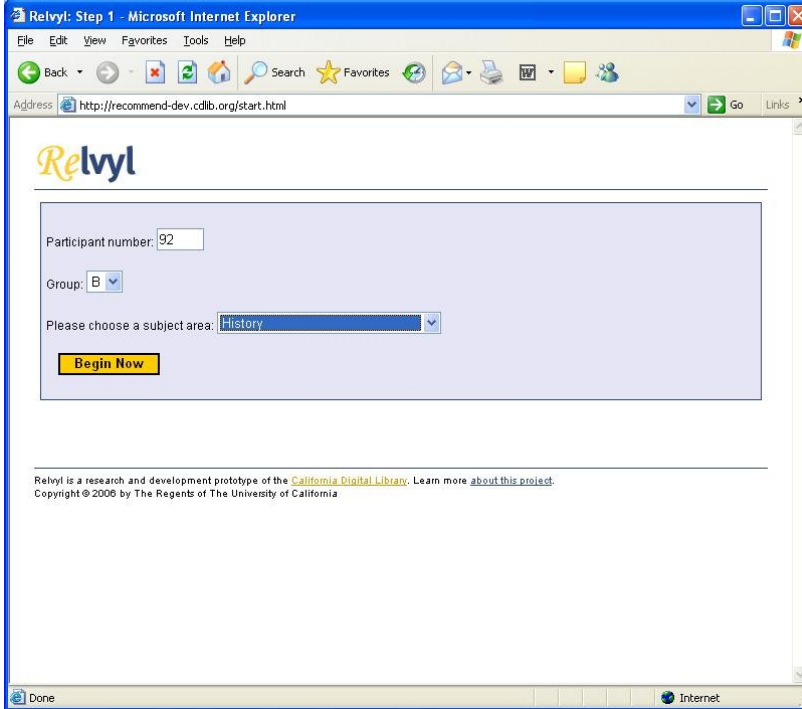
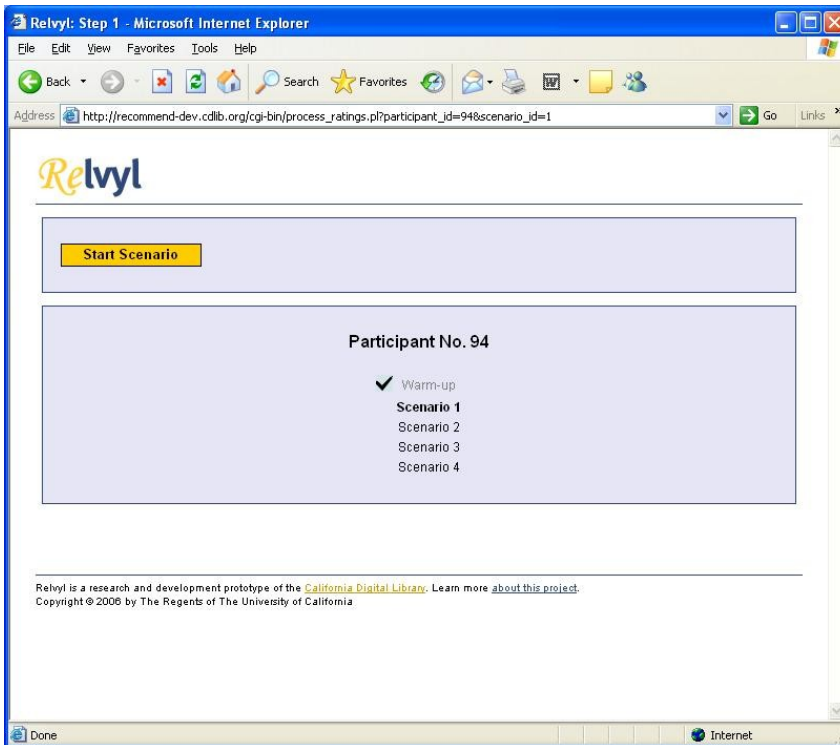


## Screen Shots from Recommending Assessment

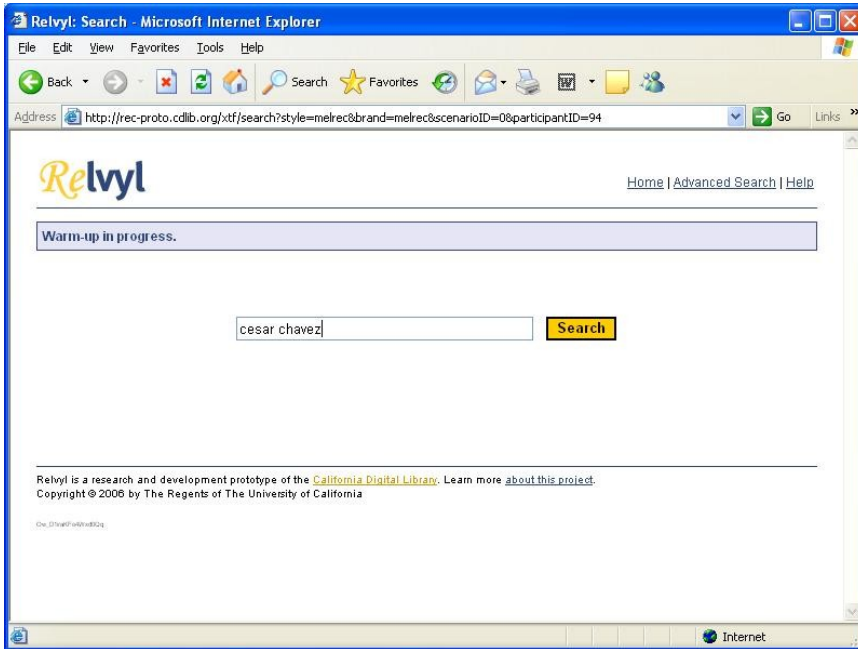
**Figure 1.** Participants were instructed to enter an assigned participant number, a group code, and a major or field of study.



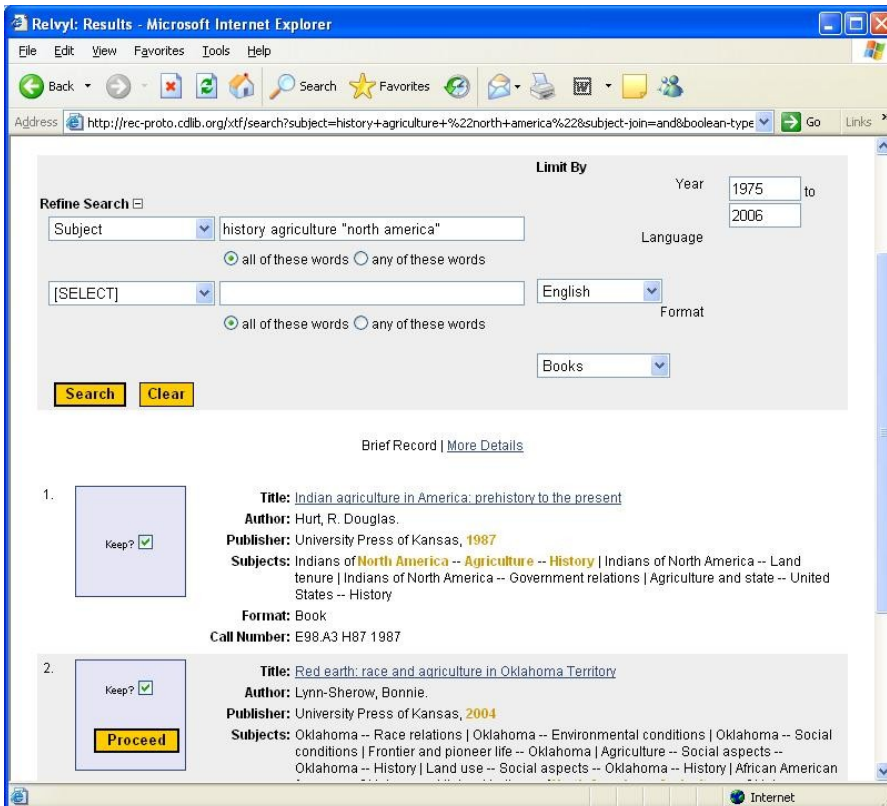
**Figure 2.** A “home page” that tracked progress through the protocol was shown between scenarios.



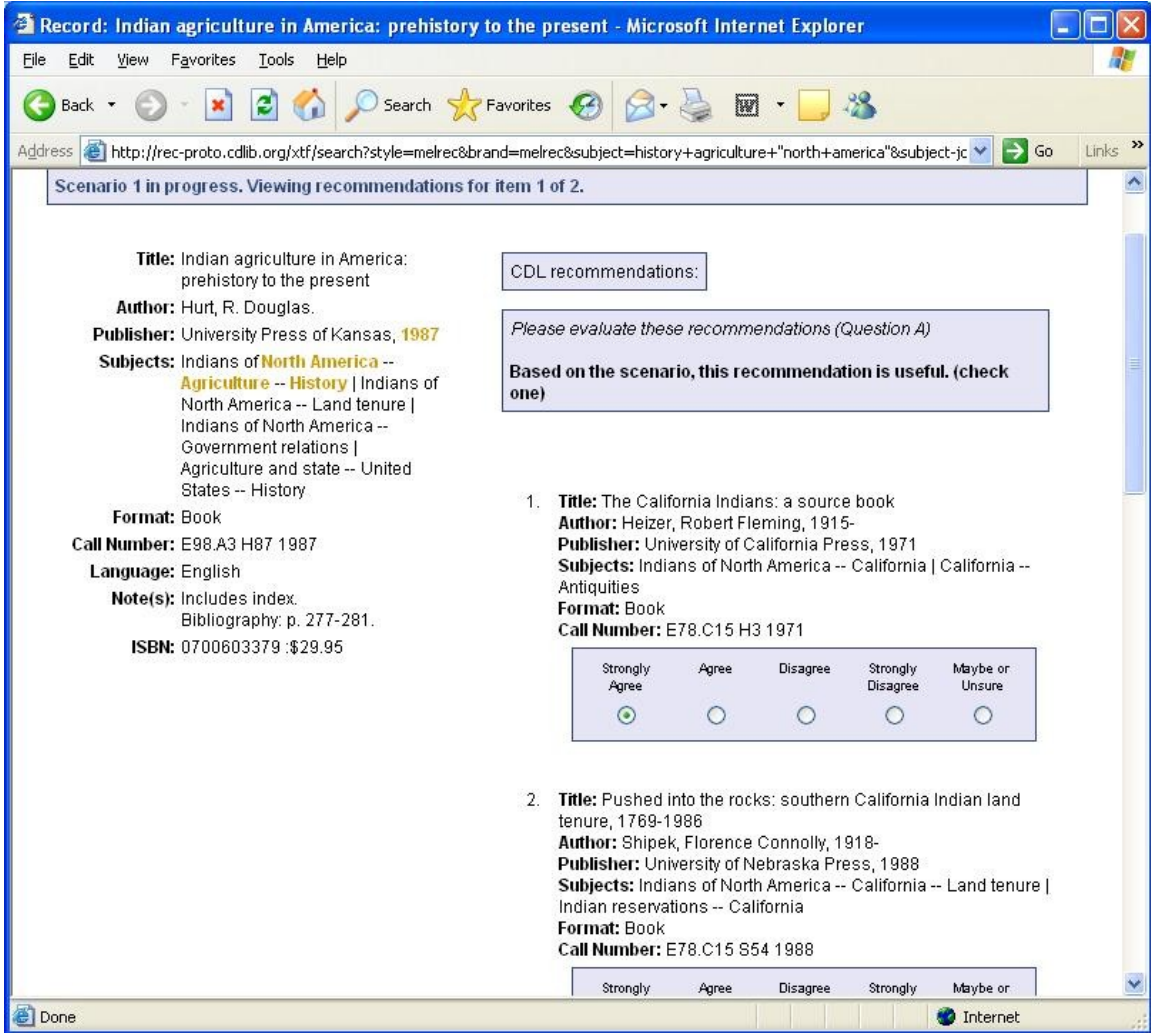
**Figure 3.** After reading the task on the paper scenario strip out loud, users entered keywords into a search box.



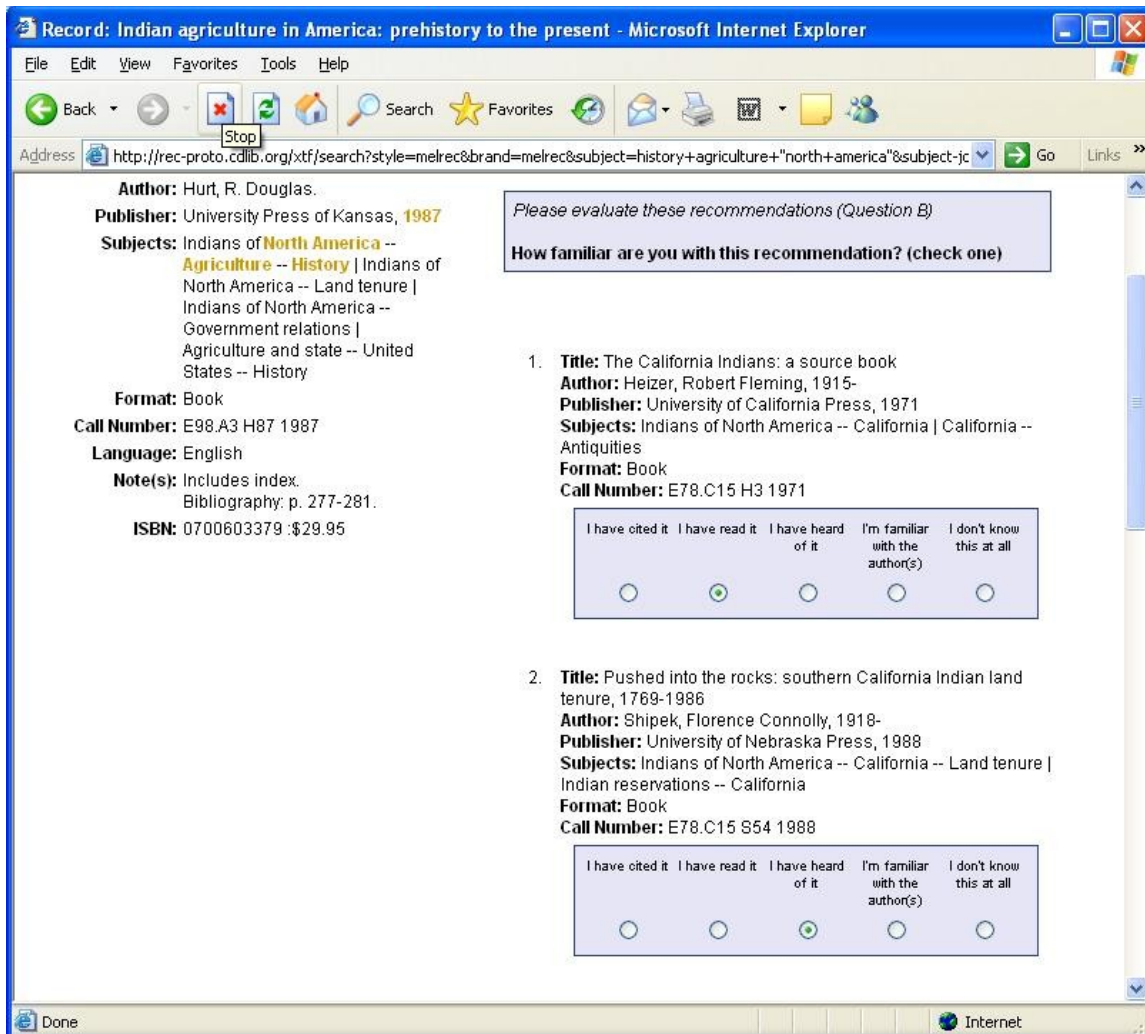
**Figure 4.** As with the relevance ranking assessment, participants were instructed to refine and conduct searches until they obtained a result set that seemed like a good starting point. Once that was accomplished, they were instructed to select up to two “seed” items that they wished to use to generate recommendations, as shown here.



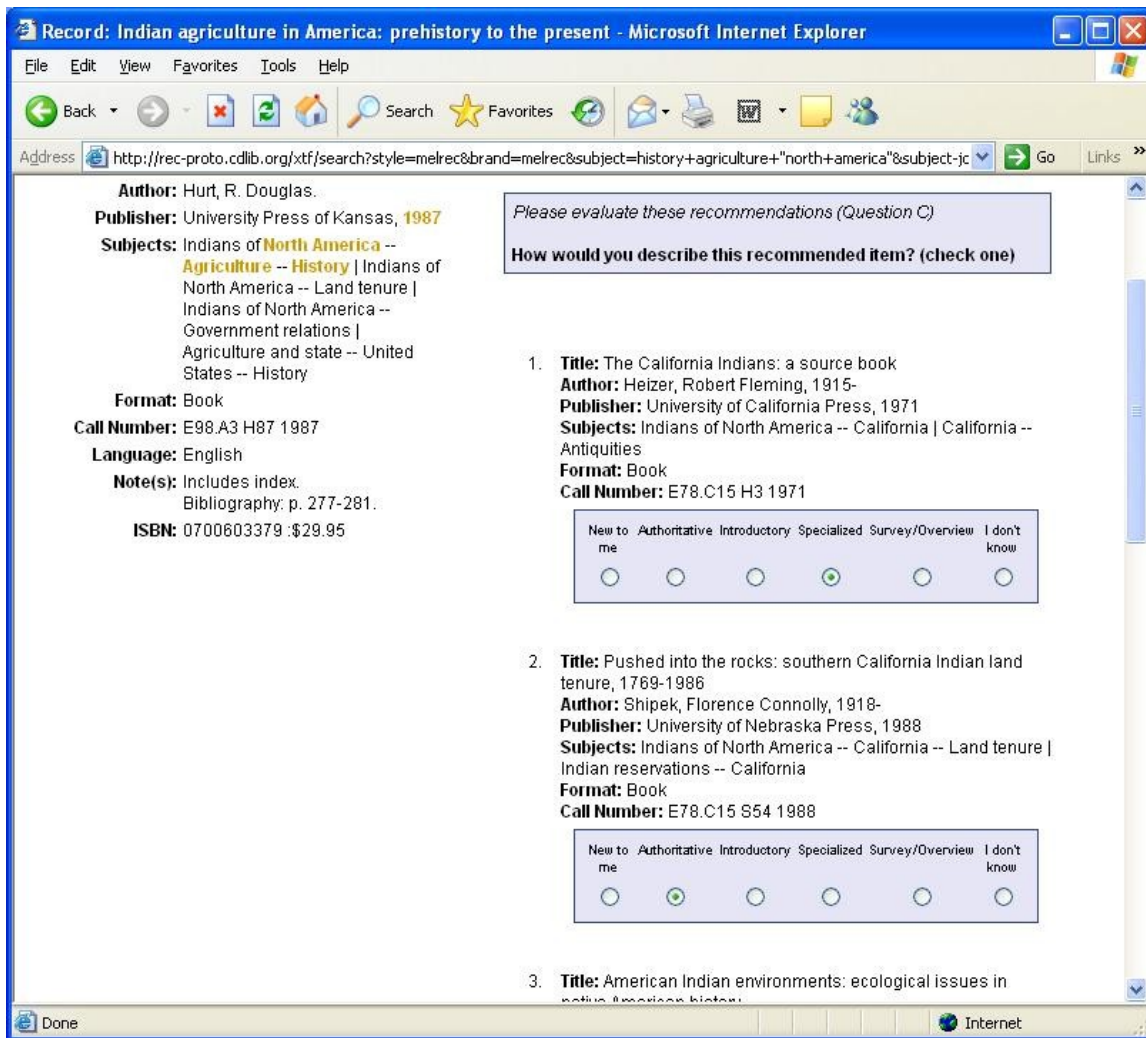
**Figure 5.** After choosing “seed” items, users were shown recommendations sets for each one. Each recommendation set had four questions associated with it. Users evaluated the recommendations associated with one seed before being presented with recommendations associated with the next seed. Shown here is the first evaluation screen; users are asked to express their agreement or disagreement with the statement “Based on the scenario, this item is useful.” and were allowed to choose “Strongly Agree,” “Agree,” “Disagree,” “Strongly Disagree,” or “Maybe/Unsure.”



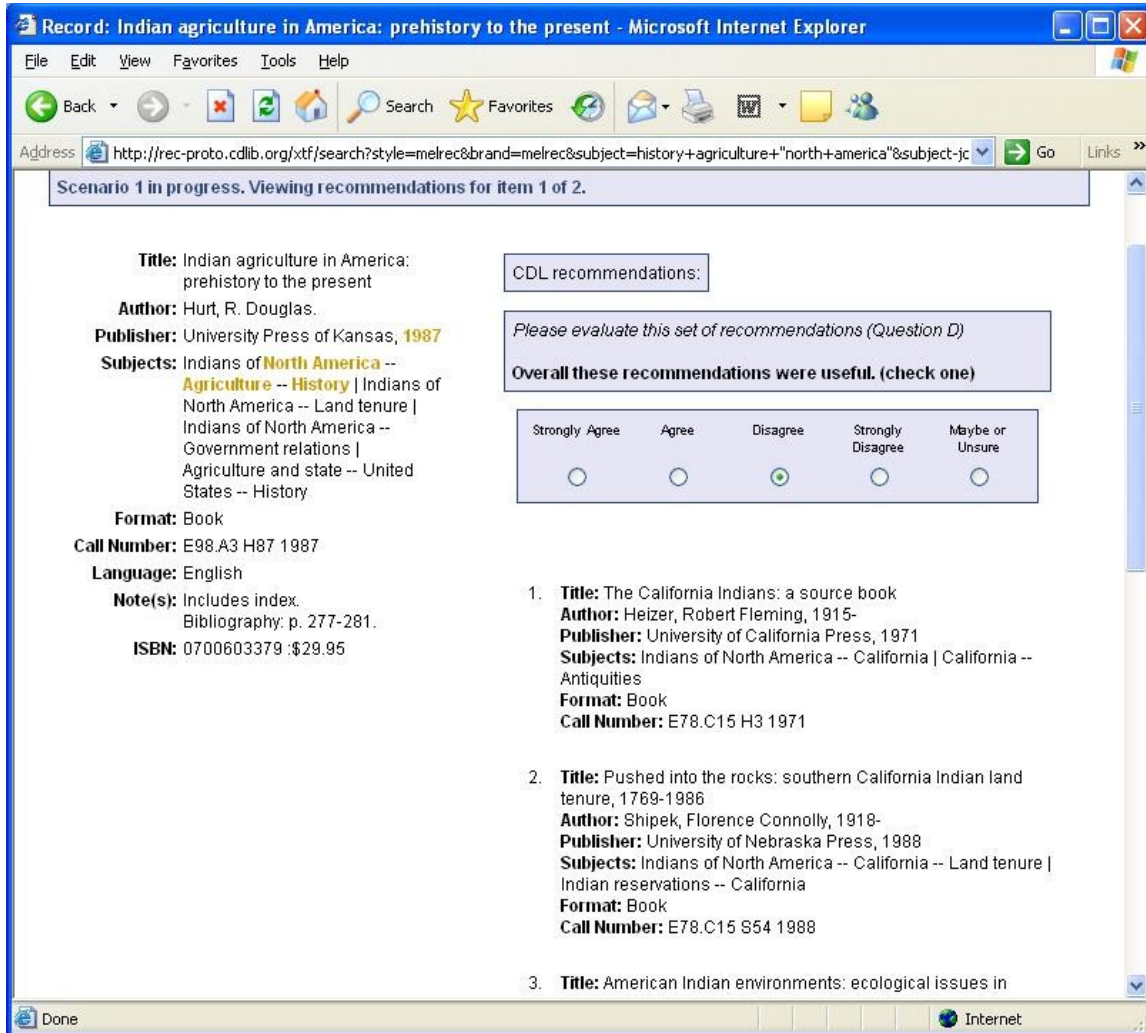
**Figure 6.** The second question, shown here, measured knowledge of the item by asking “How familiar are you with this recommendation?” Users chose from “I have cited it,” “I have read it,” “I have heard of it,” “I’m familiar with the author(s),” “I don’t know this at all.”



**Figure 7.** The third question asked users to assess the nature of the recommended item, regardless of familiarity with the item. Choices ranged from “Authoritative” to “Survey/Overview”, as shown.



**Figure 8.** Users evaluated the quality of the entire set of recommendations as an entire unit, in terms of how useful the set was. As shown here, participants chose the degree to which they agreed or disagreed with the statement “Overall these recommendations were useful.”



**Figure 9.** If the participant had selected more than one “seed” item, they were asked the same set of evaluation questions for the next item, as shown.

