CARMA: Assessing Usability through a Non-biased Online Survey Technique

John D. Hastings & Anatole Mirasano
Dept. of Computer Science & Info. Sys.
University of Nebraska-Kearney
Kearney, NE, USA
{jdhastings, miraanat}@gmail.com

Alexandre V. Latchininsky & Scott P. Schell
Dept. of Renewable Resources
University of Wyoming
Laramie, WY, USA
{latchini, sschell}@uwyo.edu

Abstract

CARMA is an advisory and research support tool for grasshopper infestations. Designed with usability as a primary goal, CARMA presents an interface so intuitive that it completely eliminates the need for a user manual. To achieve this goal, CARMA interacts with the user through a goal-oriented, guided style reminiscent of a natural conversation between an advice seeker and an expert. Usability is furthered by its modeling of four important characteristics of human expert problem solving (speed, graceful degradation, explanations, and opportunism). In order to gain non-biased user feedback about CARMA's interface, we surveyed a group of novice users not previously familiar with CARMA. Positive survey results suggest that CARMA's approach to usability is a success. Furthermore, our survey approach illustrates a simple anonymous online technique which elicits candid non-biased feedback from participants about a product, and is particularly applicable to practitioners short on staff or time.

1. Introduction

CARMA, short for *CA*se-based *R*angeland grasshopper *M*anagement *A*dvisor, is an advisory system for grasshopper infestations that has been successfully used since 1996 [3, 7]. CARMA employs a variety of artificially-intelligent (AI) techniques to provide advice about the most environmentally and economically effective responses to grasshopper infestations. In the process, CARMA illustrates an approach to providing advice concerning the behavior of a complex biological system by leveraging multiple, individually incomplete, knowledge sources [6] including the introduction of a technique known as *approximate-model-based adaptation*[1] which integrates case-based reasoning (CBR) [1, 12] with model-based reasoning for the purpose of prediction within complex physical systems.

[1]Approximate-model-based adaptation is defined and contrasted with perfect-model-based adaptation by Branting [4].

Initially focused on rangeland grasshopper infestations within the state of Wyoming, in 2003 CARMA was expanded to include a prototype cropland grasshopper advising module [8] in order to handle situations when grasshopper populations build up at the rangeland-cropland interface and spread into cropland, such as small grains. In recent years, CARMA's advising capabilities were extended beyond Wyoming in response to interest by several states including Colorado, Montana, Nebraska, New Mexico, North Dakota, Oregon, and South Dakota. The recent extensions saw a conversion of the graphical user interface (GUI) to Java in a manner which illustrates a technique for integrating an artificially-intelligent Lisp reasoner with a Java GUI [9, 13]. The implementation follows a philosophy called *platform freedom* which emphasizes freedom from both platform dependence and software costs, and in the process demonstrates an approach to creating a Lisp application with an appealing GUI which is also web capable.

Initially aimed at non-experts, CARMA was designed with usability as a primary goal with the intention being to present an interface so intuitive that it completely eliminates the need for a user manual. In order to achieve this goal, CARMA interacts with the user through a goal-oriented, guided style reminiscent of a conversation between an advice seeker and an expert. CARMA's usability is furthered by its modeling of four important characteristics of human expert problem solving (speed, graceful degradation, explanations, and opportunism). Recently, in order to gain non-biased user feedback about CARMA's interface, a group of novice users not previously familiar with CARMA were surveyed using a modified form of the desirability toolkit [2]. The survey results suggest that the approach employed in CARMA's interface is a success.

Section 2 describes the relevance of grasshoppers as economic pests and CARMA's role as an advisory tool within that domain including its consultation process, the desirable expert characteristics it emulates, and the usability features incorporated in its design. Section 3 describes the user survey including the survey design which utilizes the modified desirability toolkit approach, followed by the survey results and discussion.

2. CARMA: Grasshopper Infestation Advisor

Grasshoppers (Orthoptera: Acrididae) are economically important agricultural pests in North America. They annually destroy about 25% of the rangeland forage in 17 western U.S. states [11] at an inflation-adjusted cost of US$950 million. Grasshopper infestations frequently reach densities of several hundred per square yard (Figure 1) over extensive areas. During outbreaks, grasshoppers inflict severe damage to rangeland and crops and require large-scale insecticidal control to prevent further economic losses. During an outbreak in the late 1980s, 20 million acres of rangeland in the western U.S. were treated with 1.3 million gallons of insecticides at a total cost of control exceeding US$75 million [16]. Since the 1980s, federal funding for grasshopper pest management on rangeland has dwindled dramatically. At the same time, the cost of insecticides and treatments has skyrocketed. Producers have to assume responsibility for grasshopper control which must be efficient and environmentally sound and yet remain economical on such low-value commodity as rangeland. Hence, there is a challenge to develop sustainable, efficient, economically and environmentally viable grasshopper management strategies. Such strategies would sustain agricultural profits, reduce insecticide and treatment expenses, and minimize adverse environmental impacts.

Figure 1. Hopper band of early instar nymphs of the Clearwing grasshopper *Camnula pellucida***, one of the most important agricultural grasshopper pests in the western U.S. Photo: A. Latchininsky.**

CARMA is a decision-support system designed for end-users (ranchers, farmers, crop consultants and pest managers) which addresses the need for appropriate grasshopper infestation response [13]. In addition to conventional, full-coverage applications of maximum registered dose rates of insecticides, CARMA also considers an option called Reduced Agent and Area Treatments (RAATs) [15] in which insecticide dose rates and coverages are reduced from conventional levels by alternating treated and untreated swaths. This results in lower environmental impacts and treatment costs without compromising efficacy. In 2003, the RAATs strategy was applied to 400,000 acres in eastern Wyoming, saving half a million US dollars for local agriculturists. CARMA has played and continues to play an important role in supporting the development and implementation of sustainable pest management strategies such as RAATs [10].

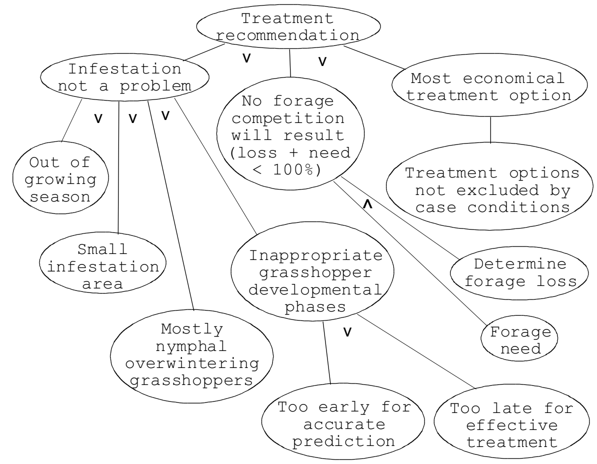

Figure 2. The goal structure that CARMA attempts to satisfy during a consultation.

As is typical of biological systems, the multiple interactions affecting grasshopper population dynamics are too complex and too poorly understood to permit precise prediction through numerical simulation [14]. CARMA is thus modeled after the reasoning approach successfully employed by human experts in the problem domain. A protocol analysis of expert reasoning indicated that a solution should consist of a recommendation supported by an explanation in terms of causal, economic, and pragmatic factors, including, when applicable, a numerical estimate of the proportion of forage destroyed by the pests and a cost-benefit analysis of the available treatment options. Because this advice cannot be produced by any individual reasoning technique, CARMA focused on integrating the multiple problem solving paradigms used by human experts. For example, in order to determine potential forage loss due to grasshoppers, CARMA compares the current infestation

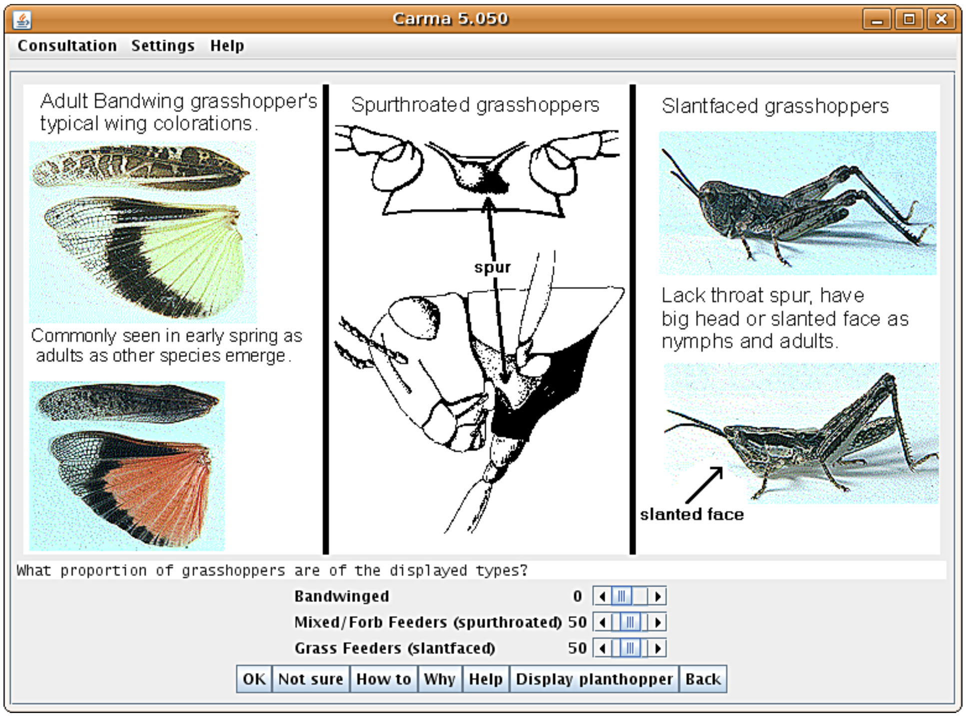

Figure 3. CARMA's grasshopper type elicitation window.

to known previous infestations (i.e., historical cases) and adapts the previous estimates accordingly. In addition, historical infestation probabilities and known treatment efficacies are used to predict re-infestations [5, 17]; statistical methods are applied to predict the economic benefits; and rules are used to select the treatments.

2.1 Consultation process

In being modeled after experts, CARMA interacts with users through the same sort of guided consultation employed by experts. Figure 2 shows the goal-structure that CARMA attempts to satisfy during a consultation, including "treatment recommendation" as the top-level goal. Subgoals separated by $\vee$'s mean that satisfying any of the subgoals results in satisfying the parent goal. Subgoals separated by $\wedge$'s mean that all subgoals must be satisfied in order to satisfy the parent goal. Subgoals are considered in a left to right order. The user is queried for information as needed in order to satisfy goals in the goal structure through interface windows such as that seen in Figure 3.

Briefly, the main steps in a consultation are:

1. Determine the relevant facts of the infestation case from information provided by the user by means of heuristic rules.

2. Predict the proportion of available forage that will be consumed by each distinct grasshopper population using approximate-model-based adaptation.

3. Compare total grasshopper consumption with the proportion of available forage needed by livestock to determine if competition for forage will occur.

4. If the predicted forage consumption will lead to economic loss, determine what possible treatment options are excluded in the current situation.

5. Provide an economic analysis for each viable treatment option and recommend the treatment or treatments that are most economical.

For a detailed description of the rangeland grasshopper infestation advising task and the implementation of the consultation process within CARMA, the reader is referred to Hastings et al. [7].

2.2 Desirable expert characteristics

The protocol analysis used in determining the consultation process identified four important characteristics of human expert problem solving in this field which we hoped to integrate into CARMA's design:

1. Speed: Human experts can provide useful advice quickly, suggesting, consistent with the process description, that human experts can use highly compiled knowledge in the form of prototypical cases.

2. Graceful degradation: Human experts can use, but do not require, highly precise information of the type required for accurate model-based reasoning. Less accurate information may degrade the quality of advice an expert can give but doesn't preclude useful advice. In the worst case, human experts can provide plausible advice based merely on the location of the rangeland and the date.

3. Explanations in terms of a causal model: Although the speed and graceful degradation of human expert performance suggest that experts can use compiled knowledge, they can also readily provide causal explanations for their conclusions. Moreover, entomologists can generate causal predictions of the effects of incremental variations on case facts.

4. Opportunism: Human experts can use a variety of different strategies to solve a single given problem depending on the available information. Human experts don't address the subgoals that arise in decision making in an invariant order but adapt their problem-solving behavior to the particular facts of a given case.

2.3 Emulating Expert Characteristics

CARMA's implementation of the process model emulates the four characteristics of human expert performance mentioned above: speed, opportunism, graceful degradation, and causal explanations. CARMA is fast because, like a human expert, it can use compiled knowledge in the form of rules and cases, rather than depending entirely on computation-intensive simulations. CARMA is opportunistic in that it can recognize when no additional information is required from the user (e.g., when it is too early in the season to provide an accurate prediction) and thus does not require the user to respond to a fixed set of facts when some of those facts will be irrelevant.

CARMA achieves graceful degradation in two ways. First, CARMA utilizes multiple levels rules ordered by reliability to infer case features. Thus, accurate information can be used if available, but the absence or inexactness of the information does not cause a drastic fall-off in accuracy. Second, the use of CBR by CARMA means that incrementally less accurate information will produce incrementally less precise matching and adaptation, but not a catastrophic inability to produce plausible predictions and advice.

Finally, CARMA is capable of providing causal explanations, notwithstanding its use of CBR, based on a knowledge of what constitutes variations from normal rangeland conditions. CARMA orders features based on their importance to forage consumption and the magnitude by which they vary from normal. An explanation is generated by describing the features that most aggravate and mitigate forage loss using standard template-instantiation techniques.

These desirable expert traits provide a positive contribution to usability as detailed in the following subsection.

2.4 Usability

CARMA's initial audience was expected to include a substantial proportion of users with little to no computer experience running computers with limited capabilities. CARMA was thus designed with user friendliness as a primary goal with the intention being to completely eliminate the need for a user manual – no user manual is available either bundled or separate. Early feedback from representative users influenced CARMA's design in order to achieve that goal. The initial software distribution included only the following instructions:

> Because our goal is to provide a piece of software that is so intuitive and accessible that it can be used with this following brief set of instructions, there is no cumbersome and alienating "User's Manual". To begin your first consultation, simply click on "Consultation" and then "New" and answer the questions as they appear. The other buttons are intended to be self-explanatory and intuitive.

In order to achieve this goal, CARMA's interface employs a guided style reminiscent of a natural conversation between an advice seeker and an expert. In this context, a consultation follows the sort of conversation a rancher might have with an extension entomologist who attempts to determine the appropriate response to a grasshopper outbreak. The initial implied question by the rancher is "I have grasshoppers, what should I do?". The consultation proceeds with questions which depend on the rancher's answers. If the rancher is unsure of how to respond, CARMA can assist the rancher in understanding the domain to the extent necessary to provide useful information. This conversation style along with the appropriate guidance contributes to CARMA's usability.

Figure 3 shows a typical CARMA input window which illustrates several elements of CARMA's usability. First and foremost is the question presented to the user, in this case "What proportion of grasshoppers are of the displayed types?". Input windows display visual aids when appropriate, for example, when the user is performing identification and a graphic can aid or guide that task. The user is rarely required to type, unless doing something more advanced like modifying/adding treatment options, and instead typically uses components such as sliders, radio buttons and push buttons. CARMA input windows present push buttons in a consistent order (so that the buttons are easily locatable and the user quickly becomes familiar with their placement) and are chosen from a common set of possible buttons. Most windows display an identical set of buttons so that the user is not learning new buttons on each screen unless a particular button is not applicable to a certain window. The set of common push buttons, each with a descriptive name, includes:

1. "OK": accept the user's input,
2. "Not sure": jump to the next window for the current input field with an alternate or easier set of inputs,
3. "How to": provide instructions for gathering the necessary inputs,
4. "Why": describe the importance of the input to the consultation,
5. "Help": provide assistance to the user in navigating the interface, and
6. "Back": return to a previous input window.

In rare cases, an input window will present a specialized button, for example, the window in Figure 3 also includes a button which displays a planthopper, a beneficial species which the "How to" window indicates shouldn't be confused with a grasshopper.

In our experience, users quickly become proficient with the system, and after time are able to breeze through a consultation by setting the inputs for each window and clicking OK. However, should the user have trouble, they can click on the other buttons for guidance.

In addition to the intuitive user interface presented in the form of a guided conversation, the human characteristics modeled within CARMA produce a positive contribution to usability:

1. Speed: The speed with which CARMA allows a user to complete a consultation allows the user to react more quickly to an infestation, and run multiple consultations if necessary. The speed of a consultation also makes CARMA an appealing alternative to phone support for some users.
2. Graceful degradation: The ability to provide useful advice even in the absence of perfect data allows CARMA to support a variety of users at different levels of experience. It also allows a novice the ability to fine tune their information if CARMA suggests that an infestation is of borderline concern. For example, if a user just returned to the computer from the field and has difficulty providing exact grasshopper numbers, they can quickly run through a consultation using approximations, receive useful approximate advice, and then return to the field to obtain more precise counts if necessary.
3. Explanations: By using a reasoning approach based on experience rather than a black box approach (e.g., neural nets), CARMA's transparent reasoning approach supports explanations which allow the user to learn about the reasoning behind CARMA's advice. Such transparency is important to land managers in gaining broad-based support for management decisions. In addition, such explanations serve an educational role by indicating to the user factors that might lead to grasshopper problems in the future (e.g., "Hot, dry conditions can exacerbate the effects of a grasshopper outbreak.").
4. Opportunism: When possible CARMA can end a consultation early by providing advice based on a smaller set of system inputs. This further contributes to consultation speed, and prevents the user from being burdened by filling in a larger unnecessary fixed set of inputs.

Our interactions with users guided the incorporation of usability features within CARMA and subsequent interactions indicated that the design is successful in that respect.

3 User Survey

In order to gain non-biased feedback about the quality of CARMA's interface, a group of novice users not previously familiar with CARMA were surveyed. The survey design and results are described in the next two subsections, followed by a discussion.

3.1 Design

This survey was designed to gauge general user reaction to CARMA's interface. Because the survey was focused more on the usability of the interface, and less on the actual usefulness or accuracy of the product and its fit within the world of sustainable grasshopper management (which represent surveys in and of themselves), we did not seek out users who had prior experience with similar systems or who had any particular domain knowledge, either specific (e.g., grasshoppers or grasshopper management) or general (e.g., agriculture or other advising systems). We thus selected our participants to be the general student population

Introduction

Thank you for taking the time to answer this survey. This survey should take approximately 5-10 minutes to complete, and you must be at least 19 years of age to participate. This survey is completely anonymous.

Instructions

CARMA is a decision-support system which provides advice to users about grasshopper outbreaks. This anonymous survey will ask you to run CARMA for a brief amount of time, and then complete a very short survey. Prior experience with the problem domain is not necessary -- this survey focuses on the software. After launching CARMA, click "Consultation" and "New", and follow the instructions. Use/navigate the software for five to ten minutes. After using the software, please return to complete the survey.

1. Click on the link below to launch CARMA.

   Launch CARMA

2. After using the software, please return to complete the survey by clicking the submit button. The survey has only two parts.

   Submit (must be 19)

Figure 4. Online survey: introduction.

Survey Part 1

Pick five words from the following group which best describe your experience with CARMA, then click submit.

☐empowering	☐sophisticated	☐unconventional	☐not valuable	☐entertaining	☐ambiguous	☐counter-intuitive	☐effective
☐simplistic	☐cutting edge	☐connected	☐relevant	☐difficult	☐clear	☐old	☐gets in the way
☐complex	☐accessible	☐fresh	☐unattractive	☐predictable	☐ineffective	☐understandable	☐compelling
☐new	☐flexible	☐boring	☐high quality	☐dated	☐stressful	☐time-saving	☐ordinary
☐friendly	☐convenient	☐obscure	☐contradictory	☐faulty	☐attractive	☐consistent	☐desirable
☐responsive	☐system-oriented	☐uncontrollable	☐busy	☐non-standard	☐insecure	☐rigid	☐satisfying
☐customizable	☐overbearing	☐poor quality	☐unrefined	☐controllable	☐easy to use	☐fast	☐straightforward
☐hard to use	☐slow	☐educational	☐patronizing	☐organized	☐intuitive	☐intimidating	☐professional
☐overwhelming	☐efficient	☐creative	☐dull	☐engaging	☐usable	☐annoying	☐energetic
☐credible	☐exciting	☐familiar	☐clean	☐simple	☐meaningful	☐business-like	☐collaborative
☐powerful	☐bright	☐impressive	☐secure	☐frustrating	☐inconsistent	☐too technical	☐advanced
☐effortless	☐confusing	☐reliable	☐useful	☐appealing	☐personal	☐inadequate	☐motivating
☐misleading	☐innovative	☐stable	☐distracting	☐unpredictable	☐incomprehensible	☐fun	☐approachable
☐awkward	☐time-consuming	☐comprehensive	☐irrelevant	☐inviting	☐stimulating	☐trustworthy	☐expected
☐cluttered	☐uninstructive	☐illogical	☐vague				

Submit

Figure 5. Online survey: term selection.

at the University of Nebraska at Kearney (UNK) and distributed the survey en masse via email to these students during the week of final exams, May 2009. UNK does not have an agricultural college, so there was a low probability that the students had used CARMA previously. Each of the participants should have been computer literate and would have computer skills on par with first-time CARMA users[2]. The students were enticed to complete the survey with the chance to win a free pizza. The survey was entirely anonymous in nature, partly in keeping with university protocol, but also in order to solicit candid feedback.

The survey method itself was a modified version of the desirability toolkit. As originally described by Benedek and Miner [2], this approach asks the user to complete a usability test for a product, then pick the five cards from a group of 118 "product reaction cards" that best match their reaction to the system. The selected cards then become the basis of a guided interview aimed at soliciting feedback.

The two greatest advantages of the desirability toolkit as opposed to a more conventional usability questionnaire, and the reasons that we selected the technique for this survey, are 1) it aims to avoid a bias toward the positive found in typical questionnaires (e.g., it has been found that if a respondent thinks that a survey intends to assess the quality of a product, they are likely to provide more positive answers about quality) and 2) it is able to more effectively uncover constructive negative criticisms in the guided interview.

We presented the survey to respondents online. The introductory page for the survey shown in Figure 4 asked the participant to briefly use CARMA (which ran through Java Web Start) and then complete the survey. In order to avoid bias, the respondents were presented with precious little information about the intent of the survey. Our goal was for each participant to use CARMA's interface just long enough to complete the survey based on their reaction to the interface. CARMA runs fairly quickly, so we felt that five to ten minutes of use would give the participant a sufficient feel

[2]Although survey participants were potentially more or less computer literate than the population of typical users, CARMA includes "How to" and "Help" features to guide even the most technologically illiterate, so such skill differences should produce a negligible effect on the survey result.

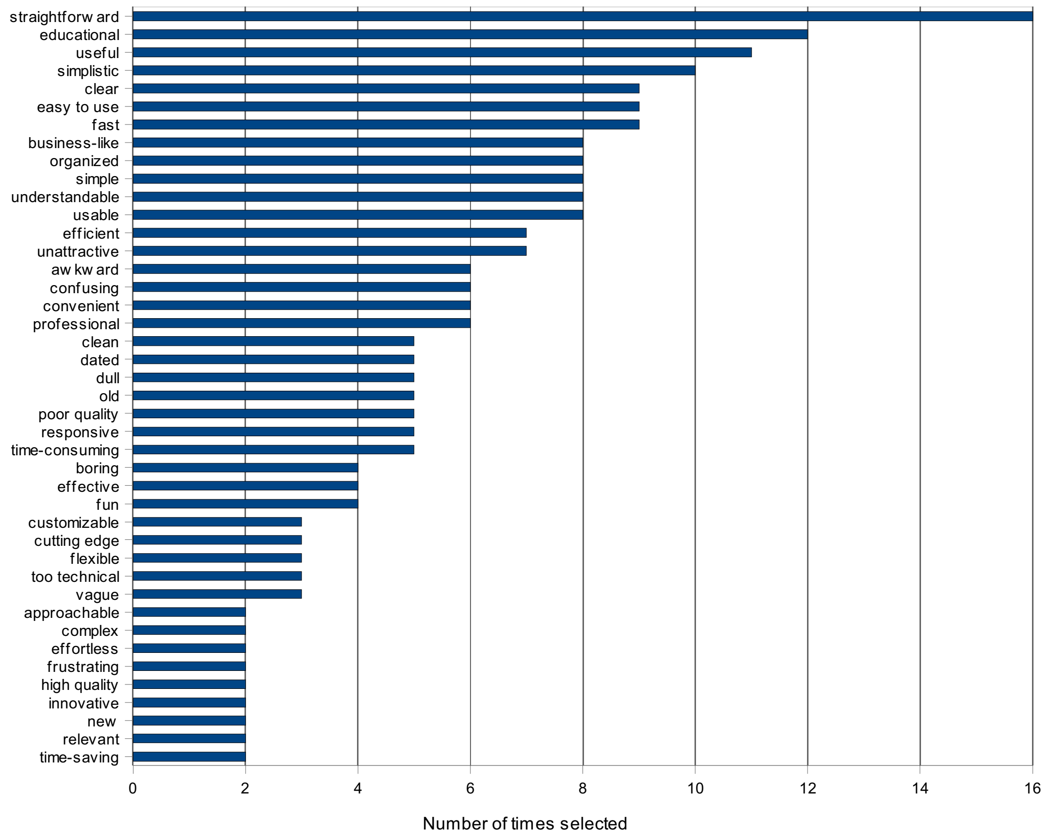

Figure 6. Survey results: number of times each term was selected.

for the interface, and would be less likely to cause busy participants to leave the survey early with an impression that the survey would be too time consuming.

Following their use of the software, on the next page of the survey (Fig. 5), the participant was asked to pick from a set of 116 terms the five terms which best described their experience with CARMA. The term $educational$, not part of the original desirability toolkit, was included because one of CARMA's goals is to be educational. The antonym $uninstructive$ was included so as to not unfairly bias the term list in a positive direction.

The list of terms was presented in a random order each time the page was loaded in order to lessen the influence of order on the selection of any particular terms. In addition, the participant was not given advance warning as to how a selected term would be used or that a followup comment would be required in an attempt to reduce the affect of bias on the selection of any term.

Because the survey was anonymous, rather than conducting a face-to-face followup interview session with each participant to discuss their selection of each term (as described in the desirability toolkit), we asked for this feedback online. On the final page of the survey (not shown), the participant was asked to provide a comment explaining their selection of each of the five terms.

3.2 Results

The survey received 89 responses. Initial analysis of the comments revealed that 37 respondents (41% of the total responses) did not provide "valid" responses. For example, a certain percentage of participants were not serious about the survey and filled in nonsense answers in order to enter the drawing for the pizza (e.g., comments such as "zscfsd sdds..."). Another group of participants were confused about the survey (perhaps partly because we did not wish to describe its full intent) and rather than provide responses about CARMA's interface through term selection and comments, chose to provide commentary on the domain (i.e., grasshopper management) or the applicability of the domain to themselves, e.g., some participants selected the term $irrelevant$ because "i don't see what i have to do with grasshopers" or "I don't live on a farm/ranch and I don't have a grasshopper infestation.".

Figure 7. Survey results: word cloud.

Before compiling the final results, we filtered out responses with comments that did not focus on CARMA's interface. This filter left 52 "valid" responses (with five terms each) which are described below. The selections made by these 52 respondents are shown in Figure 6. The word cloud, an alternate result representation common in desirability studies, seen in Figure 7, shows terms with a font based in proportion to the frequency of selection. Note that terms selected only once (of which there were 25) are not displayed in Figures 6 and 7.

The top four terms chosen by participants to describe CARMA's interface were *straightforward*, *educational*, *useful*, and *simplistic*. The comments for *straightforward* included "Presents information in an understandable and effective manner." and "The instructions were very clear.". Comments for *educational* included "I learned things I never knew about grasshoppers." and "Learned something about grasshopper levels". The top negative term was *unattractive* with comments such as "program seemed dated" and "I really do not like the font chosen".

As a means to quantitatively summarize the user experience, we classified the individual responses by the participants into three categories based on the terms used and the survey context: positive (e.g., *straightforward*), negative (e.g., *complex*), and neutral (e.g., *simplistic*). As seen in Fig. 8, according to this measure, CARMA's interface was described as positive in 68% of the responses, negative in 25%, with the remainder as neutral.

The participants' comments provide an alternate means to rank the user experience into positive and negative categories. In most cases, comments were either completely positive or negative, but occasionally, individual comments included both positive and negative components and thus were more neutral in nature, e.g., one selection of *confusing* included the comment "I was not sure exactly what I was supposed to do, but once I clicked the how to, everything was much clearer." Fig. 8 shows the categorization of the responses based on the comments and indicates that CARMA's interface was described positively in 73% of the responses, negative in 25%, with 2% neutral.

CARMA was ranked more positively on the basis of comments than when evaluated via terms because the use of terms categorized as neutral were mostly accompanied by comments descriptive of a positive experience. For example, the comments for *simplistic*, which could be perceived as a positive or a negative depending on context, show that the term was generally viewed as positive, e.g., "easy to use. not a lot of buttons" and could mostly be combined with the term "simple", although two of the comments were negative in nature, e.g., "Not very engaging".

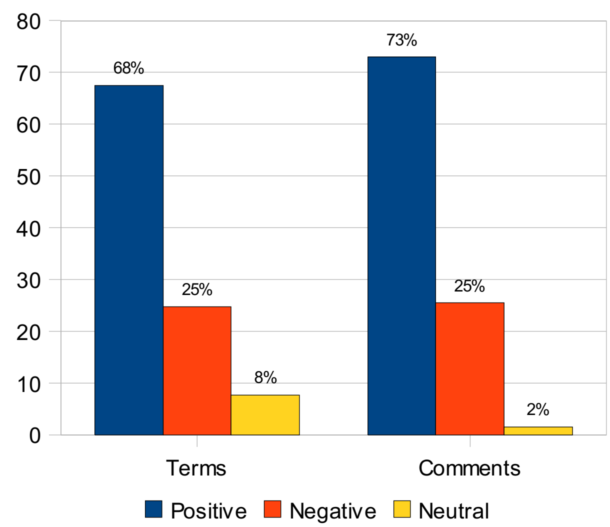

Figure 8. Survey results: categorization of responses based on terms and comments.

3.3 Discussion

Our survey did not receive as many responses as we had hoped, although due to timing issues it was distributed during final exam week which surely cut down on the availability and interest of the student participants. Despite the smaller number of participants, the modified desirability survey toolkit enabled the acquisition of a perfectly usable, reliable and unbiased response.

In analyzing the survey results, it is obvious that CARMA's interface is perceived in a positive light by the majority of respondents. The top twelve selected terms were positive in nature (either by definition, or by reviewing the comments). The comments were also exceedingly positive.

A smaller percentage of negative terms were selected, although the explanations revealed that their choice was generally tied to the "old-fashioned" look of the interface and not the usability or the performance of the software. However, some users had troubles with the interface, e.g., "I was confused as to how to get the percentages right until I figured it out". The comments about the "old-fashioned" look of the software are somewhat surprising given that Java is used for the interface, but perhaps the comments are somehow influenced (consciously or subconsciously) by the initial CARMA splash screen, which shows a western-style cartoon drawing of a grasshopper wearing a cowboy hat and boots.

The comments in some cases will prove helpful in improving CARMA's look. In addition, the comments were valuable if not crucial in determining the context of a chosen term (e.g., $simplistic$) and categorizing the survey responses. While it would be time prohibitive to wade through comments in the event of a larger number of responses, such analysis was easily manageable in this case.

In terms of the applicability of the survey method, the modified desirability toolkit method that we employed should be generally applicable to product evaluation where higher-level impressions of that product are desired. A survey using this approach is quick and easy to construct and distribute, and is thus applicable in situations where time and staff are in short supply.

Although the desirability toolkit is intended to be qualitative in nature, in our opinion, with a categorization of the terms and comments (into positive and negative), a quantitative score can be produced. Generating a score from the terms is fairly quick and represents a good first cut. Given that the selection of a term can vary depending on the context, if time permits, an analysis of the comments is necessary for a more precise product score.

This survey approach does seem to elicit candid non-biased feedback. In our opinion, care must be taken to avoid bias in all portions of the survey, e.g., we made the survey anonymous and online because even interacting with a participant could potentially bias the result. We further attempted to avoid bias by giving a scant set of instructions. In a sense, by providing minimal instructions (from which it would be harder to determine the purpose of the survey) and by splitting the online survey into separate pages, we tricked the survey participants into providing feedback without informing them as to the meaning of their choices. That is not to say that the participant could not figure out that the survey was about software or evaluating the software in some way, but at least they were not completely sure of the exact meaning of the term choices until the point that they were asked to explain their choice in comments.

However, it's nearly impossible to remove all bias. This online approach could be prone to a type of participant selection bias called voluntary response bias in which participants are self selected. In this case it might be argued that an individual who agrees to participate in an online survey could bias the survey with a sort of preselection of individuals who are more comfortable interacting online or with software in general and who would thus find software of any sort easier to use.

Ideally, our survey would have received more responses and this limitation tempers the associated confidence of the quantitative results. However, although the sample size was relatively small, a very positive reaction was produced, valuable and candid comments were obtained, and a pattern developed (e.g., 30% of the participants selected $straightforward$ even though they had 116 words from which to choose). Future work might involve expanding the sample size, or surveying users with some familiarity of the problem domain in order to obtain comments from a different perspective.

In summary, the survey indicates that users have a positive impression of CARMA's interface providing confirmation of the approach we employed in incorporating positive usability traits in CARMA, including the goal-oriented guided style, and the emulation of desirable expert characteristics. Furthermore, our survey method demonstrates a simple anonymous online survey approach which can elicit candid non-biased feedback from participants about a product and should be generally applicable to a wide range of products. The approach is a nearly effortless means of conducting a survey, and should be especially relevant to practitioners who are short on the staff or time needed to design and distribute a survey.

4 Availability and status

Since its inception in 1996, CARMA has been presented to pest managers in all 17 western states in which grasshoppers present economic problems. The most recent version of CARMA, 5.050, with rangeland grasshopper advising capabilities for Colorado, Montana, Nebraska, New Mexico, North Dakota, Oregon, South Dakota and Wyoming is available free of charge for noncommercial purposes and can be downloaded and installed from http://carma.unk.edu or run as a Java Web Start application. CARMA is currently being extended to Idaho and Utah.

5 Conclusion

CARMA is an advisory system for grasshopper infestations that has been successfully used since 1996. CARMA was designed with usability as a primary goal with the intention being to present an interface so intuitive that it completely eliminates the need for a user manual. In order to achieve this goal, CARMA interacts with the user through a goal-oriented, guided style reminiscent of a conversation between an advice seeker and an expert. CARMA's usability is furthered by its modeling of four important characteristics of human expert problem solving (speed, graceful degradation, explanations, and opportunism). Recently, in order to gain non-biased user feedback about CARMA's interface, a group of novice users not previously familiar with CARMA were surveyed using a modified form of the desirability toolkit. Our simple anonymous online survey approach can elicit candid non-biased feedback from participants about a product, and is particularly applicable to practitioners who are short on the staff or time needed to design and distribute a survey. The positive survey results suggest that the approach employed in designing CARMA's interface, including its emulation of desirable expert characteristics, is a success.

6 Acknowledgments

CARMA's development since 2003 was supported through funds from Cooperative Agreements between USDA-APHIS-PPQ (Western Region) and the University of Wyoming (grants USDAAPHIS5112, USDAAPH44906, USDAAPH44909GHS and USDAAPH44913GHM).

References

[1] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1):39–59, 1994.

[2] J. Benedek and T. Miner. Measuring desirability: New methods for evaluating desirability in a usability lab setting. In *Proceedings of the Usability Professionals' Association Conference*, Orlando, FL, USA, July 2002.

[3] K. Branting, J. Hastings, and J. Lockwood. CARMA: A case-based range management advisor. In *Proceedings of The Thirteenth Innovative Applications of Artificial Intelligence Conference (IAAI-01)*, pages 3–10, Seattle, Washington, August 2001. AAAI Press.

[4] L. K. Branting. Integrating cases and models through approximate-model-based adaptation. In *Proceedings of the AAAI 1998 Spring Symposium on Multimodal Reasoning (SS-98-04)*, pages 1–5, Menlo Park, CA, USA, March 1998. AAAI Press.

[5] L. K. Branting, J. D. Hastings, and J. A. Lockwood. Integrating cases and models for prediction in biological systems. *AI Applications*, 11(1):29–48, 1997.

[6] J. Hastings, K. Branting, and J. Lockwood. A multi-paradigm reasoning system for rangeland management. *Computers and Electronics in Agriculture*, 16(1):47–67, 1996.

[7] J. Hastings, K. Branting, and J. Lockwood. CARMA: A case-based rangeland management adviser. *AI Magazine*, 23(2):49–62, 2002.

[8] J. Hastings, K. Branting, J. Lockwood, and S. Schell. CARMA+: A general architecture for pest management. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03) Workshop on Environmental Decision Support Systems (EDSS-03)*, pages 18–21, Acapulco, Mexico, August 2003.

[9] J. D. Hastings and A. V. Latchininsky. CARMA: Platform freedom for a graphical lisp application through armed bear common lisp. In *Proceedings of the International Lisp Conference (ILC-09)*, Cambridge, MA, USA, March 2009.

[10] J. D. Hastings, A. V. Latchininsky, and S. P. Schell. Sustainability of grasshopper management and support through CARMA. In *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS-42)*, pages 10 pages, CDROM, Los Alamitos, CA, USA, January 2009. IEEE Computer Society.

[11] G. Hewitt and J. Onsager. Control of grasshoppers on rangeland in the United States: a perspective. *Journal of Range Management*, 36(2):202–207, 1983.

[12] J. Kolodner. *Case-based reasoning*. Morgan Kaufmann, San Mateo, CA, USA, 1993.

[13] A. V. Latchininsky, J. D. Hastings, and S. S. Schell. Good CARMA for the high plains. In *Proceedings of the 2007 Americas' Conference on Information Systems (AMCIS 2007)*, Keystone, CO, USA, August 9-12 2007.

[14] J. A. Lockwood and D. Lockwood. Rangeland grasshopper (Orthoptera: Acrididae) population dynamics: insights from catastrophe theory. *Entomological Society of America*, 20(4):970–980, 1991.

[15] J. A. Lockwood and S. P. Schell. Decreasing economic and environmental costs through reduced area and agent insecticide treatments (RAATs) for the control of rangeland grasshoppers: Empirical results and their implications for pest management. *Journal of Orthoptera Research*, 6:19–32, 1997.

[16] National Grasshopper Management Board. In *Proceedings and Resolutions of the 1995 meeting of the National Grasshopper Management Board*, Rapid City, SD, USA, January 1995.

[17] K. M. Zimmerman, J. A. Lockwood, and A. V. Latchininsky. A spatial, Markovian model of rangeland grasshopper (Orthoptera: Acrididae) population dynamics: Do long-term benefits justify suppression of infestations? *Environmental Entomology*, 33(2):257–266, 2004.