

In Your Own Words: Using Full Sentences as Feedback

Jacob O. Wobbrock <jrock@cs.cmu.edu>
Human Computer Interaction Institute
Carnegie Mellon University

The Problem: Understanding Complex Settings

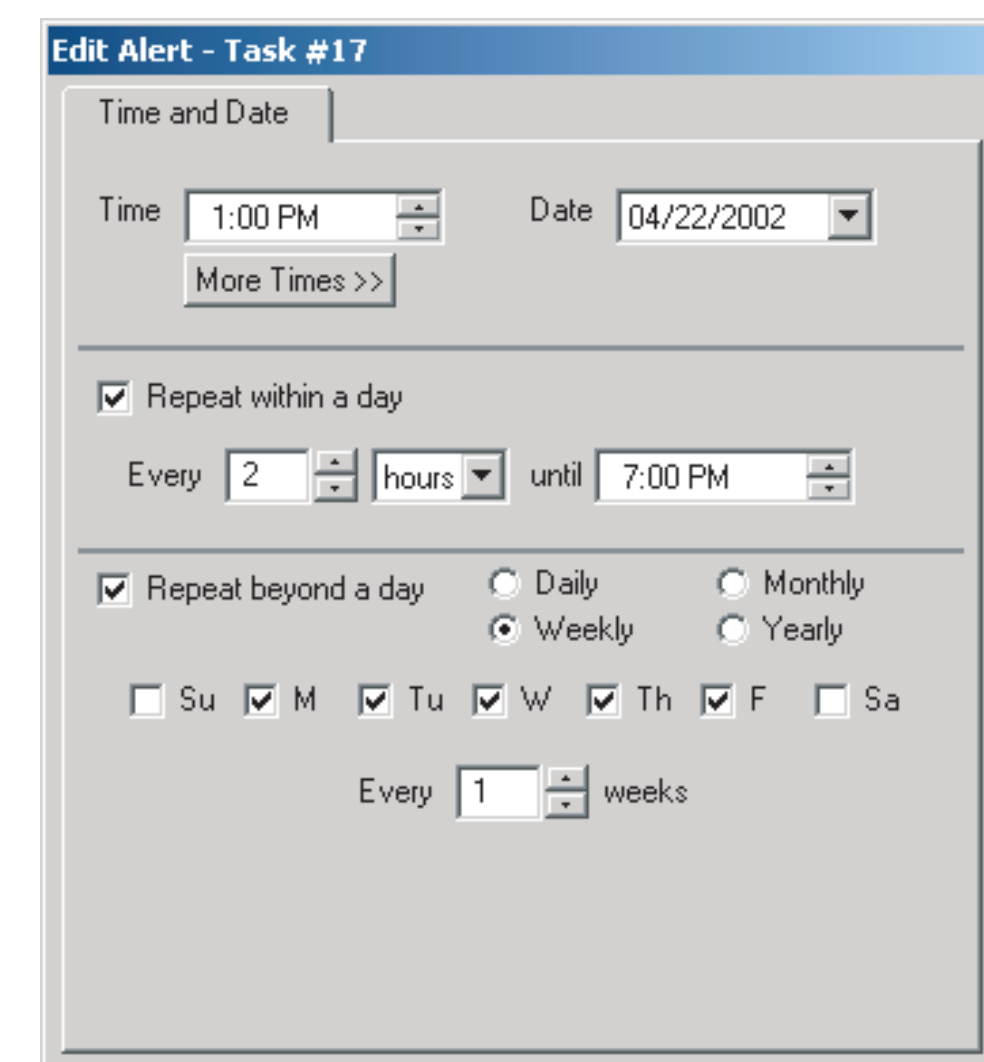
Modern software is complex. Dialogs often exhibit this complexity in the overuse of widgets. Widgets can be interdependent: the settings of some widgets may affect the states of others. While users can understand widget-level settings, understanding the overall configuration of complex dialogs can be difficult and error prone.

The Approach: Full Sentence Feedback

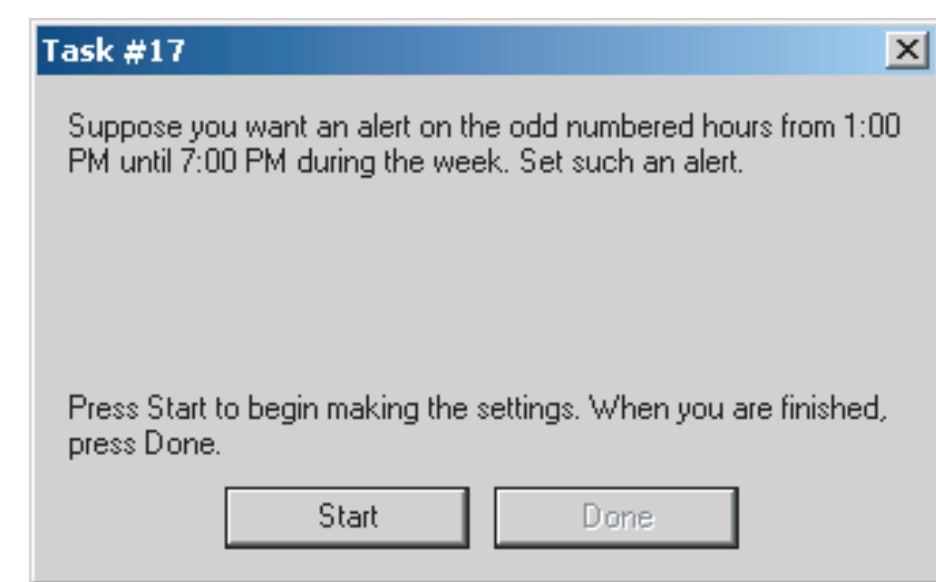
The states of all widgets on a dialog can be sent through a grammar, which generates an easily readable sentence. The sentence summarizes the overall configuration of the dialog, and uses color to distinguish changes from the previous configuration sentence. This high-level feedback helps users understand and feel confident in their settings.

The Experiment: Scheduling Alerts

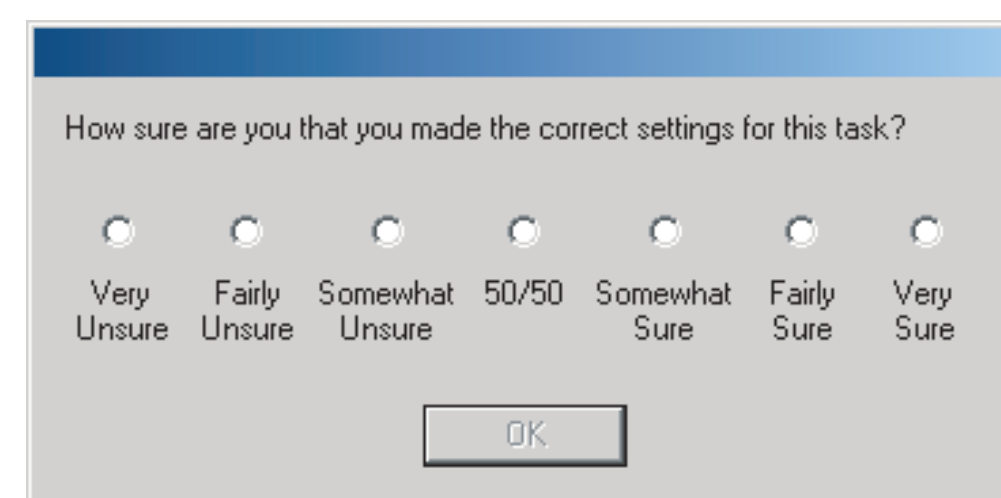
Twenty participants - 10 experts and 10 novices - used one of the two interfaces below to perform 30 progressively difficult alert scheduling tasks. The tasks were the same for all subjects. Half the participants used the interface on the left: no full sentence feedback. The other half used the same interface augmented with full sentence feedback, on the right. Measurements were taken for correctness, speed, and confidence.



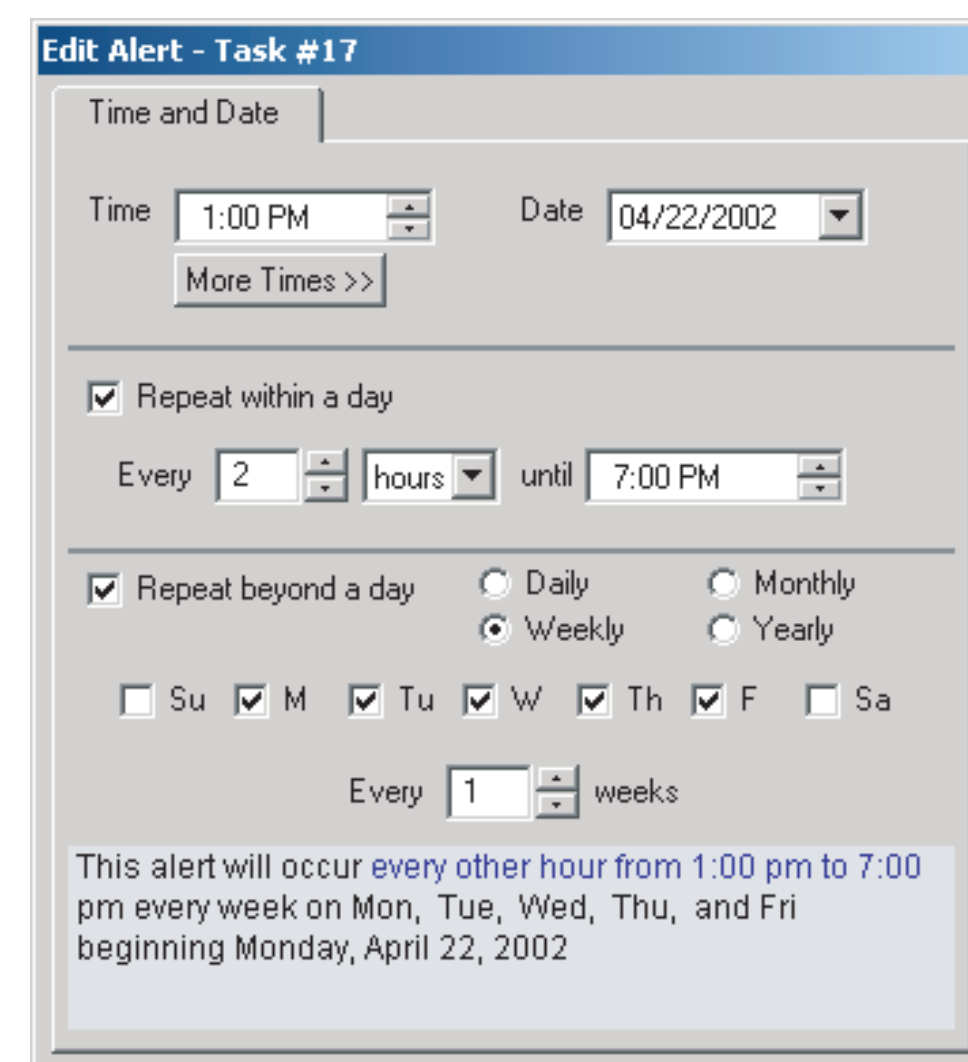
Condition 1: Interface without full sentence feedback



Task prompt



Confidence post-task measure



Condition 2: Interface with full sentence feedback

Experiment Results

When compared to experts without feedback, experts with feedback:

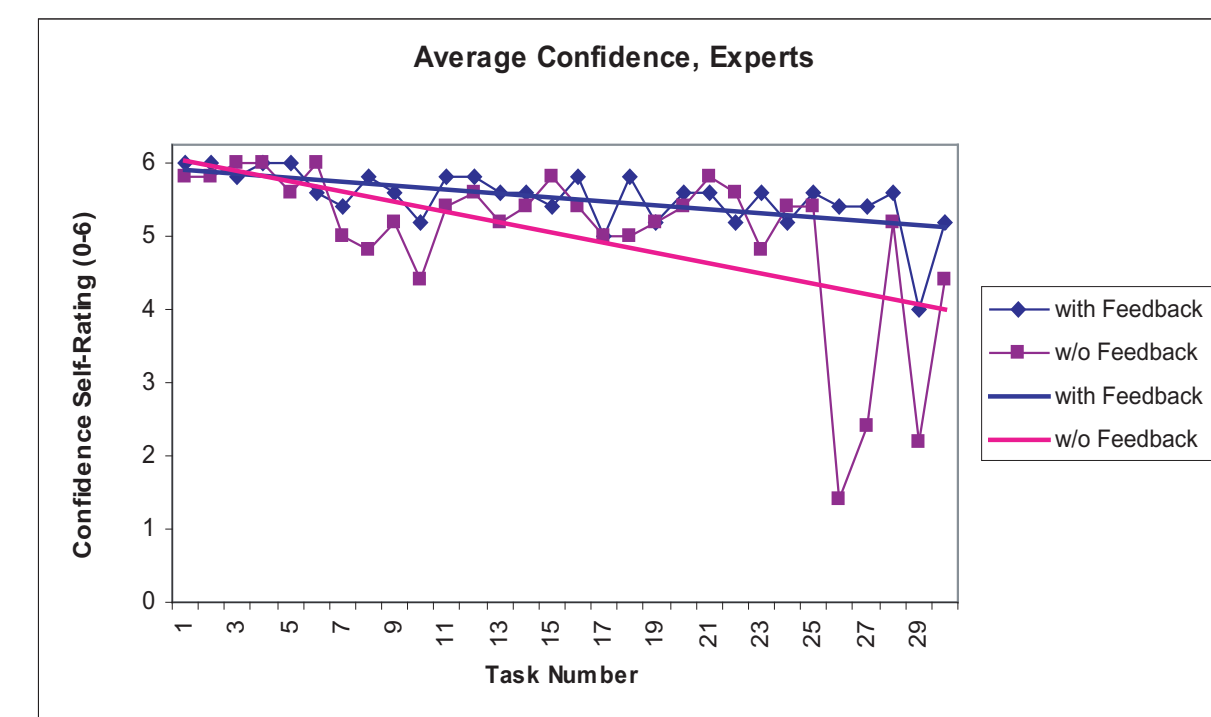
- were more confident
- were slower in completing tasks

When compared to novices without feedback, novices with feedback:

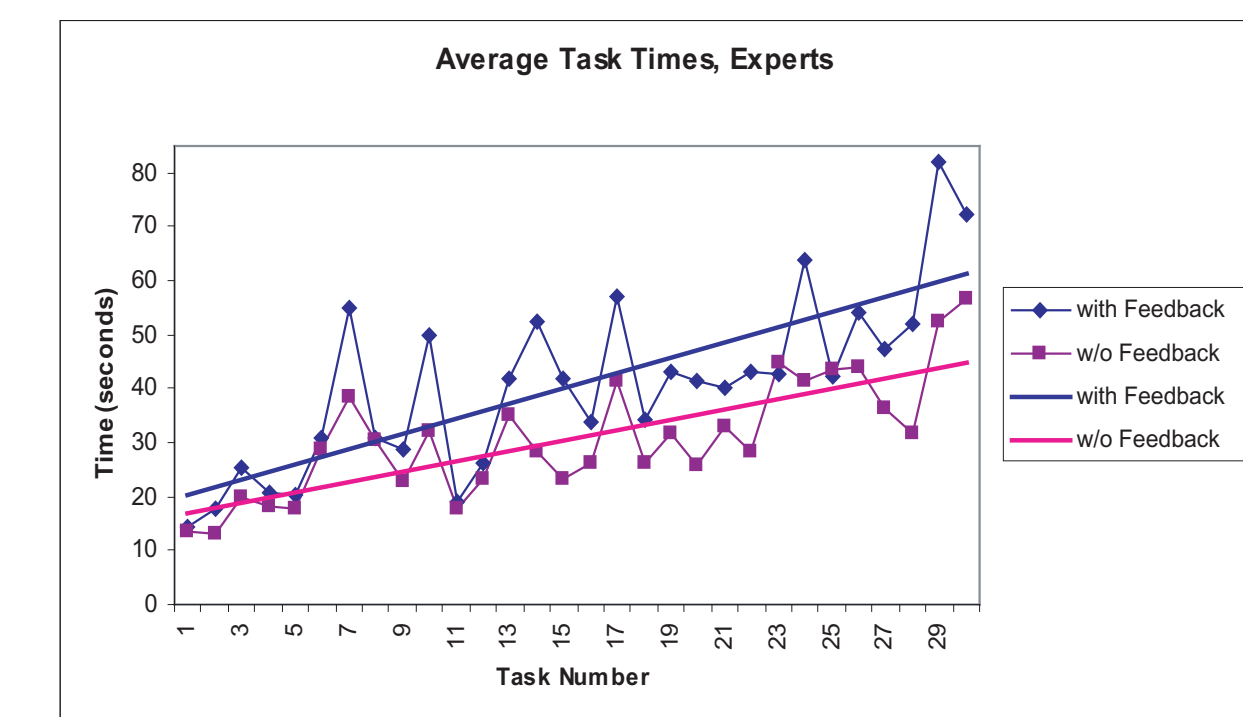
- spent less time viewing task prompts
- were more often correct
- were faster in completing tasks

When compared to novices, experts were [expertise validation]:

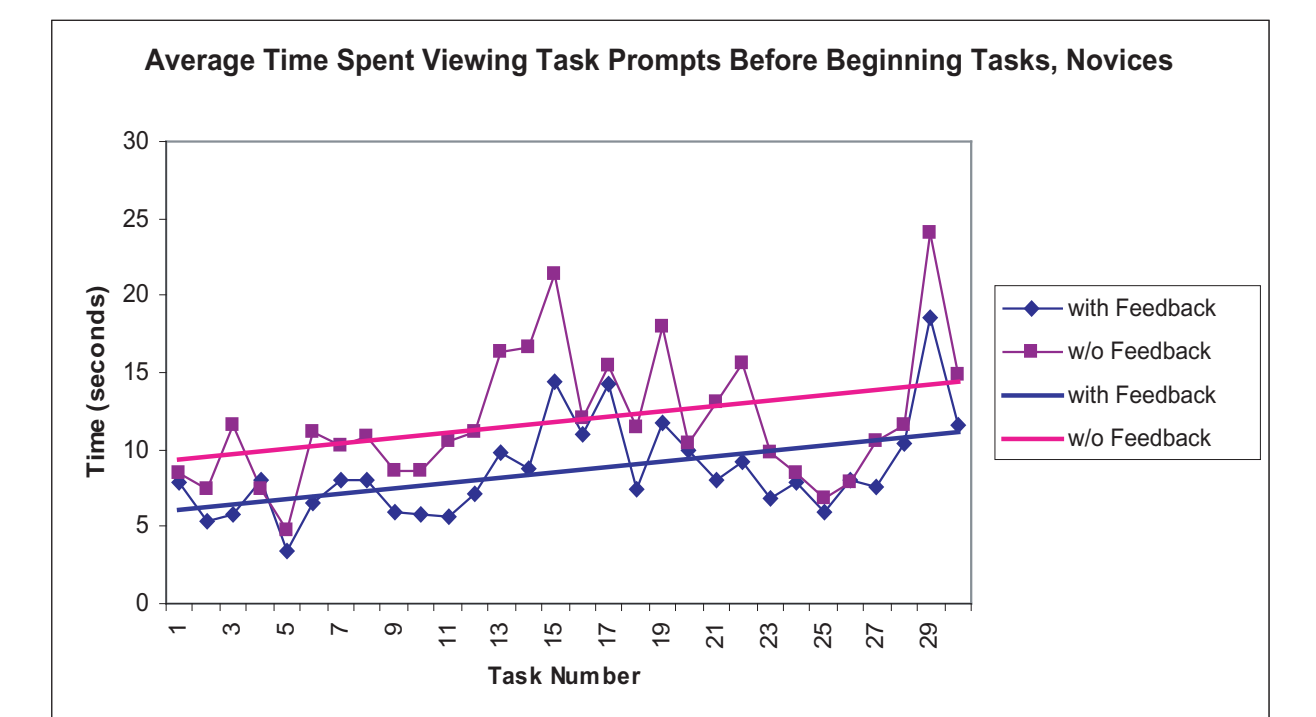
- more often correct when neither had feedback
- indistinguishable in correctness when both had feedback
- faster in task completion in both conditions



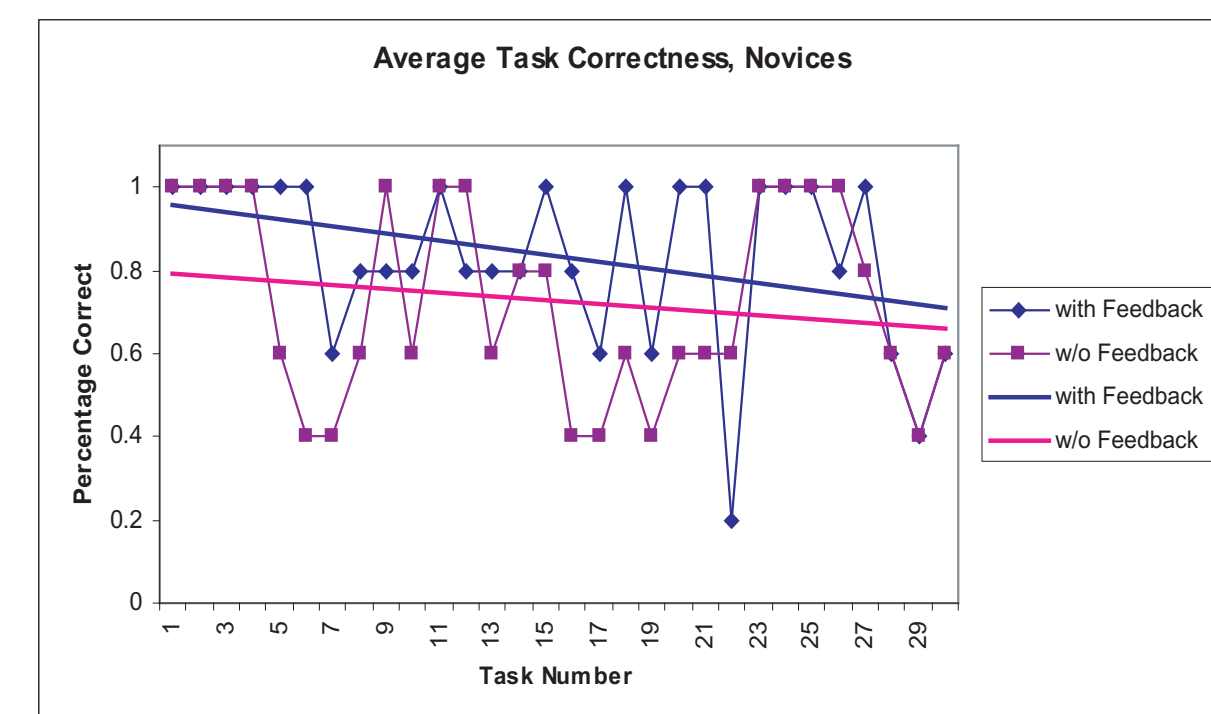
Experts with feedback were more confident than experts without it, 5.53 vs. 5.02 on a scale 0-6, $t(8) = -2.47, p < .05$.



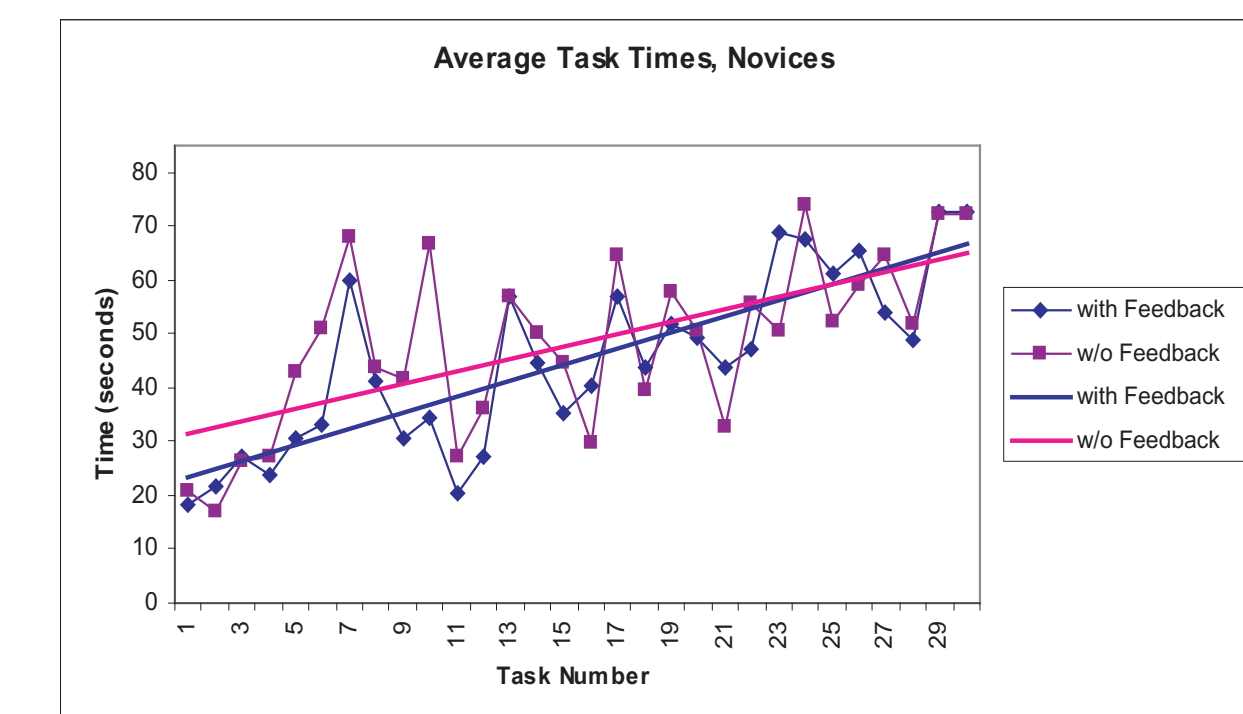
Experts with feedback were slower in completing tasks than experts without it, 40.85 vs. 30.87 seconds, $t(8) = -2.77, p < .05$.



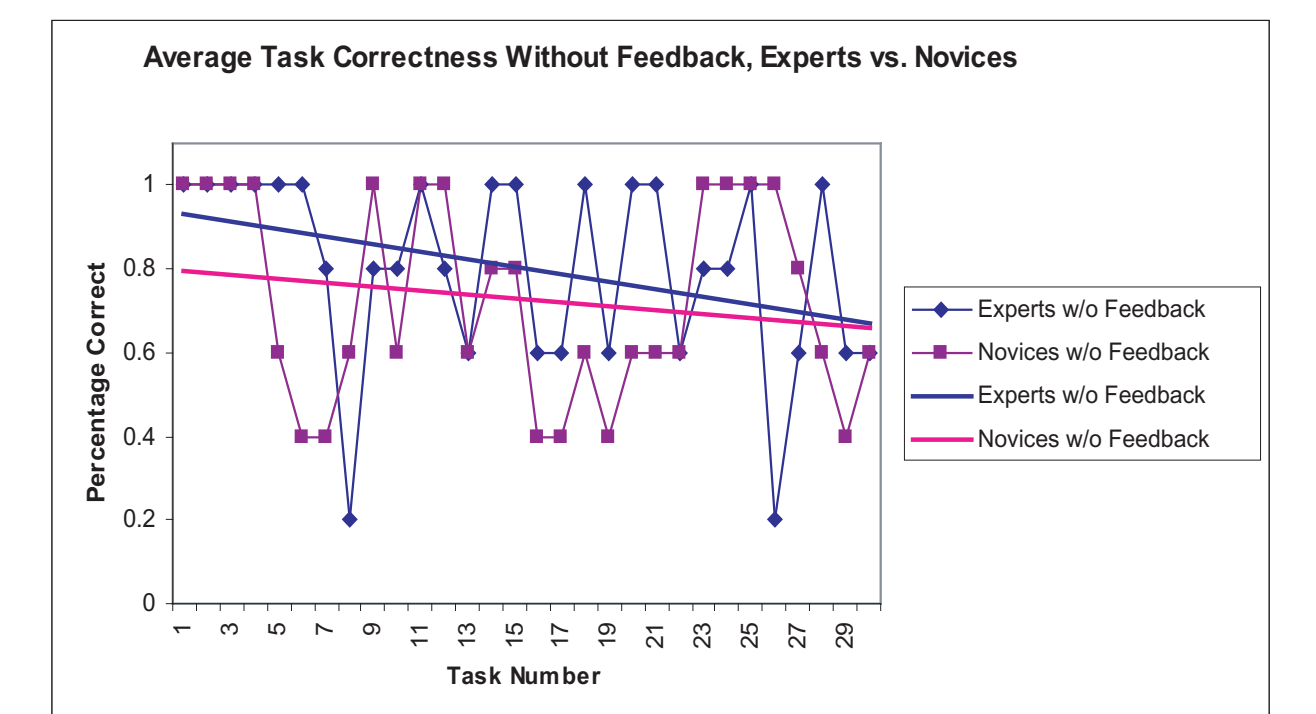
Novices with feedback spent less time viewing task prompts than novices without it, 8.63 vs. 11.83 seconds, $t(8) = -2.91, p < .05$.



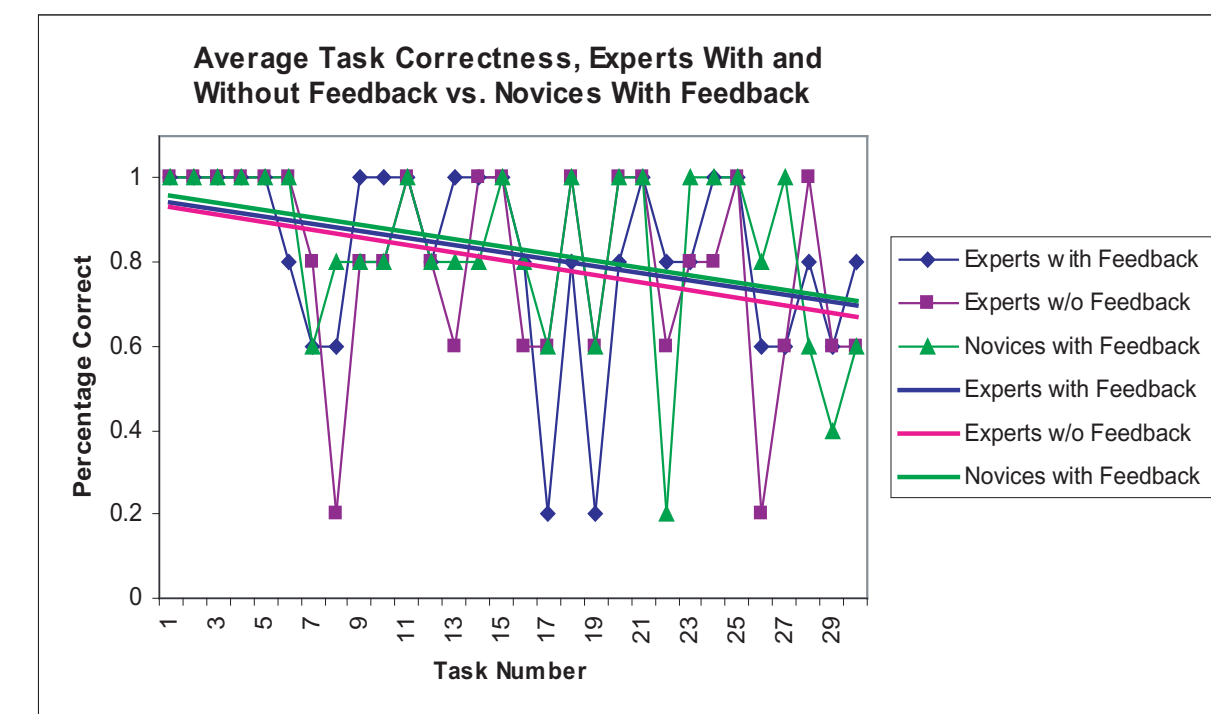
Novices with feedback were more often correct than novices without it, 83.3% vs. 72.7%. These averages fell just shy of desired significance.



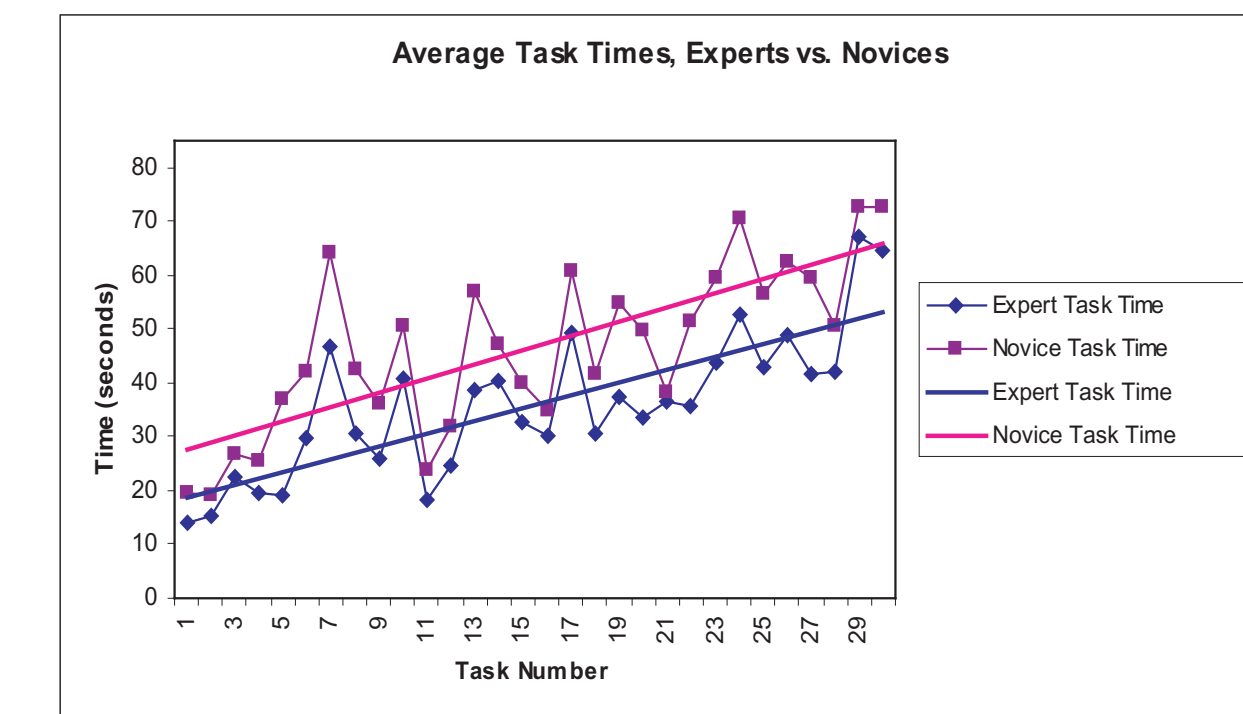
Novices with feedback did not exhibit the slow-down that experts did; in fact, on average they sped up 4.2 seconds with feedback.



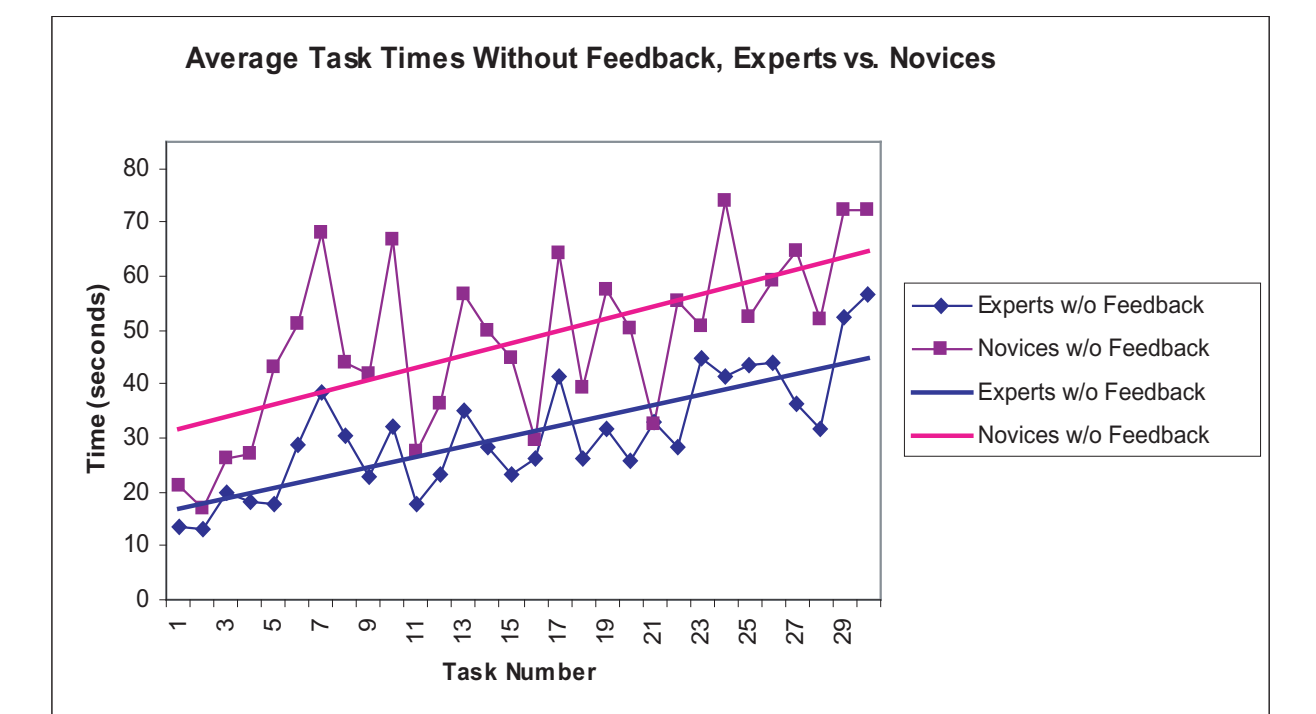
When neither group had feedback, experts were more often correct than novices, 80.0% vs. 72.7%, validating the expertise measure.



With feedback, novices were made indistinguishable from experts in task correctness. Note the contrast to the previous graph.



Overall, experts were faster in completing tasks than novices, 35.86 vs. 46.64 seconds, $t(18) = 2.57, p < .05$, validating the expertise measure.



When neither group had feedback, experts were faster in completing tasks than novices, 30.87 vs. 48.36 seconds, $t(8) = -2.84, p < .05$.

Test Yourself! For each of these interfaces, try to discern the overall configuration without full sentence feedback. Then, uncover the same interface with feedback, and check your answer.

