of SRT duration appears in Table 2. As shown, participants estimated longer SRTs when the actual duration was longer, with the shortest mean estimate for the 3-sec SRT and longest mean estimate for the 18-sec SRT. Participants showed increasing variability in their time estimates as the SRT became longer. Post hoc t tests on the difference

Table 2: Subjective SRT Estimates for Each Actual SRT Duration and Auditory Stimulus (Experiment 2)

Auditory Stimulus	*Mean Estimate of Mean Actual SRT (Sec)* [a]	*SD (Sec)*	*Actual Length of SRT*	*Mean Subjective Estimate of SRT (Sec)*	*SD (Sec)*
Silence	8.89	5.63	3 sec	3.01	1.46
0.5-sec ticking	10.42	6.12	8 sec	7.42	2.88
1-sec ticking	9.52	6.08	13 sec	11.45	4.36
2-sec ticking	8.72	7.58	18 sec	15.66	5.49

Note. SRT = system response time.

[a] The mean actual SRT duration was 10.5 seconds (mean of 3, 8, 13, and 18 seconds).

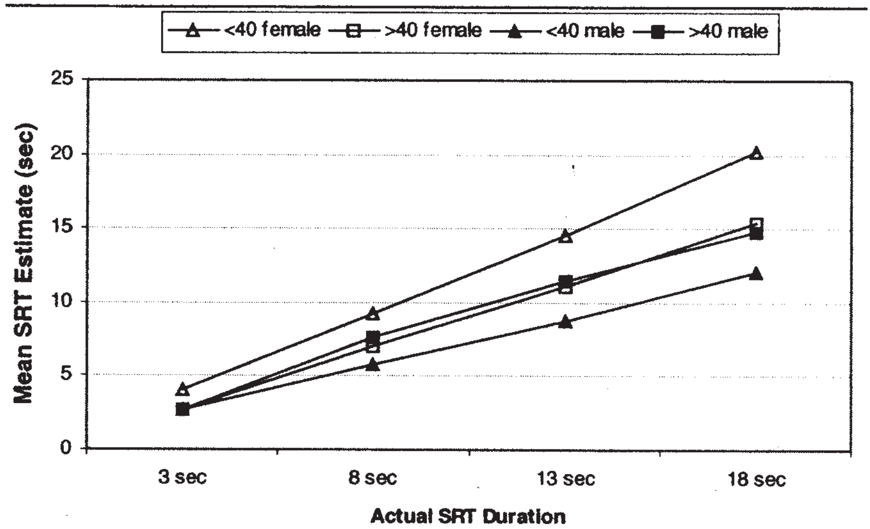

FIGURE 4 System response time (SRT) duration estimates by gender and age group.

scores again indicated statistically significant differences among estimates for all four SRT durations ($p < .000001$).

Analysis of the time estimate data also indicated two significant interactions with SRT duration. The duration–age–gender interactions are shown in Figure 4. In general, women under 40 years of age provided the longest estimate of each SRT compared to the other three participant groups. Men under 40 years of age estimated the shortest SRT.

Negative affect. A mixed model ANOVA demonstrated a main effect of auditory stimulus, $F(3, 36) = 4.11$, $MSE = 6.56$, $p = .013$; SRT duration, $F(3, 36) = 21.39$, $MSE = 3.85$, $p < .0001$; and affective rating, $F(2, 24) = 17.60$, $MSE = 0.73$, $p < .0001$. Significant interactions occurred between auditory stimulus and affective rating, $F(6, 72) = 2.31$, $MSE = 0.18$, $p = .042$; auditory stimulus and SRT duration, $F(9, 108) = 2.02$, $MSE = 1.39$, $p = .044$; and SRT duration and affective rating, $F(6, 72) = 8.72$, $MSE = 0.23$, $p < .0001$; as well as a three-way stimulus–rating–age–interaction, $F(6, 72) = 2.72$, $MSE = 0.18$, $p = .019$. The four-way interaction of duration–affect–age–gender was marginally significant, $F(6, 72) = 2.08$, $MSE = 0.23$, $p = .066$. All other main effects and interactions were not statistically significant ($p > .095$).

The effects of interest indicated that participants' negative affect (a combined rating consisting of anxiety, stress, and impatience) decreased as the rate of ticking decreased. Figure 5 shows that participants rated higher negative affect when they heard the fastest ticking rate (0.5-sec ticking) and rated lower negative affect with slower ticking rates (1- and 2-sec ticking). However, their negative affect increased slightly as the ticking rate became very slow (2-sec ticking). Post hoc t tests indicated marginally significant differences between the silence and ticking conditions: si-

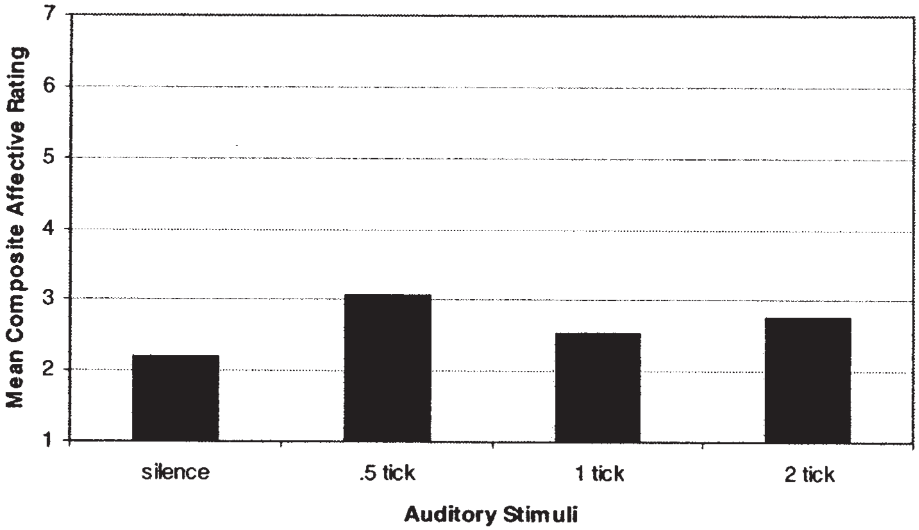

FIGURE 5 Main effect of auditory stimulus.

lence–0.5-sec ticking, $t(15) = -2.83, p = .013$; and silence–2-sec ticking, $t(15) = -2.69, p = .017$; but other mean comparisons were not significant ($p > .04$, using the Bonferroni correction with $\alpha = 0.01$). In general, 1-sec ticking was rated as similar to silence, but both 0.5- and 2-sec ticking resulted in increased perception of negative affect.

Figure 6 illustrates the interaction between auditory stimulus and affective rating. Perceived anxiety, stress, and impatience all decreased as the ticking rate became slower but were slightly higher with the slowest ticking rate (2-sec rate). Accordingly, post hoc t tests indicated marginally significant mean differences between the perceived stress during silence and 0.5-sec ticking, $t(15) = -3.28, p = .005$; and between the perceived impatience during silence and 2-sec ticking, $t(15) = -3.13, p = .007$. All other within-affective category mean differences were not significant ($p > .01$, using the Bonferroni correction with $\alpha = 0.004$).

3.3. Comparison of Experiments 1 and 2

A mixed model ANOVA revealed a nonsignificant main effect of study on subjective time estimates, $F(1, 24) = 0.94, MSE = 72.07, p = .343$; this indicates that participants in both studies provided similar estimates of the silence and 0.5-sec ticking rate conditions. Significant main effects occurred for gender, $F(1, 24) = 4.39, MSE = 72.07, p = .047$; auditory stimulus, $F(1, 24) = 9.27, MSE = 16.17, p = .006$; and SRT duration, $F(3, 72) = 216.38, MSE = 10.06, p < .0001$. Marginally significant interactions occurred for duration–gender, $F(3, 72) = 2.41, MSE = 10.06, p = .074$; and auditory stimulus–SRT duration–study, $F(3, 72) = 2.51, MSE = 5.83, p = .066$. In general, these results indicated that participants estimated the silence and 0.5 conditions similarly in both studies.

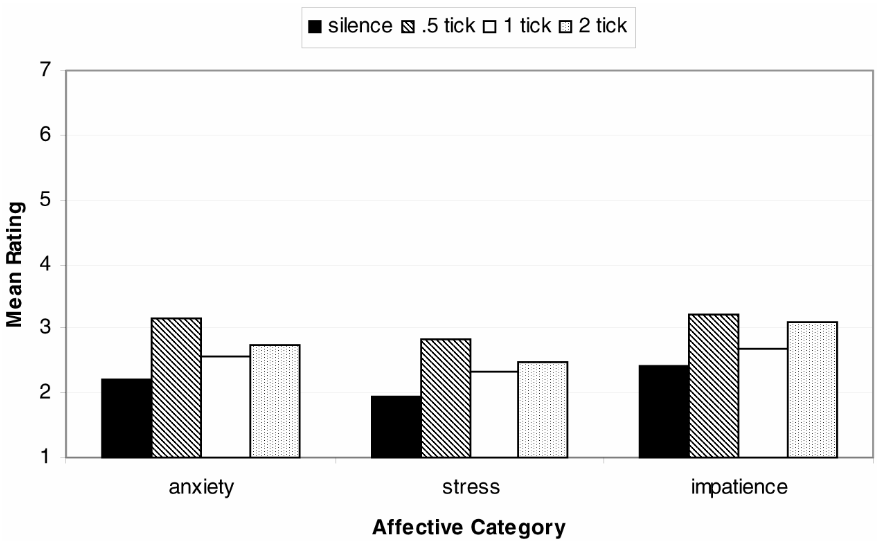

FIGURE 6 Interaction between auditory stimulus and affective rating.

In addition to the analysis of the raw data in both studies, we also analyzed difference scores (actual–perceived) and ratio (actual–perceived) data in both studies. The analyses were virtually identical and are therefore not reported in detail. However, the ratio data help to illuminate how time estimates varied overall.

As shown in Figure 7, participants underestimated SRT durations in the slow ticking rate conditions (1- and 2-sec ticking) and silence but overestimated SRT with the more rapid rates (0.5-, 0.25-, 0.125-sec ticking). In the figure, a ratio less than 1.00 indicates an underestimate of the actual SRT, and a ratio greater than 1.00 indicates an overestimate of the actual SRT. Because the main effect of study was nonsignificant, Figure 7 presents the means of the 0.5-sec ticking and silence conditions calculated from both experiments (32 participants).

We also completed a final analysis on the ratio estimates. As noted in the previous analyses, significant differences occurred between 0.25-sec ticking and silence, 0.125-sec ticking and silence, 0.5-sec ticking and silence, and 0.5-sec ticking and 2-sec ticking conditions. A set of four independent t tests among the unreplicated ticking rate conditions (using the Bonferroni correction with $\alpha = 0.01$) indicated that time estimates with the slow ticking rates were significantly shorter than time estimates with the fast ticking rates:

- 1-sec ticking SRT estimate less than 0.25-sec ticking SRT estimate, $t(126) = 3.31$, $p = .001$.
- 1-sec ticking SRT estimate less than 0.125-sec ticking SRT estimate, $t(126) = 3.78$, $p < .0001$.
- 2-sec ticking SRT estimate less than 0.25-sec ticking SRT estimate, $t(126) = 4.51$, $p < .0001$.
- 2-sec ticking SRT estimate less than 0.125-sec ticking SRT estimate, $t(126) = 4.65$, $p < .0001$.

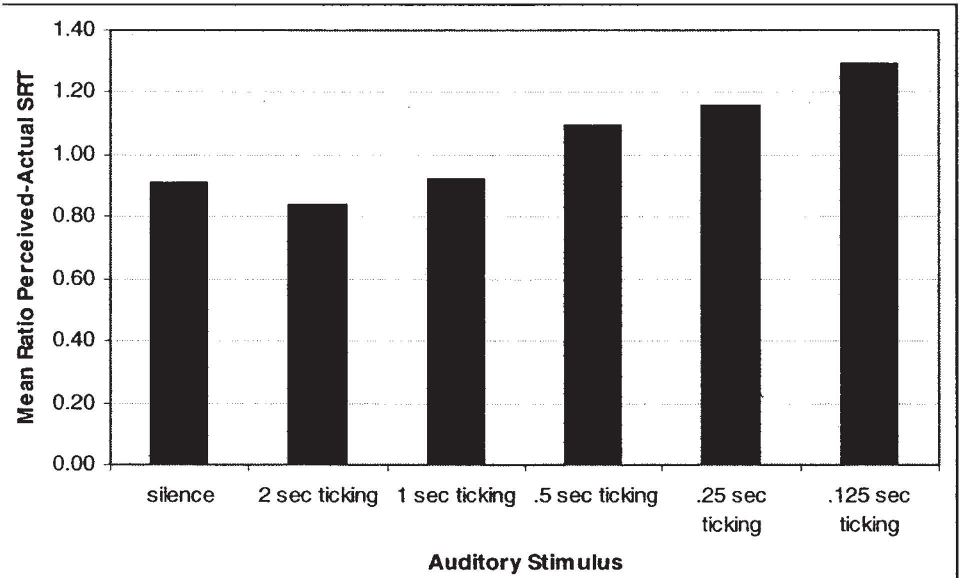

FIGURE 7 Ratio of perceived–actual system response time (SRT; Experiments 1 and 2).

In summary, this final analysis sharpens our findings in Experiments 1 and 2: Not only did participants estimate longer SRTs with the fast rates than with silent SRTs, they also estimated longer SRT durations as the ticking rate increased.

4. GENERAL DISCUSSION

In two experiments, we investigated the effect of SRT duration and rate of change of a system processing tone (ticking rate) on participants' subjective time estimates and perceived affect. In general, the results indicated that participants' SRT estimates increased with ticking rate, and they overestimated SRTs with fast ticking rates relative to silent SRTs. As the ticking rate increased, participants' ratings of perceived anxiety, stress, and impatience also increased. However, with a slow rate of ticking (2-sec ticking), participants underestimated the duration of SRT but indicated a significant increase in negative affect as compared with silent SRTs. The results confirmed our initial hypotheses that more rapid processing tones would increase both subjective SRT estimates and perceived negative affect. Our hypotheses that slower processing tones would result in underestimates of SRT was also confirmed for the 2-sec ticking rate.

4.1. Critical Evaluation

These experiments improve on previous research because they are the first SRT studies that investigate the effect of an auditory stimulus, as opposed to a visual display, on user's perception of SRT duration. These studies are also the first to address the unique design environment of speech recognition telephony applications. Finally, because time estimation was a within-subjects variable, we can draw conclusions about individuals' estimates under several auditory stimulus conditions: This sensitive, economical, and powerful design has had limited use in the previous literature.

There are also potential limitations in the design of the two current experiments. As Meyer et al. (1996) pointed out, a display that minimizes an apparent duration may not be the one that users prefer. In our studies, we did not determine whether users prefer the ticking tone. Preliminary investigation of this issue (Polkosky, 2001) suggests that users prefer jazz music and silence to a ticking tone (0.5-sec ticking) during system processing. Previous studies of music and waiting periods have demonstrated that music can influence emotional states during a waiting period (Chebat, Gelinas-Chebat, & Filiatrault, 1993; Hui, Dube, & Chebat, 1997), and individuals wait longer when they hear music instead of silence (Hargreaves, 1999). Ramos (1993) found that the type of music is an important design consideration, with jazz music producing the fewest lost calls to a state abuse hotline (followed by country, classical, popular, and relaxation music causing progressively more lost calls). However, the effect of music on time perception is not clear in this research. North, Hargreaves, and McKendrick (1999) found that two music conditions resulted in similar waiting time estimates as a "please hold" message spoken at 10-sec intervals. Chebat et al. demonstrated that musical tempo (fast vs. slow) did not have a direct influence on time perception while customers waited in bank lines. Instead, musical tempo had a complex, moderating relation to mood and attention, which in turn influenced time perception; the tempo rate also interacted with the amount of visual information, an interaction reminiscent of Glicksohn (1992). These studies provide some preliminary empirical support for music during SRT, but the results may not generalize the specific context of a telephony application. Continued research is needed to further investigate the effects of type and tempo of music on user preferences and time estimation in telephony applications.

Another possible limitation of our work centers on the generalizability of results. The participant group included volunteers who were all employees of IBM. However, there is little reason to expect that IBM employees would have unique perceptual abilities as compared to a broader population of adults. Indeed, a study of IBM employees and individuals hired from a temporary agency demonstrated no difference between these participant groups in their preferences for auditory tones (Polkosky & Lewis, 2001). Therefore, these results, especially when combined with previous empirical evidence with a variety of participant groups, should generalize to a broad population of male and female adults.

4.2. Design Implications

Prior to our studies, the most notable style guide for telephony applications offered a single "good practice" guideline related to user waiting times: "Provide auditory cues for wait times of a few seconds or more" (Balentine & Morgan, 1999, p. 139). Balentine and Morgan also advocated "low-level ticking or [a] similar sound" (p.

141) as appropriate auditory cues. They mention music as a potential waiting cue only to caution designers against it: "There are cases in which this [music or a spoken message] is not possible or desirable" (p. 140). Our results provide several more specific guidelines for telephony interfaces, expanding on the previous guideline and confirming that designers must consider SRT to create truly user-centered telephony designs.

Guideline 1. Do include a slow-rate waiting cue in telephony applications. Balentine and Morgan (1999) cautioned that their auditory waiting cue guideline is an item that "appear[s] to be common sense but [is] not supported by any evidence, or which seem[s] to work in practice but [is] easily missed by designers" (p. 26). Our studies begin to provide empirical evidence that support this guideline as a requirement of telephony interfaces. In our studies, 0.5-, 1-, and 2-sec ticking cues have advantages for the user. These tones confirm that the user is connected and the system is working. The 1- and 2-sec ticking cues have two additional benefits: They caused participants to underestimate the actual time they were waiting in our simulated applications and decreased the negative affect associated with waiting as compared to faster ticking rates. Our evidence indicates that designers should include waiting cues during SRTs in telephony applications.

Guideline 2. Limit the duration and variability of SRTs as much as possible. We found that even 3-sec SRTs result in user anxiety, stress, and impatience in simulated telephony applications. Our results are consistent with the results of previous studies that have shown that SRT itself can negatively impact users' emotional and physiological states (Guynes, 1988; Komatsubhara, Yokomizo, Yamamoto, & Noro, 1985; Kuhmann et al., 1987; Schleifer & Amick, 1989), and that anxiety is associated with SRTs (A. Eisler & Eisler, 1994; Guynes, 1988). This guideline is especially important in telephony applications because the cognitive load on users is greater than in desktop or other highly visual applications. Users must hold spoken commands and the structure of the interface itself in working memory while completing their primary task, process synthetic speech, and appropriately manage information exchange with the technology, all unique and unfamiliar aspects of telephony applications that are likely to increase user anxiety for these interfaces. In addition, it is these sophisticated abilities of telephony interfaces (speech recognition, information retrieval, and synthetic speech), coupled with the familiarity of a telephone receiver, that are also likely to make users less tolerant of SRT.

Guideline 3. Use a ticking tone with a 1- or 2-sec rate, which provides the best combination of shortened SRT perception with least negative emotional effects. Based on these findings, we recommend that if a speech recognition–telephony interface uses a ticking tone during system processing (as suggested by Balentine & Morgan, 1999), a slow rate of ticking is better than a fast ticking rate. A 2-sec ticking rate resulted in underestimation of SRT. A 1-sec ticking rate resulted in similar affect as silence but did not have a perceptual advantage of shortened SRT duration. An intermediate ticking rate (approximately 1.5 sec) may combine perceptual advantages with lower negative affect; however, this observation is speculative and should be evaluated by interface designers prior to its use in a particular application.

Guideline 4. Avoid ticking tone rates of greater than 0.5 sec. The most unambiguous finding in these studies is that fast ticking rates had perceptual disadvantages, causing participants to overestimate their waiting time, as well as increase their negative affect. Rapid ticking should be rejected as a system processing tone. For the designer, any auditory cue used during SRT should have a relatively slow rate of change or tempo to minimize user overestimation of SRT and negative affect.

Guideline 5. Select a waiting tone based on evaluation of proposed tones with the targeted user population. We found statistically significant interactions based on participants' age and gender. North et al. (1999) found that music that callers liked and fit their expectations positively influenced the waiting period in a telephone survey. These results confirm that design teams must carefully identify their user population and test their application with participants from the target population to ensure they provide optimal, user-centered feedback.

Guideline 6. Evaluate the priority of limiting SRT duration based on the targeted user population. Our studies suggest that user groups over 40 years of age (both men and women) may be less tolerant of long and variable SRTs because they provide longer estimates of time durations. These groups may especially benefit from slow tempo music or ticking to reduce their anxiety and make their waiting period seem shorter. Conversely, men under 40 years of age may be more tolerant of SRTs, even relatively long delays, because they consistently underestimate their waiting time.

Guideline 7. Evaluate proposed waiting cues with both long and short SRT durations. Interface designers should evaluate any proposed auditory processing tone with a variety of SRT durations. The finding of interaction effects in Experiment 1 suggests that a specific tone may have different effects with short and long SRTs.

In terms of application design, our studies more completely specify the range and type of waiting cues that should be included in telephony applications. They provide the weight of empirical data to extend and clarify guidelines for this emerging, highly constrained, and unique design environment. The studies also unmistakably highlight the need for a clearly defined user population for a particular application, as well as the importance of user-centered design and evaluation throughout the development process.

At a broader level, these studies justify the need for thoughtful interface design, informed and enhanced through its contact with documented cognitive-psychological phenomena. Our first study indicated that uninformed use of a ticking tone, an aspect of the telephony interface that designers may easily disregard, can have very negative consequences. Conversely, the second study demonstrated that the application of empirical work in the cognitive sciences to interface design can reduce these limitations. Therefore, these studies are consistent with the goal of cognitive engineering identified by various researchers (Falson, 1990; Gerhardt-Powals, 1996; Hollnagel & Woods, 1983): to identify guidelines for interface design based on human information processing abilities. As Shneiderman (1987) asserted more than 1 decade ago, this approach is not "the paint put on the end of a project" but "the steel frame on which the structure is built" (p. v).

REFERENCES

Albers, M. (1995). The Varese system, hybrid auditory interfaces, and satellite–ground control: Using auditory icons and sonification in a complex, supervisory control system. In *Proceedings of ICAD'94 The Second International Conference on Auditory Display* (pp. 3–13). Santa Fe, NM: Santa Fe Institute.

Albers, M. (1996). Auditory cues for browsing, surfing, and navigating the WWW: The audible Web. In *Proceedings of ICAD'96 The Third International Conference on Auditory Display* (pp. 85–90). Palo Alto, CA: International Conference on Auditory Display.

Albers, M., & Bergman, E. (1995). The audible Web: Auditory enhancements for Mosaic. In *Proceedings of CHI'95: The ACM Conference on Human Factors in Computing Systems* (pp. 318–319). Denver, CO: ACM.

Balentine, B., & Morgan, D. (1999). *How to build a speech recognition application: A style guide for telephony dialogues.* San Ramon, CA: Enterprise Integration Group.

Barber, R., & Lucas, H. (1983). System response time, operator productivity, and job satisfaction. *Communications of the ACM, 26,* 972–976.

Beaudouin-Lafon, M., & Conversy, S. (1996). Auditory illusions for audio feedback. In *Proceedings of CHI'96 Conference on Human Factors in Computing Systems* (pp. 299–300). New York: ACM.

Block, R. (1974). Memory and the experience of duration in retrospect. *Memory and Cognition, 2,* 153–160.

Block, R. (1990). *Cognitive models of psychological time.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Block, R., Zakay, D., & Hancock, P. (1998). Human aging and duration judgments: A meta-analytic review. *Psychology and Aging, 13,* 584–596.

Bradley, J. V. (1958). Complete counterbalancing of immediate sequential effects in a Latin square design. *Journal of the American Statistical Association, 53,* 525–528.

Butler, T. (1984). Computer response time and user performance during data entry. *AT&T Bell Laboratories Technical Journal, 63,* 1007–1018.

Buxton, W. (1989). Introduction to this special issue on nonspeech audio. *International Journal of Human–Computer Interaction, 4,* 1–9.

Carrasco, M., Guillem, M., & Redolat, R. (2000). Estimation of short temporal intervals in Alzheimer's disease. *Experimental Aging Research, 26,* 139–151.

Chebat, J., Gelinas-Chebat, C., & Filiatrault, P. (1993). Interactive effects of musical and visual cues on time perception: An application to waiting lines in banks. *Perceptual and Motor Skills, 77,* 995–1020.

Craik, F., & Hay, J. (1999). Aging and judgments of duration: Effects of task complexity and method of estimation. *Perception and Psychophysics, 61,* 549–560.

Dannenbring, G. (1983). The effect of computer response time on user performance and satisfaction: A preliminary investigation. *Behavior Research Methods and Instrumentation, 15,* 213–216.

Eisler, A., & Eisler, H. (1994). Subjective time scaling: Influence of age, gender, and type A and type B behavior. *Chronobiologia, 21,* 185–200.

Eisler, H. (1976). Experiments on subjective duration 1868–1975: A collection of power functions exponents. *Psychological Bulletin, 83,* 1154–1171.

Ezrati-Vinacour, R., & Levin, I. (2001). Time estimation by adults who stutter. *Journal of Speech Language and Hearing Research, 44,* 144–155.

Falson, P. (1990). *Cognitive ergonomics: Understanding, learning, and designing human computer interaction.* New York: Academic.

Fink, A., & Neubauer, A. (2001). Speech of information processing, psychometric intelligence and time estimation as an index of cognitive load. *Personality and Individual Differences, 30,* 1009–1021.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35,* 1–36.

Galloway, G. (1981). Response times to user activities in interactive man/machine computer systems. In *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 754–758). Santa Monica, CA: Human Factors and Ergonomics Society.

Gaver, W. (1997). Auditory interfaces. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 1003–1041). Amsterdam: Elsevier.

Gerhardt-Powals, J. (1996). Cognitive engineering principles for enhancing human–computer performance. *International Journal of Human–Computer Interaction, 8,* 189–211.

Glicksohn, J. (1992). Subjective time estimation in altered sensory environments. *Environment and Behavior, 24,* 634–652.

Goodman, T., & Spence, R. (1982). The effects of potentiometer dimensionality, system response time, and time of day on interactive graphical problem solving. *Human Factors, 24,* 437–456.

Guynes, J. (1988). Impact of system response time on state anxiety. *Communications of the ACM, 31,* 342–347.

Hargreaves, D. (1999). Can music move people? The effects of musical complexity and silence on waiting time. *Environment and Behavior, 31,* 136–149.

Hellstrom, A., & Almkvist, O. (1997). Tone duration discrimination in demented, memory-impaired, and healthy elderly. *Dementia and Geriatric Cognitive Disorders, 8,* 49–54.

Hellstrom, A., Lang, H., Portin, R., & Rinne, J. (1997). Tone duration discrimination in Parkinson's disease. *Neuropsychologica, 35,* 737–740.

Hollnagel, E., & Woods, D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man–Machine Studies, 18,* 583–600.

Hui, M., Dube, L., & Chebat, J. (1997). The impact of music on consumers' reactions to waiting for services. *Journal of Retailing, 73,* 87–104.

Jacko, J., Sears, A., & Borella, M. (2000). The effect of network delay and media on user perceptions of Web resources. *Behavior and Information Technology, 19,* 427–439.

Johansson, G., & Aronsson, G. (1984). Stress reactions in computerized administrative work. *Journal of Occupational Behavior, 5,* 159–181.

Jones, E., & Natale, T. (1973). Information processing theory of time estimation. *Perceptual and Motor Skills, 36,* 226.

Kamm, C., & Helander, M. (1997). Design issues for interfaces using voice input. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 1043–1059). Amsterdam: Elsevier.

Kohlisch, O., & Kuhmann, W. (1997). System response time and readiness for task execution: The optimum duration of inter-task delays. *Ergonomics, 40,* 265–280.

Komatsubhara, A., Yokomizo, S., Yamamoto, S., & Noro, K. (1985). Mental strain in a VDT task imposed by computer system response time. In *Proceedings of the 9th International Ergonomics Association Meeting* (pp. 316–318). Bournemouth, England: International Ergonomics Association.

Kuhmann, W., Boucsein, W., Schaefer, F., & Alexander, J. (1987). Experimental investigation of psychophysiological stress-reactions induced by different system response times in human–computer interactions. *Ergonomics, 30,* 933–943.

Lange, K., Tucha, O., Steup, A., Gsell, W., & Naumann, M. (1995). Subjective time estimation in Parkinson's disease. *Journal of Neural Transmision-Supplement, 46,* 433–438.
Levin, I., & Zakay, D. (Eds.). (1989). *Time and human cognition.* Amsterdam: Elsevier.
Lewis, J. R. (1993). Pairs of Latin squares that produce digram-balanced Greco-Latin designs: A BASIC program. *Behavior Research Methods, Instruments, & Computers, 25,* 414–415.
Martin, G., & Corl, K. (1986). System response time effects on user productivity. *Behavior and Information Technology, 5,* 3–13.
Meyer, J., Bitan, Y., & Shinar, D. (1995). Display a boundary in graphic and symbolic "wait" displays: Duration estimates and users preference. *International Journal of Human–Computer Interaction, 7*, 273–290.
Meyer, J., Shinar, D., Bitan, Y., & Leiser, D. (1996). Duration estimates and users' preferences in human–computer interaction. *Ergonomics, 39,* 46–60.
Meyer, J., Shinar, D., & Leiser, D. (1990). Time estimation of computer "wait" message displays. In *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 360–364). Santa Monica, CA: Human Factors and Ergonomics Society.
Miller, L., & Thomas, J. (1977). Behavioral issues in the use of interactive systems. *International Journal of Man–Machine Studies, 9,* 509–536.
Moore, B. (1989). *An introduction to the psychology of hearing* (3rd ed.). London: Academic.
Murray, R., & Abrahamson, D. (1983). The effect of system response time delay variability on inexperienced videotext users. *Behavior and Information Technology, 2,* 237–251.
Nichelli, P., Venneri, A., Molinari, M., Tavani, F., & Grafman, J. (1993). Precision and accuracy of subjective time estimation in different memory disorders. *Cognitive Brain Research, 1,* 87–93.
Nickerson, R. (1969). Man–computer interaction: A challenge for human factors research. *IEEE Transactions on Man–Machine Systems, 10*(4), 164–180.
North, A., Hargreaves, D., & McKendrick, J. (1999). Music and on-hold waiting time. *British Journal of Psychology, 90,* 161–164.
Ornstein, R. (1969). *On the experience of time.* New York: Penguin.
Planas, M., & Treurniet, W. (1988). The effects of feedback during delays in simulated teletext reception. *Behavior and Information Technology, 7,* 183–191.
Polkosky, M. (2001). *User preference for system processing tones* (Tech. Rep. No. 29.3436). Raleigh, NC: IBM.
Polkosky, M., & Lewis, J. (2001). *User preference for turntaking tones 2: Participant source issues and additional data* (Tech. Rep. No. 29.3447). Raleigh, NC: IBM.
Poynter, W., & Homa, D. (1983). Duration judgment and the experience of change. *Perception and Psychophysics, 33,* 548–560.
Rammsayer, T., & Rammstedt, B. (2000). Sex-related differences in time estimation: The role of personality. *Personality and Individual Differences, 29,* 301–312.
Ramos, L. (1993). The effects of on-hold telephone music on the number of premature disconnections to a statewide protective services abuse hot line. *Journal of Music Therapy, 30,* 119–129.
Ramsay, J., Barbesi, A., & Preece, J. (1998). A psychological investigation of long retrieval times on the World Wide Web. *Interacting With Computers, 10,* 77–86.
Rauterberg, M. (1998). About the importance of auditory alarms during the operation of a plant simulator. *Interacting With Computers, 10,* 31–44.
Riesen, J., & Schnider, A. (2001). Time estimation in Parkinson's disease: Normal long duration estimation despite impaired short duration discrimination. *Journal of Neurology, 248,* 27–35.
Roast, C. (1998). Designing for delay in interactive information retrieval. *Interacting With Computers, 10,* 87–104.
Schaefer, F. (1990). The effect of system response times on temporal predictability of work flow in human–computer interaction. *Human Performance, 3,* 173–186.

Schleifer, L., & Amick, B. (1989). System response time and method of pay: Stress effects in computer based tasks. *International Journal of Human–Computer Interaction, 1,* 23–39.

Schumacher, R., Hardzinski, M., & Schwartz, A. (1995). Increasing the usability of interactive voice response systems: Research and guidelines for phone-based interfaces. *Human Factors, 37,* 251–264.

Shneiderman, B. (1984). Response time and display rate in human performance with computers. *ACM Computer Surveys, 16,* 265–285.

Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human–computer interaction.* Cambridge, MA: Winthrop.

Spiegal, M., & Streeter, L. (1997). Applying speech synthesis to user interfaces. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 1061–1084). Amsterdam: Elsevier.

Thadhani, A. (1981). Interactive user productivity. *IBM Systems Journal, 20,* 407–423.

Thum, M., Boucsein, W., Kuhmann, W., & Ray, W. (1995). Standardized task strain and system response times in human–computer interaction. *Ergonomics, 38,* 1342–1351.

Watson, B., Walker, N., Ribarsky, W., & Spaulding, V. (1998). Effects of variation in system responsiveness on user performance in virtual environments. *Human Factors, 40,* 403–414.

Weiss, S., Boggs, G., Lehto, M., Shodja, S., & Martin, D. (1982). Computer system response time and psychophysiological stress. In *Proceedings of the 26th Annual Meeting of the Human Factors Society* (pp. 698–702). Santa Monica, CA: Human Factors and Ergonomics Society.

Yoblick, D., & Salvendy, G. (1970). Influence of frequency on the estimation of time for auditory, visual, and tactile modalities: The kappa effect. *Journal of Experimental Psychology, 86,* 157–164.

Zakay, D. (1989). Subjective time and attentional resource allocation: An integrated model of time estimation. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 365–397). Amsterdam: Elsevier.

Zakay, D. (1993). Time estimation methods—Do they influence prospective duration estimates? *Perception, 22,* 91–101.

Zakay, D., Lomranz, J., & Kaziniz, M. (1984). Extraversion–introversion and time perception. *Personality and Individual Differences, 5,* 237–239.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception & Psychophysics, 34,* 451–456.