Journal of Experimental Psychology:
Human Perception and Performance
2005, Vol. 31, No. 4, 749–761

0096-1523/05/$12.00 DOI: 10.1037/0096-1523.31.4.749

Learning to Achieve Perfect Timesharing: Architectural Implications of Hazeltine, Teague, and Ivry (2002)

John R. Anderson and Niels A. Taatgen
Carnegie Mellon University

Michael D. Byrne
Rice University

E. Hazeltine, D. Teague, and R. B. Ivry (2002) have presented data that have been interpreted as evidence against a central bottleneck. This article describes simulations of their Experiments 1 and 4 in the ACT-R cognitive architecture, which does possess a central bottleneck in production execution. The simulation model is capable of accounting for the emergence of near-perfect timesharing in Experiment 1 and the detailed data on the distribution of response times from Experiment 4. With practice, the central bottleneck in ACT-R will be reduced to a maximum of 50 ms (1 production cycle) and can often be much less, depending on timing of stages and variability in their times. The authors also show, with a mathematical analysis of E. Hazeltine et al.'s Experiment 2, that the expected dual costs for these kinds of highly practiced tasks will be small in many circumstances, often under 10 ms.

Keywords: dual task, perfect timesharing, ACT-R architecture, central bottleneck

Usually, people find it more difficult to perform two tasks at once than to perform a single task, even when the tasks involve different perceptual and response modalities. Such difficulties are often taken as evidence for a central bottleneck (Pashler, 1994; Welford, 1952). Recently, however, Schumacher et al. (2001) provided evidence that with enough practice and with enough incentive, participants could come to perform an aural–vocal task and a visual–manual task simultaneously with very little cost. This research was taken as evidence for the Meyer and Kieras (1997) *executive-process interactive control* (EPIC) theory, which postulates that central cognition is controlled by a parallel production system that is not subject to capacity limitations. More recently, Hazeltine, Teague, and Ivry (2002) followed up Schumacher et al. with a more extensive series of experiments that addressed some possible questions about the original research, and they also concluded that there was very little if any central bottleneck after considerable practice (a few thousand trials). The research in the Hazeltine et al. article and its implications are the focus of this article.

Byrne and Anderson (2001) published a model showing that the basic Schumacher et al. (2001) results could be accommodated in the ACT-R theory (Anderson & Lebiere, 1998), a production system that postulates that production-rule execution is serial and, therefore, constitutes a central bottleneck. Our purpose in the present article is to show that the ACT-R theory is compatible with the detailed data that Hazeltine et al. (2002) presented and that the learning mechanisms in the theory are capable of accounting for the reduction of dual-cost effects with practice. However, at the outset, we want to say that we suspect the ACT-R model we are offering is not correct in every detail. The real goal of this article is to show that assumptions like those in ACT-R are compatible with the Hazeltine et al. results and, in particular, that these results do not contradict a detailed central bottleneck theory. The main ACT-R-specific contribution of this article is to show how ACT-R's learning theory and its perceptual–motor theory can combine successfully in a way that is compatible with the Hazeltine et al. data. At the end of the article, after having covered the theory and the experiments, we note some details about the ACT-R account that could be problematic.

First, we review the task used by Schumacher et al. (2001) and by Hazeltine et al. (2002) and the basic ACT-R model that Byrne and Anderson (2001) proposed for this task. Then, we elaborate on how ACT-R can account for the learning results and the detailed Hazeltine et al. data.

In the original Schumacher et al. (2001) version of the task, which served as the basis for the first experiment described in Hazeltine et al. (2002), participants responded to the presentation of a circle and a tone. The circle appeared in one of three horizontal locations, and participants made a spatially compatible response with their right hand, pressing their index, middle, or ring finger to the left, middle, or right location, respectively. The 150-ms tone was either 220 Hz, 880 Hz, or 3520 Hz, and participants responded "one," "two," or "three." In the single-task condition, participants did just the visual–manual task or just the aural–vocal task. In the dual-task condition, both stimuli were presented simultaneously, and participants were asked to do both tasks simultaneously. Over 5 days of practice, participants came to respond virtually as fast at each task in the dual-task condition as

John R. Anderson and Niels A. Taatgen, Department of Psychology, Carnegie Mellon University; Michael D. Byrne, Department of Psychology, Rice University.

This research was supported by National Aeronautics and Space Administration Grant NCC2-1226 and Office of Naval Research Grant N00014-96-01491. We thank Eliot Hazeltine for his comments on drafts of the article and his assistance in understanding and modeling the data. We also thank David Meyer for his review and for subsequent discussions that helped us better understand EPIC and the issues raised by our modeling.

Correspondence concerning this article should be addressed to John R. Anderson, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: ja@cmu.edu

they did in the single-task condition. Thus, participants were able to perform two tasks at once with virtually no cost.

Figure 1 displays the original schedule chart for the task published in Byrne and Anderson (2001). The line labeled *Cognition* represents production firing in which each production takes 50 ms (an assumption shared by ACT-R and EPIC). There is one production that converts the visual stimulus into the manual response and another production that converts the tone into the speech act. The important observation is that because it takes longer to encode the sound, the two productions are offset from one another and do not interfere. (It did take participants considerably longer to perform the aural–vocal task than the visual–manual task.) The conditions in the later experiments of Hazeltine et al. (2002) eliminated this convenient offset of times by introducing a delay in the presentation of the visual condition and by increasing the difficulty of the visual–manual task either by making visual discrimination more difficult (Experiment 2) or by making the stimulus–response mapping incompatible (Experiments 3 and 4). Despite these changes, Hazeltine et al. continued to find virtually perfect timesharing. Nonetheless, as we show below, the ACT-R model does a fairly good job of simulating Hazeltine et al.'s results.

The ACT-R Model for Hazeltine et al. (2002)

In this section, we do not attempt an elaborate explanation of the ACT-R theory. Rather, we refer readers to Byrne and Anderson (2001), in which the perceptual–motor details are developed, and to Taatgen and Anderson (2002), in which the ACT-R assumptions about production learning are specified (see also Anderson et al., 2004, for the most current statement of the entire theory).

The production-learning model developed for ACT-R is one that takes declarative task instructions that are interpreted to perform the task and, with practice, converts them into production rules for directly performing the task. Early on, the task is heavy in demand on central cognition to interpret these instructions, but later, central cognition becomes a minimal bottleneck (as shown in Figure 1). This accounts for both the speedup in performance of the task and the elimination of much of the dual-task cost. The key to understanding the ACT-R learning model for this task is to understand the beginning and end states of the model as it learns to perform in the dual task. The model starts out with instructions for the task represented declaratively. These instructions can be rendered in English as follows:

a. When doing a pure aural block, prepare to respond to the detection of an aural stimulus with the aural task instructions.

b. When doing a pure visual block, prepare to respond to the detection of a visual stimulus with the visual task instructions.

c. When doing a mixed block, prepare to respond to the detection of a visual stimulus with the visual task instructions and to respond to the detection of an aural stimulus with the aural task instructions.

d. To perform the visual task, translate the visual location into a key, press that key, and check for success.

e. To perform the aural task, translate the aural tone into a word, say that word, and check for success.

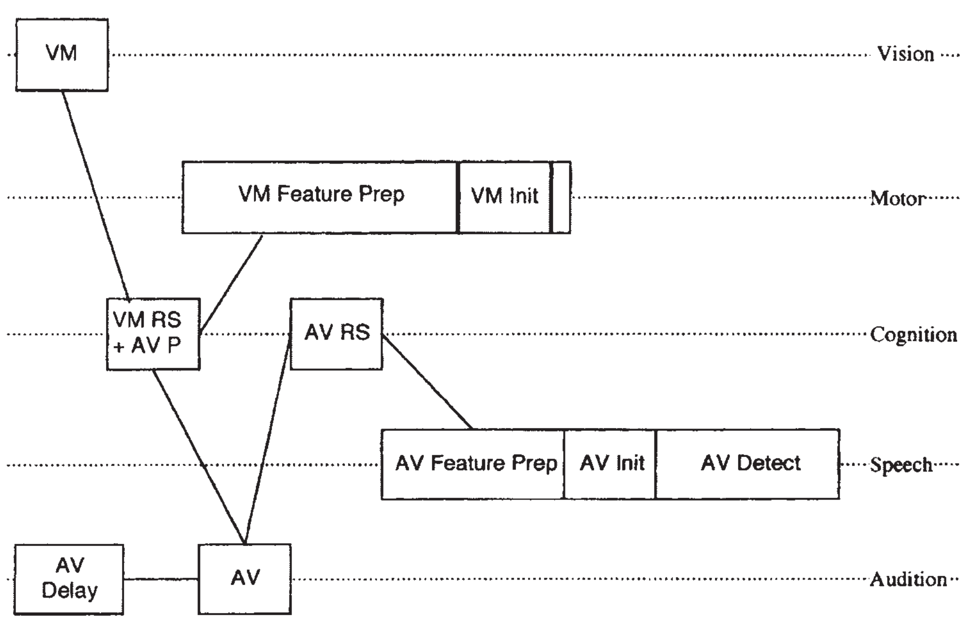

Figure 1. Byrne and Anderson's (2001) ACT-R schedule chart for Schumacher et al. (2001). VM = visual–manual task; Prep = preparation; Init = motor initiation; RS = response selection; AV = auditory–vocal task; P = initiate perception. From "Serial Modules in Parallel: The Psychological Refractory Period and Perfect Time-Sharing," by M. D. Byrne and J. R. Anderson, 2001, *Psychological Review, 108,* Figure 6, p. 856. Copyright 2001 by the American Psychological Association.

In addition, the model has committed to memory the mappings of the locations and sounds:

f. A left location translates to the index finger of the right hand.

g. A middle location translates to the middle finger of the right hand.

h. A right location translates to the ring finger of the right hand.

i. A low tone (220 hz) translates to saying "one."

j. A middle tone (880 hz) translates to saying "two."

k. A high tone (3520 hz) translates to saying "three."

As described in Anderson et al. (2004), ACT-R has general interpretative procedures for converting such declarative instructions into task behavior. Figure 2A illustrates the sequence of productions involved in interpreting these instructions in the mixed condition, in which both tasks are presented. Below, we step through the production rules.

A. Set Up

1. Retrieve Instruction: This retrieves Instruction c above.
2. Retrieve Steps: This retrieves the steps involved in that instruction—in this case, preparing to respond to stimuli in both modalities.
3. Prepare Visual: This sets the system to respond with Instruction d above when the visual stimulus is encoded.
4. Prepare Aural: This sets the system to respond with Instruction e above when the aural stimulus is encoded.
5. Ready: The system notes that it is finished processing the instruction and ready to respond to a stimulus.
6. Attend Visual: This requests encoding of the visual stimulus.
7. Attend Aural: This requests encoding of the aural stimulus.

B. Perform Visual–Manual Task

1. Focus on Visual: When the location is encoded, it requests retrieval of instruction (in d above).
2. Retrieve Instruction: This retrieves Instruction d above.
3. Retrieve Steps: This retrieves the steps involved in that instruction—in this case, translating the location into a finger, pressing that finger, and checking the result.
4. Translate Position: A request is made to retrieve the finger corresponding to the location.
5. Retrieve Finger: One of Facts f–h is retrieved.

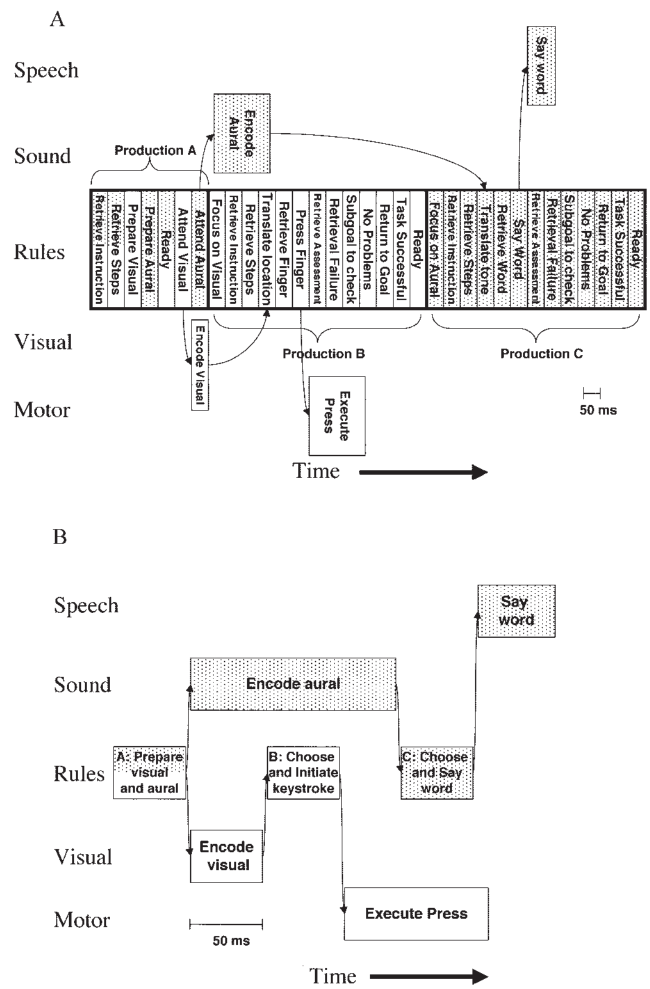

Figure 2. The operations of the ACT-R model at the beginning of the learning of Hazeltine et al. (2002) task (A) and at the end of the learning of the task (B). The lengths of the boxes reflect the average time for each operation. Those operations concerned with the aural–vocal task are shaded.

6. Press Finger: That finger is pressed.
7. Retrieve Assessment: In response to the step of checking, first a check is made whether the result is known.[1]
8. Retrieval Failure: At this starting point in the experiment, nothing can be retrieved.
9. Subgoal to Check: Therefore, set a subgoal to check the outcome.

[1] Both the visual–manual and the aural–vocal task end with an assessment of whether the task has been performed successfully. This assessment is critical to production learning. The task can be judged successful either by setting a subgoal to judge it or by retrieving information that this action has been successful in the past. The retrieval route is tried first but, initially, will fail until a reliable memory is built for the outcome of such a check. This enables us to model the process by which an explicit check is dropped out and built into the learned production rules.

10. No Problems: There is no negative feedback from the experiment.

11. Return to Goal: Return this determination to the main goal.

12. Task Successful: The task has been successfully accomplished.

13. Get Ready: The system notes that it is finished processing the instruction and ready to respond to a stimulus.

C. Perform Aural–Vocal Task

1. Focus on Aural: When the tone is encoded, it requests retrieval of instruction (in e above).

2. Retrieve Instruction: This retrieves Instruction e above.

3. Retrieve Steps: This retrieves the steps involved in that instruction—in this case, translating the tone into a word, saying that word, and checking the result.

4. Translate Tone: A request is made to retrieve the word corresponding to the tone.

5. Retrieve Word: One of Facts i–k is retrieved.

6. Say Word: That word is generated.

7. Retrieve Assessment: In response to the step of checking, first a check is made whether the result is known.

8. Retrieval Failure: At this starting point in the experiment, nothing can be retrieved.

9. Subgoal to Check: Therefore, set a subgoal to check the outcome.

10. No Problems: There is no negative feedback from the experiment.

11. Return to Goal: Return this determination to the main goal.

12. Task Successful: The task has been successfully accomplished.

13. Get Ready: The system notes that it is finished processing the instruction and ready to respond to a stimulus.

When the task is a single task, only B or C will be performed, and when it is a pure block, the preparation in A will be simpler. However, in all cases, the above is a rather laborious (if logical) interpretation of the instructions. As Figure 2A illustrates, production execution at this point in time will pose a significant central bottleneck. All of the productions for the aural task have to wait for completion of the productions from the visual task (or vice versa—there is no requirement that the visual task be performed first). In Figure 2A, there are perceptual encodings and motor actions, but they are not part of the critical path.

Production compilation will collapse pairs of productions together. In the limit, only three productions are required to do this task. All of the productions in Part A above can be collapsed into a single production that responds to the simultaneous presentation of a tone and a location with a request to encode them. The acquired production can be paraphrased as follows:

Production A
IF the goal is to perform in a mixed block
    and a tone has been sounded
    and a circle has appeared,
THEN encode the frequency of the sound
    and encode the location of the circle
    and prepare to respond to an encoding of the frequency with the aural task instructions
    and prepare to respond to an encoding of the location with the aural task instructions
    and note that things are ready.

Similarly, all of the productions in Part B above can be collapsed into a single production that responds to the appearance of the location with an appropriate keypress. This requires learning that the keypress will be successful. There are three productions learned for the three locations. The one for the left location can be paraphrased as follows:

Production B
IF the location has been encoded on the left,
THEN press the index finger
    and note that things are ready.

Similarly, Part C above can be compressed into single productions like the following:

Production C
IF the frequency of the tone has been encoded as 220 Hz,
THEN say "one"
    and note that things are ready.

Figure 2B illustrates the situation after production compilation. The situation is like that in Figure 1, in which the two productions for the two tasks are offset and do not interfere with one another.[2]

Figure 2B shows the extent to which learning can effectively proceed. On the one hand, neither of these perceptual–motor productions (B or C) can be collapsed with the first preparation, Production A, because the preparation production makes perceptual requests that require encoding from the environment before Productions B or C can fire. This is one example of how the perceptual events define the limits on combining productions. On the other hand, the system can attempt to create a combination of the last two productions:

Production B and C
IF the location has been encoded on the left
    and the frequency of the tone has been encoded as 220 Hz,

[2] There are some slight differences between this and Figure 1, because the production rules ACT-R learns are not identical to those Byrne and Anderson (2001) hand coded, but the differences do not affect the basic explanation of perfect timesharing.

THEN press the index finger
and say "one"
and note that things are ready.

However, because the location and sound are never encoded at the same moment in Hazeltine et al.'s (2002) first experiment (but see our model of Hazeltine et al.'s, 2002, Experiment 4), this production never gets to fire. This result, which is a natural outcome of the production-compilation mechanism, is critical to explaining one of the Hazeltine et al. results in their first experiment—namely, that participants trained on a subset of six of the nine possible combinations of three locations and three tones were able to transfer to the remaining three without showing any deficit. Separate productions are always required in this experiment to handle the two modalities, and such combination rules never get to be used.

There were a number of critical parameters determining the behavior of the models for these experiments. Among these are the timings of the operations shown in Figure 2. We assumed that (a) each production took an average of 50 ms, (b) the time to encode a visual location will be 50 ms, (c) the time to encode the auditory stimulus will be 130 ms, (d) the time to complete a fingerpress will be 100 ms, and (e) the time to trigger the voice key with an utterance will be 50 ms. The production-execution time is a basic parameter of ACT-R and EPIC, and the visual encoding time and manual times are close to their standard values in ACT-R (85 ms and 100 ms, respectively). The aural and vocal times were estimated in light of the data but are within the constraints suggested by Hazeltine et al. (2002). In addition, we assumed that all times had a 100% variability (the EPIC model has 67% variability[3])—that is, if the mean time was T, the actual times on a trial varied uniformly between $T/2$ and $T + T/2$. For instance, production times, with a mean of 50 ms, vary between 25 and 75 ms. These assumptions, especially those about variability, are a bit arbitrary, but they serve to establish plausible benchmarks for showing that the basic results of Hazeltine et al. can be predicted within the ACT-R framework, which is not that different from the EPIC framework, except for the assumption of a central bottleneck. In addition, two parameters controlled the rate of production learning: the learning rate for production utility was .05, and the s parameter controlling noise in utilities was .056. The first is a standard value for many models (e.g., Anderson et al., 2004), but the second was estimated to fit the learning data.

Hazeltine et al.'s (2002) Experiment 1

Hazeltine et al. (2002) performed four experiments. The first involved 9 participants. Seven of these participants continued to work through the remaining three experiments. We are concerned with modeling in detail the results of the first and fourth experiments. The first experiment followed a procedure very similar to that in the first experiment of Schumacher et al. (2001). On the 1st day, participants practiced just the single tasks. There then followed up to seven sessions on different days in which participants performed dual-task and single-task blocks. In the dual-task blocks, participants experienced six of the nine possible combinations of tone–location pairs. The goal was for participants to reach the point of performing as well in the dual task as in the single task. Seven of the participants reached this goal, but the data reported for this experiment were from all 9 participants. After completing this phase of the experiment, participants performed two more sessions during which they had to deal with the three remaining combinations of locations and tones that had been withheld as well as the other six. Our simulation of this experiment involved seven sessions—the first just a single task, followed by four sessions in which dual-task blocks were intermixed with single-task blocks, followed by two more sessions in which the transfer stimuli were introduced. In each session, the simulation experienced the same presentation sequence as the participants. We ran 20 simulated participants, which resulted in standard errors of less than 1 ms per estimated mean.

The learning results from Hazeltine et al.'s (2002) first experiment are illustrated in Figure 3A, and the results of the simulation are shown in Figure 3B. In Hazeltine et al.'s experiment and in our simulation, there were two kinds of single-task trials: trials that occurred in homogeneous blocks, in which participants were only responsible for these items, and trials that were interspersed among dual-task trials. Because there was virtually no difference between these two trial types in the data or in our simulation, we collapsed over these. Also, there were two types of dual-task trials: those that involved the original six pairings and those that involved the new ones. Because Hazeltine et al. found no difference between these two types of items and our model produced none, we also collapsed over those. Thus, Figure 3 only presents performance on single-task trials and dual-task trials for the visual–motor task and the aural–vocal task. Hazeltine et al. reported data for three periods of the experiment—the first two sessions, in which dual tasks were used; the last two sessions; and the two, even later transfer sessions. In our simulation, we collected the means of Sessions 2 and 3, Sessions 4 and 5, and Sessions 6 and 7, respectively, to correspond to these sessions.

The simulation reproduced the overall trend of reduced differences among conditions, particularly the reduced dual-task cost. The model starts out somewhat better in the single-task aural–vocal condition and somewhat worse in the single-task visual–manual condition, but it ends up at close to the same point as Hazeltine et al.'s (2002) participants. Although the correspondence is not perfect, we have reproduced the magnitude of the learning effects (both data and simulation show approximately a 100-ms improvement) and the dropout of the dual-task cost with practice (in both data and simulation, the average dual-task cost effect dropped from approximately 50 ms to approximately 10 ms). Also, the model predicts no difference between the new and old stimulus combinations in transfer, as was observed. Given the variability among participants contributing to these data, getting such ballpark effects is all that we would expect from the learning model. The

[3] One must be careful in comparing the EPIC and ACT-R variance assumptions, even if they both use uniform distributions. This is because the variability across stages in EPIC is perfectly correlated. That is, if one stage is 50% longer than the average, all stages are. In contrast, the variability in the times of different components in ACT-R is totally uncorrelated. Curiously, if the total response time involved three equal-length components, the predictions of the two models would be approximately the same. EPIC with its 67% assumptions predicts the standard deviation of times should be 19.2% of the mean, whereas ACT-R with its 100% assumption predicts the standard deviation will be 16.7% of the mean.

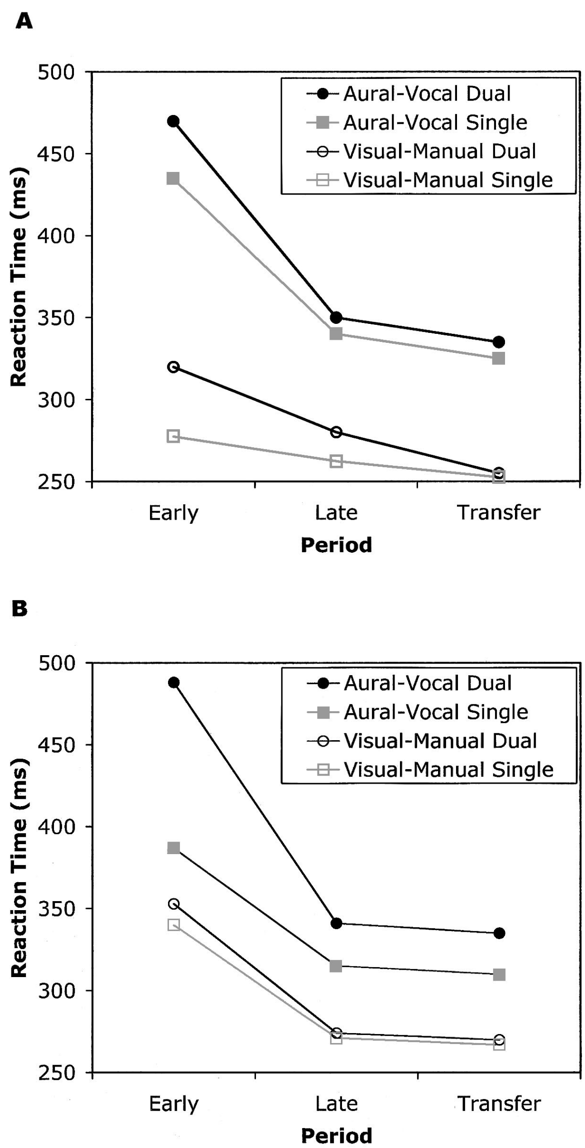

Figure 3. Learning to timeshare. A: Experiment 1 from Hazeltine et al. (2002). B: ACT-R simulation.

simulation of Experiment 4 deals with data in which participants are apparently more uniform and in which there is much more detail reported. There, we are concerned with more precise matches. From this simulation, we simply conclude that ACT-R can produce a reduction in dual-task cost that is qualitatively similar to what was observed.

Hazeltine et al.'s (2002) Experiment 2

The reason the model predicted more of a dual-task cost effect for the aural–vocal than for the visual–manual task is that the visual–manual task can complete its encoding more quickly than the aural–vocal task and, therefore, its central bottleneck production has a better chance of occurring in a position to block the production of the aural–vocal task. The second experiment of Hazeltine et al. (2002) used a discriminability manipulation (the introduction of distractor circles of a different intensity) to slow the visual–manual task without much effect on the dual-task cost for these materials. We do not provide a detailed simulation of this experiment but, rather, a mathematical analysis of its potential dual-task cost, to show—perhaps more transparently—why ACT-R does not predict much of a dual-task cost even when there is not a convenient offset in the average encoding times for the two tasks.

In this second experiment, the times were almost identical in the hard visual discrimination condition and the aural condition. Although this is probably not the exact model for each task, let us assume that both the aural encoding and the visual encoding took the 130 ms assumed for the aural encoding in Experiment 1. This means that on average, the two tasks would complete at the same time and should result in maximal interference. This is the worst-case analysis of the experiment. Should one encoding complete before the other on average (and we suspect that the visual encoding was still a little quicker), there would be less interference. The important complication concerns the variability in the encoding times. Although the mean encoding time is 130 ms, our model assumes a uniform distribution from 65 to 195 ms for each task, and these two distributions are independent of one another. For simplicity of analysis, we assume that the central bottleneck takes a constant of 50 ms, but the simulation for Experiment 4 allows for variability in the central bottleneck times as well.[4]

The following is an analysis of the delay that Task A will cause to the central processing of Task B, assuming that the asymptotic state in Figure 2B, in which each task only requires 50 ms of central processing, has been reached. Note that Task A can be either the visual–manual task or the aural–vocal task in this analysis, with Task B being the other.[5] The advantage of assuming that the encoding time for each task is 130 ms is that the analysis is symmetric and gives us the expected dual-task cost for either the visual–manual or the aural–vocal task. Task B will be delayed only if its encoding (Encoding B) finishes 0–50 ms after the encoding for Task A (Encoding A). If Encoding B finishes x (<50) ms after the Encoding A, its central processing will be delayed by $50 - x$ ms. There are two cases to consider:

a. Encoding A completes between 65 and 145 ms, leaving a full 50 ms for Encoding B to complete. Because the distribution is uniform, the probability that Encoding A will complete in this interval is $80/130 = 8/13$, and the probability that Encoding B will complete in the following 50 ms is 5/13. The mean delay will be 25 ms, because any delay between 0 and 50 ms is equally likely.

[4] But, in fact, there is virtually no difference if we include variability in this central 50 ms. In the example that follows, there is 8.9 ms delay with variability in the central bottleneck rather than 8.4 ms delay without variability.

[5] Please note that the terminology *Task A* and *Task B* are being used just to allow us to analyze the impact of the auditory–vocal task on the visual–manual task, and vice versa, in a single analysis. The terminology carries no implication about which encoding finishes first. We are calculating the impact of Task A on Task B, averaged over which encoding finishes first. Of course, Task B can have a negative impact on Task A only if its encoding does occur first.

b. Encoding A completes between at time t between 145 and 185 ms, leaving just $185 - t$ ms for Encoding B to complete at a time that will result in a delay to Task B. The probability of Encoding A completing in this period is 5/13, the probability of Encoding B completing afterward is on average 2.5/13, and the mean expected delay is 33.3 ms (a little calculus is required to establish this last number).

Thus, the total expected delay is:

$$\frac{8}{13} * \frac{5}{13} * 25 \text{ ms} + \frac{5}{13} * \frac{2.5}{13} * 33.3 \text{ ms} = 8.4 \text{ ms}.$$

Note that because this analysis is symmetric, there will be an average 8.4-ms slowing in Task A as well as in Task B. Hazeltine et al. (2002) observed a 19-ms slowing for the auditory task and a 1-ms slowing for the visual task. Thus, the average observed slowing is 10 ms, and the average predicted is 8.4 ms. However, empirically, there is an asymmetry in the sizes, suggesting the visual encoding is still completing somewhat before the auditory encoding.

The above analysis depends on a number of specific assumptions, and the reader might well wonder how much the expected delay would vary if we changed these assumptions. We performed a number of simulations to address this question:

a. We assumed that each encoding process lasted 130 ms and had a 130-ms range from 65 to 185 ms (which implies a standard deviation of 38 ms, or 29% of the mean). What would happen if we changed these assumptions about length of the interval and range? Treating this case as 100% range with a standard deviation that is .29 of the mean, Figure 4A shows the expected delay for different range ranges from 20% (or .06 standard deviation) to 180% (.52 standard deviation). As can be observed, over a wide range of values, the expected delay is below 15 ms and often below 5 ms. If we assume 67% range, as in the EPIC model, but still 130-ms processes, the expected delay is 11.7 ms. It is never more than 25 ms, because in the worst-case scenario, only one of the processes will be delayed, and never more than 50 ms, resulting in an average delay per process of $50/2 = 25$ ms.

b. One might wonder what would happen if we used a nonuniform distribution more like a reaction time distribution. Therefore, we simulated Weibull distributions with means and standard deviations equivalent to the distributions in Figure 4A. These results are displayed in Figure 4B. Weibull distributions are determined by two parameters, v and b, where v determines the shape, and b largely determines the scale. The v and b values are obtained in each case to get the corresponding standard deviation to a curve in Figure 4A. In the case of b, it is given as proportion of the mean—for example, if the mean is 100 ms and b is listed as 1.11, the actual value of b in this case is 111. As can be seen, there is virtually no difference between the results in Figure 4A with the uniform distribution and these results in Figure 4B. In the reference case of 130-ms times and 29-ms standard deviation, the expected delay is 8.6 ms, as compared with 8.4 ms for the uniform distribution. We should note that by assuming a Weibull distribution, our analysis becomes more similar to the stage analysis of Lien, McCann, Ruthruff, & Proctor (2005).

c. We have assumed so far that the two encoding processes take the same mean time. In Figure 4C, we assume the reference uniform distribution with a range equal to that of the mean and a standard deviation 29% of the mean. The different curves in Figure 4C give the mean delay of the process when it is competing with an encoding process that has a mean 50%, 75%, 100%, 150%, and 200% as long as it, with a standard deviation similarly 29% of its mean. As would be expected, the slower process is more likely to be delayed, but in most cases, the mean delay is under 15 ms, and in many cases it is under 5 ms.

d. Finally, just for completeness, we have included the same analysis as in c (above), using the corresponding Weibull distribution (i.e., $v = 3.88$ and $b = 1.11$ of the mean). Again, there are not strikingly different results with this distribution. We should note that one reason for using the uniform distribution is that it is easy computationally, and apparently the choice does not greatly matter for the issue of mean expected delay.

In conclusion, although the exact dual-task delay depends on assumptions, over a wide range of assumptions about distribution of times and differences in encoding-process lengths, there is relatively little expected delay assuming a maximum central bottleneck of 50 ms. In most cases, one encoding process does not finish often enough just before the other encoding process in a way that will cause a significant delay in the central bottleneck that follows the second encoding process.

Hazeltine et al.'s (2002) Experiment 4

To get a larger dual-cost one needs more central processing than the 50-ms single production. Hazeltine et al.'s (2002) third experiment introduced a stimulus–response compatibility manipulation for the visual–manual task, which, as we detail below, does require an additional rule in the ACT-R model. Still, that manipulation did not have much of an effect on the dual-task cost. Hazeltine et al.'s fourth experiment both used the compatibility manipulation and involved a variation in the onset of the two tasks. Because this fourth experiment was the most complex and reported the most elaborate data, we attempt to model it in detail. This series of experiments involved 7 of the participants from the original experiment who also took part in the second and third experiments. By the fourth experiment, they had come to display very rapid (about 250 ms for the visual–manual task and under 300 ms for the aural–vocal task) and stable responses. Thus, it is a good data set to look for detailed matches with the simulation.

These participants, after using only compatible stimulus–response pairings in the first two experiments, were asked to also be responsible for incompatible pairings in the third experiment. These incompatible mappings involved responding to the left location with the index finger (mapping unchanged), the middle

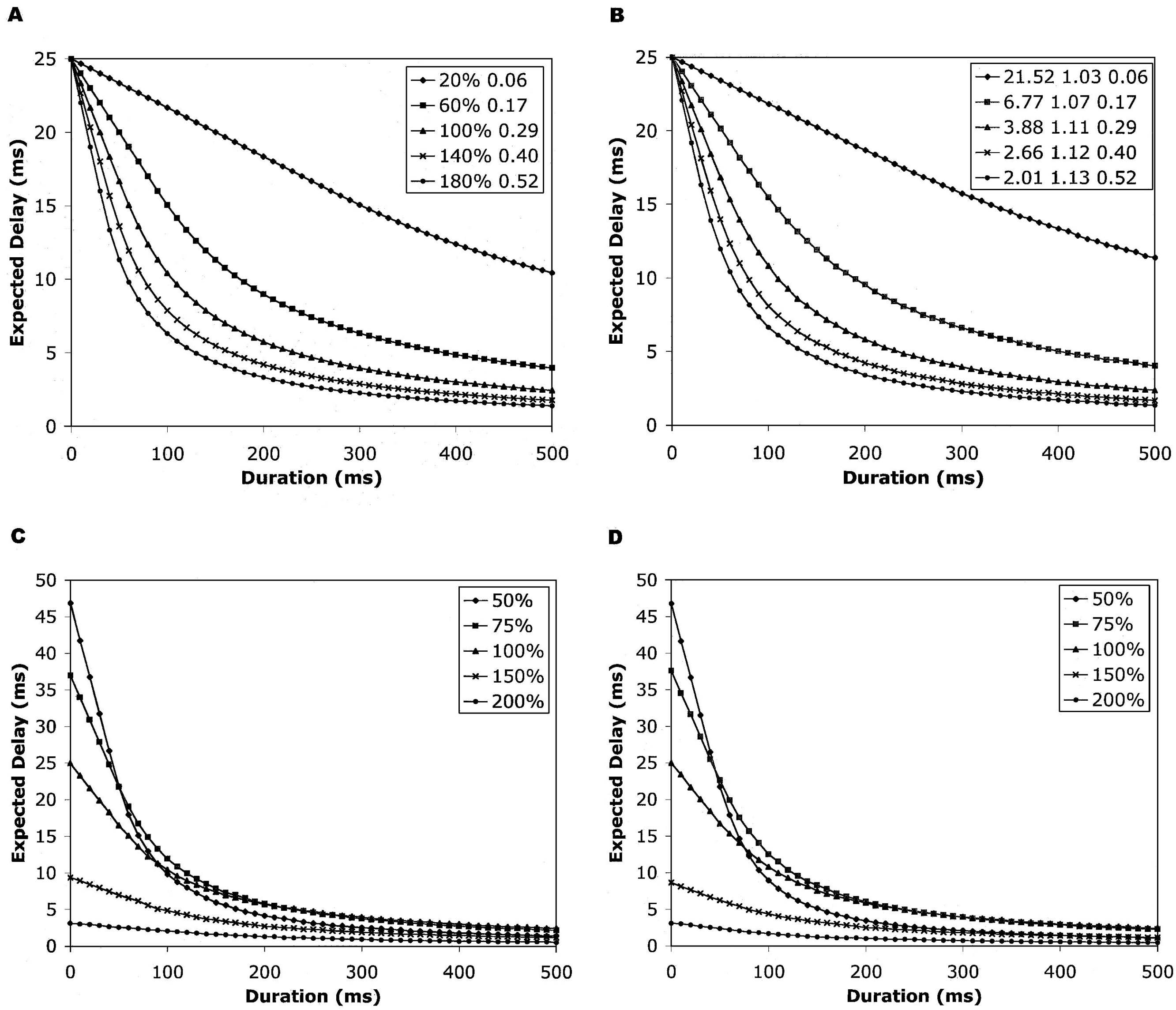

Figure 4. Mean expected delay to the 50-ms bottleneck for (A) uniform distributions of different mean duration (*x*-axis) and variability (different lines—range as percentage of mean and standard deviation as proportion of mean); (B) Weibull distributions of different mean duration (*x*-axis) and parameters (different lines for *v* and *b* given as proportion of *b*); (C) when the other distribution has different mean length (length as proportion of mean) in the case of a uniform distribution whose standard deviation is .29 of mean; and (D) when the other distribution has different mean length (length as proportion of mean) in the case of a Weibull distribution whose standard deviation is .29 of mean.

location with the ring finger (mapping changed), and the right location with the middle finger (mapping changed). Two test sessions were used in Experiment 4, and both involved compatible mapping blocks and incompatible mapping blocks. Because participants were responding so rapidly, we turned learning off in our simulation and assumed it was always behaving according to the terminal model in Figure 2B (thus speeding the simulation program and allowing us to collect large numbers of observations). The way we modeled the effect of compatibility was to assume that, in the incompatible condition, the participants processed the stimuli as the compatible condition and then converted their response to the incompatible response. Asymptotically, this conversion took just a single production, which made the incompatible conditions 50 ms slower—close to the observed deficit.

As an added manipulation, Hazeltine et al. (2002) presented the two stimuli either simultaneously or offset one from another by 50 ms. Thus, there were six dual-task conditions defined by whether the visual task involved a compatible mapping or not, crossed with whether there was a 50-ms head start for the visual task, or the two tasks were simultaneous, or there was a 50-ms head start for the aural task. In addition, there were two single-task aural conditions (homogeneous and heterogeneous) and four single-task visual conditions (compatible vs. incompatible crossed with homogeneous vs. heterogeneous). We ran 500,000 simulated trials in each condition to reduce the error in the estimate of ACT-R's predictions to less than 0.1 ms. This is excess precision for predicting mean times, but we also wanted to predict the distribution of times. To achieve the 500,000 trials, we actually created a simulation of the ACT-R simulation that just reproduced the keying timing relationships and did not have the full generality of the ACT-R model, because this is hardly required for modeling this experiment.

Figure 5A compares the predictions of the ACT-R model with the data for the visual–motor task, and Figure 5B compares the predictions and data for the aural–vocal task. For the visual task, the model and data show a strong 50-ms effect of the compatibility manipulation. The model does predict participants will be 7 ms faster on the visual task when given a head start on that task than when given a head start on the aural task, whereas there is no significant difference in the data. However, in the case of the aural task, participants averaged a significant 14 ms longer when the tone came first, and the model predicts 12 ms. The model predicts a disadvantage for both the aural–vocal and visual–manual task when there is a head start for the tone because the head start speeds up the aural task to the point at which the visual task is more likely to overlap. The data also show a significant 13-ms slowing of the aural–vocal task in the case of the incompatible visual mapping, whereas the model predicts 9 ms. Again, in the model, this is because the incompatible mapping slows the visual task to the point at which there is more likely to be a conflict between the two tasks.

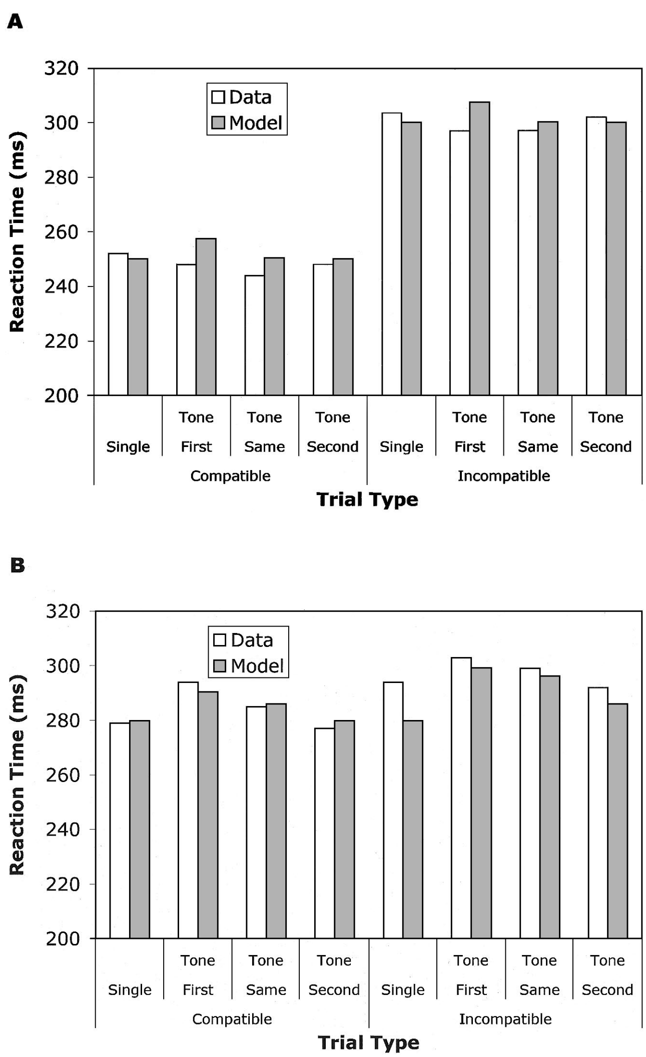

Figure 5. Comparison of data and simulation for Experiment 4 from Hazeltine et al. (2002). A: Visual–manual task. B: Aural–vocal task.

The correspondence between model and data is close if not perfect. The model predicts that the condition in which the tone sounds first will result in slower responses for both the visual–manual task and the aural–vocal task. It also predicts a larger dual-task deficit for the aural–vocal task. Although the observed deficit is larger in the case of the aural–vocal task, as predicted, there is basically no effect in the visual–manual task, unlike the prediction.

We get the largest dual-task cost when the aural task comes first, because giving it a 50-ms head start puts it into a range where its central bottleneck is more likely to compete with the central bottleneck for the visual task. This can be seen by inspecting Figure 6, which displays the mean timing of the various operations in the six dual-task conditions. Figures 6A–6C reflect the various compatible conditions. When there is a delay between the onset of the aural and visual stimuli (Figure 6A or 6C), separate productions are required to initiate the aural and visual encoding. Figures 6D–6F reflect the incompatible condition, in which the visual condition requires two productions—B1 (which is basically the same as B and produces the compatible mapping) and B2 (which converts that mapping). In Figure 6D, note that Production B1 fires but does not complete before the aural encoding is complete. In this case, a composite production (B2 & C) fires, combining the aural task and the second half of the visual task. Such composite productions never got an opportunity to fire in the compatible task (see Figures 6A–6C) because there was never a point at which productions for both tasks were simultaneously applicable.

Looking at Figure 6, it might seem particularly straightforward what the dual-task costs would be. The aural task is delayed 20 ms in the conditions illustrated in Figures 6A, 6D, and 6E, and it is not at all delayed in the conditions illustrated in Figures 6B, 6C, and 6F. The completion of the visual task is never delayed. Although this analysis would rather roughly correspond to the data, it ignores the complexities produced by the variability in the times. The maximum delay in any condition can be as large as 50 ms and as small as 0. It is also quite possible, when the aural task has a head start, for its production to intrude on the visual task and delay that.

Hazeltine et al. (2002) reported a simulation to determine how long the bottleneck could be to produce the dual-task deficits that they observed. They estimated that the bottleneck most plausibly is in the range of 20 to 30 ms. Their simulation did not allow for the possibility of variable length of stages and distribution of costs between both tasks. However, even so, their estimate was only a factor of 2 smaller than our maximum bottleneck cost in this task, which is the 50-ms production time, and only a factor of 2 larger than the predicted differences between the conditions.

It is important to realize that there is relatively little bottleneck possible in our model for the task. Bottlenecks become more

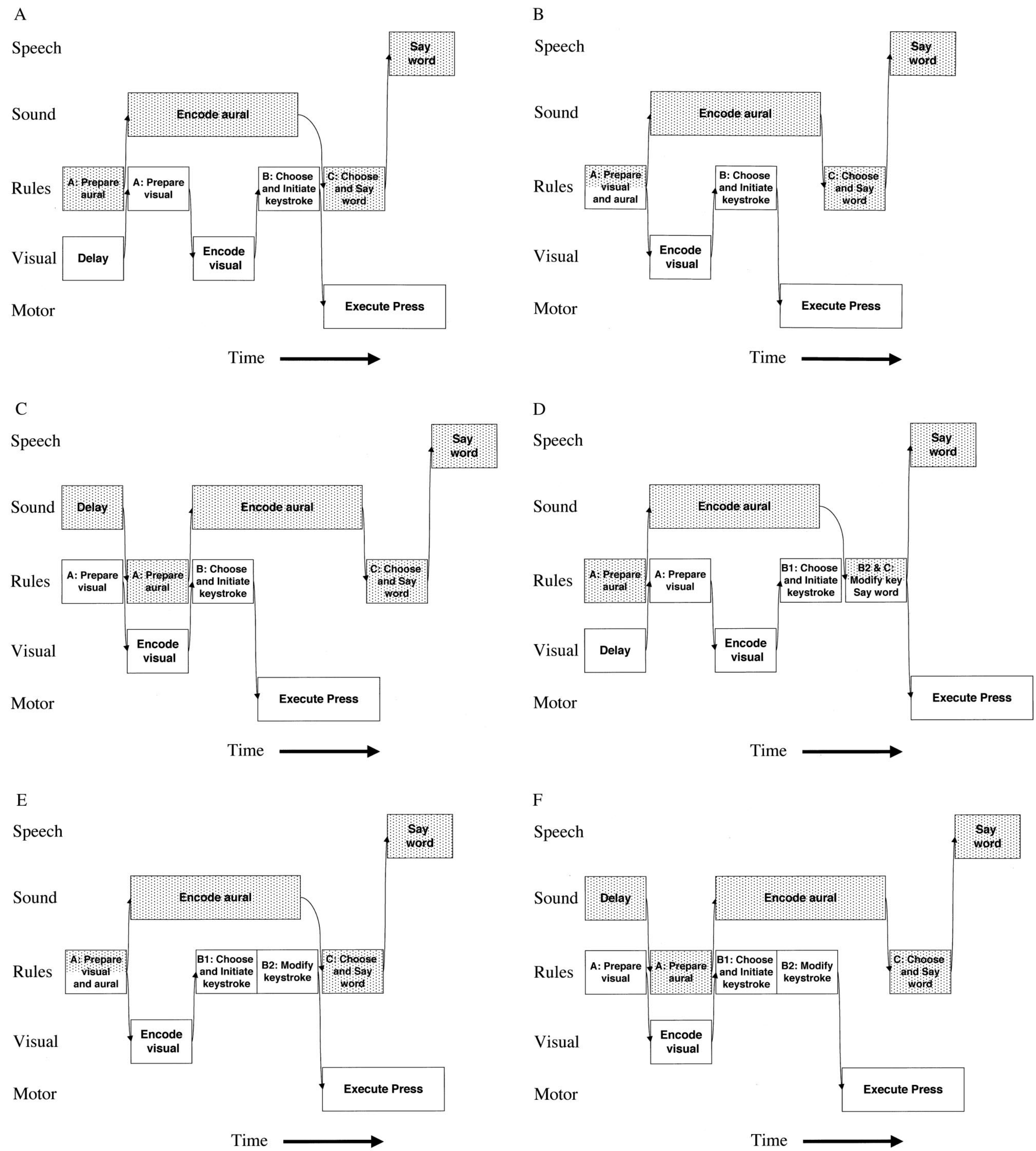

Figure 6. Timing of the ACT-R operations in the simulation of the six conditions in Experiment 4 of Hazeltine et al. (2002): compatible, tone first (A); compatible, simultaneous (B); compatible, location first (C); incompatible, tone first (D); incompatible, simultaneous (E); and incompatible, location first (F). The lengths of the boxes reflect the average time for each operation. Those operations concerned with the aural–vocal task are shaded.

significant when there is more cognition, as in Figure 2A. When one realizes that participants are producing only 250-ms latencies in the compatible visual task and 300-ms latencies in the incompatible visual task and the aural task, it should be apparent that there is not much time for central cognition to intervene between perception and action. Byrne and Anderson (2001) described tasks in which cognition is much more substantial and interference is much more substantial (over 1 s). However, they did not provide the extensive training of the current task. In principle, with enough practice, any task would be converted into one in which central cognition is almost totally eliminated and there is at most 50–100 ms of overlap in the central component, depending on compatibility. However, this requires converting all knowledge and contingencies into specific production rules. Although this is more than possible for the Hazeltine et al. (2002) tasks, the combinatorics for complex tasks would become overwhelming.

Our argument depends critically on the distribution of times for the two tasks. Hazeltine et al. (2002) provided some data to allow us to judge how closely our model corresponds to the variability in the actual data. Although Hazeltine et al. did not report data on the variability in times to reach the bottleneck, they did provide data on the variability in the difference between the completion times for the two tasks. These data are reproduced in Figure 7 for the six conditions along with our simulation. Plotted there are the probabilities that the difference between the visual and aural completion times will fall in various 25-ms bins. Given our rather blunt assumptions about stage variability, we think the correspondence between the distributions of time is stunning. One thing we would like to stress is the approximate correspondence between the variance of our distributions and the observed distributions. Across the six conditions, the average standard deviation for both data and theory is 58 ms. This indicates that our assumptions about the variability of the component processes are approximately correct.

One curious feature of the distributions (E. Hazeltine, personal communication) is that, both in the data and the model, there is a certain tendency for the distributions to be skewed with a negative tail. In the model, this is produced by the fact that the central bottleneck for the visual–manual task tends to occur before the bottleneck for the aural–vocal tasks. On trials in which the visual–manual task is by chance slower than usual or the aural–vocal task is faster than usual, we will have negative interresponse times in Figure 7. However, these trials will tend to move the aural–vocal bottleneck to where it delays the visual–manual, accentuating this effect and producing the negative skew. Just the opposite happens on those trials in which by chance the aural–vocal task is fast or the visual–manual task is slow. These are the trials that produce positive interresponse times in Figure 7. However, these trials will tend to move the auditory–vocal bottleneck beyond the range of the visual–manual bottleneck, thus eliminating the potential for large positive interresponse times.

Another interesting statistic reported by Hazeltine et al. (2002) concerns the correlation between the completion times for the two tasks. When the aural task came first, the correlation was .03; when the tasks were simultaneous, the correlation was .20; and when the visual task came first, the correlation was .24. Hazeltine et al. expressed puzzlement at why the correlation was different in the aural-first task. Our model's correlations were −.04, .20, and .17 in the three conditions. Hence, we are able to reproduce this pattern. The reason for the positive (if weak) correlation in the simultaneous and visual-first conditions is that both processes wait on the firing of the first production and will share its variability. Although this is still true in the aural-first condition, the interference in the bottleneck means that when one process is fast, it may interrupt and delay the other process, producing a negative relationship between response times.

Conclusions

In their conclusions, Hazeltine et al. (2002) write that "these results present a serious challenge to models of dual-task performance that postulate a unitary [central bottleneck]" (p. 541). We do agree that they pose a serious challenge, and we think we have shown that the ACT-R theory is up to accounting for (a) the learning trends, (b) the magnitude of the dual-task interference, and (c) the distribution of response latencies. It is interesting that ACT-R can achieve this despite the fact that it is really just a special case of a classical response-bottleneck model, which Hazeltine et al.'s data had been thought to contradict. However, ACT-R has a set of assumptions, largely similar to those in EPIC about stage duration and stage variability, that enabled it to predict the actual magnitude of the interference effects and the latency distributions (b and c above). What are most unique to ACT-R are its learning assumptions, and these enabled it to do a fair job of accounting for the learning trends in these data as well (a above). Altogether, these results display the merit of having a precise simulation model with a priori commitments to parameter values. ACT-R allows us to discover things that can be surprising—we had not anticipated capturing all of these trends and, indeed, had claimed we could not in a prior publication (Anderson et al., 2004). It also allows us to make exact value predictions so that we can truly judge whether the theory is compatible with the data. Exact value predictions are particularly important if one wants to judge the significance of the small dual-task costs that Hazeltine et al. found.

This being said, we would like to close by acknowledging that it is unlikely that the model presented here corresponds exactly to what is happening in the participants. We offer it to show that a theory with a central bottleneck like ACT-R is compatible with the reported data and that the ACT-R production-learning mechanisms can produce the emergence of near perfect timesharing with practice. As a sign that the current model is not perfect, there are effects in the data that our model does not account for. Notable in our minds are two. First, in Figure 2, our model is unable to explain the greater speedup that occurs with practice in the auditory single task than in the visual single task. This may reflect some perceptual learning that ACT-R does not model. Second, our model predicts that there should be small dual-task costs for both the visual–manual and the aural–vocal tasks in both Experiments 2 and 4. Although the dual-cost effects are *on average* about what our model predicts, they are almost exclusively found in the performance of the aural–vocal task. Although our model predicts slightly smaller effects for the visual–manual task, it does not produce the apparent 0 dual-task cost for that task. We should note that this asymmetry between the two tasks does not seem to have occurred in the original Schumacher et al. (2001) experiment. Rather, both tasks showed some residual dual-task cost as predicted by our model.

Also, we would like to acknowledge that we have doubts about the current ACT-R account of the stimulus–response compatibility effects in this experiment. The two-production asymptotic behavior of ACT-R in this task depends on having first learned the compatible response. As a consequence, the compatible production intruded first and had to be reversed by a second production. Had participants been first trained up on the incompatible condition, they would have suffered the same deficit in transferring to the

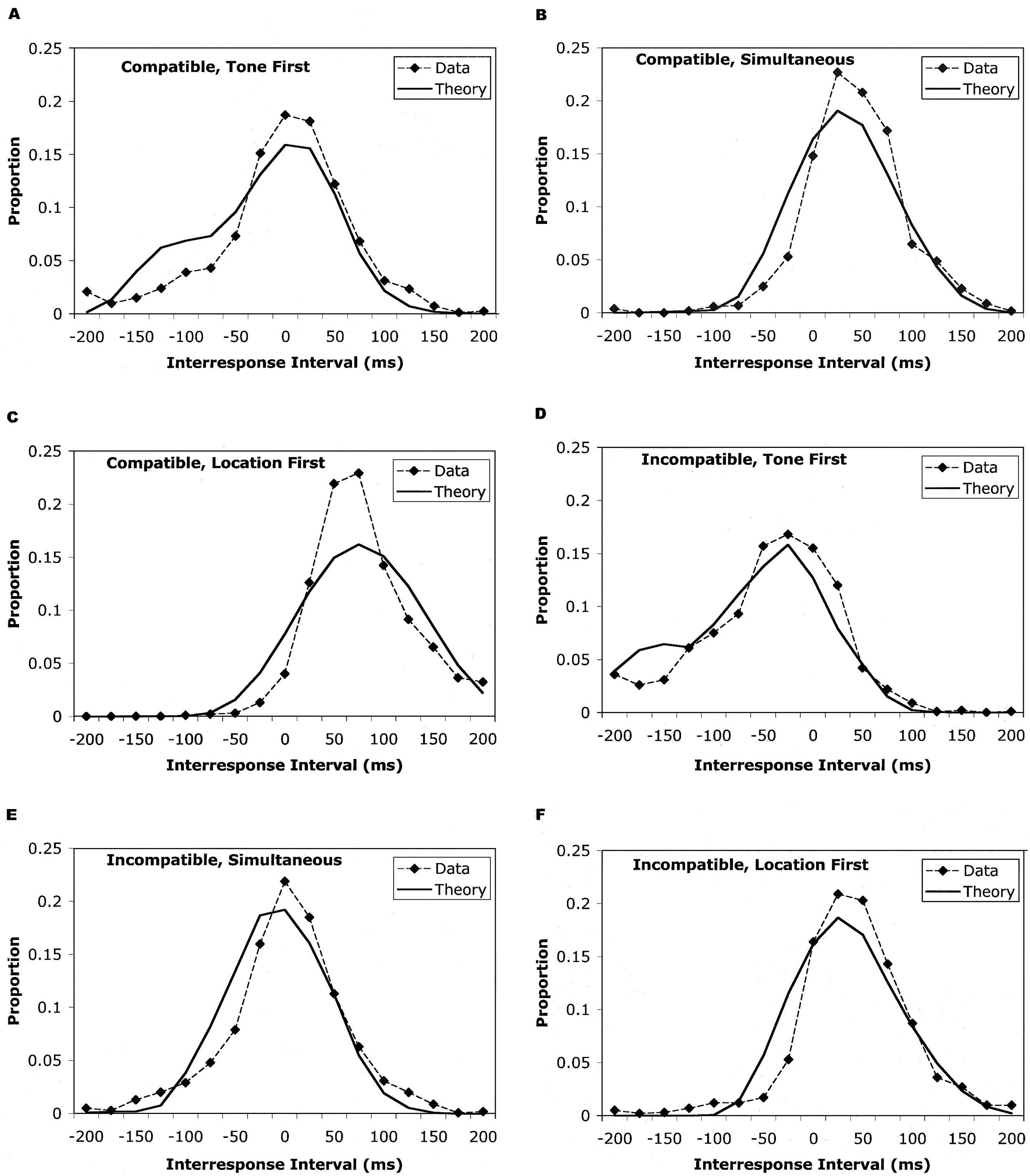

Figure 7. Observed and predicted distribution of intervals between the two responses in Experiment 4 of Hazeltine et al. (2002). These are computed by subtracting the reaction time for the visual task from the reaction time for the aural task. The six panels reflect the six conditions of the experiment.

compatible condition. This seems very unlikely. Basically, ACT-R probably does not have the right explanation of why two productions are required in the incompatible condition and only one is required in the compatible condition. However, given such a two-production model, ACT-R does seem to fit the data reasonably well and make the point that these results are not incompatible with the assumption of a central bottleneck.

Finally, we would like to state what we see as the main point of this article: If one assumes variable duration stages and a brief central bottleneck, as ACT-R does for these tasks, it will be difficult to discriminate such an account from an account without a central bottleneck. These experiments were carefully designed to serve that purpose, and we have shown that they are unable to rule out an account like that in ACT-R.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036–1060.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108,* 847–869.

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 527–545.

Lien, M.-C., McCann, R. E., Ruthruff, E., & Proctor, R. W. (2005). Dual-task performance with ideomotor–compatible tasks: Is the central processing bottleneck intact, bypassed, or shifted in locus? *Journal of Experimental Psychology: Human Perception and Performance, 31,* 122–144.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104,* 3–65.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12,* 101–108.

Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say "broke"? A model of learning the past tense without feedback. *Cognition, 86,* 123–155.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance—A review and a theory. *British Journal of Psychology, 43,* 2–19.

Received March 4, 2004
Revision received January 22, 2005
Accepted February 1, 2005 ■