A heuristic evaluation instrument for e-government online software

Jerald Hughes*, Punit Ahluwalia and Vishal Midha

Computer Information Systems,
College of Business Administration,
University of Texas – Pan American,
1201 W. University Drive,
Edinburg, 78539 TX, USA
E-mail: jhughes1@utpa.edu
E-mail: punit@utpa.edu
E-mail: vmidha@utpa.edu
*Corresponding author

Abstract: An important component of e-government maturity is online software tools for financial analysis directly on the website. This paper develops a novel extension of website heuristic evaluation methodology, contributing a set of guidelines which encompasses not only website content, but also online software quality, based on established software evaluation concerns. Interactive online retirement and tax planning software tools are used as exemplars to demonstrate the evaluation methodology. Online software tools from five nations in four regions of the globe are evaluated for three different kinds of tasks. Results indicate that such tools have considerable room for improvement.

Keywords: electronic government; e-government usability; heuristic evaluation; online software; online calculators; interactive software.


Biographical notes: Jerald Hughes is an Assistant Professor in the Computer Information Systems Department at the University of Texas – Pan American. He holds a PhD in Computer Information Systems from Baruch College of the City University of New York. His teaching and research interests include digital information goods, consumer networks and virtual communities, digital multimedia and information security.

Punit Ahluwalia currently is an Assistant Professor of Computer Information Systems at the University of Texas – Pan American. He received a PhD in CIS from Georgia State University in 2006. He also has Graduate degrees in CIS from Georgia State University and in Management and Systems from IIT-Delhi. His research interests include ‘quality of service’ in wireless networks, mobile transactions, quality of information products and services, understanding of acceptance of technology through alternative theoretical models and changing landscape of IT/IS users. He has authored many papers in several prestigious journals/conferences. He has also received much recognition for his teaching.

Copyright © 2013 Inderscience Enterprises Ltd.
Vishal Midha is an Assistant Professor of Computer Information Systems at the University of Texas – Pan American. He has a PhD in Information Systems from the University of North Carolina at Greensboro. His current other research interests include open source software development, information piracy concerns and knowledge management. He has published in Communications of AIS, Electronic Markets, Journal of CIS and many national and international conferences including ICIS, AMCIS, DSI and GITMA. Presently, he also serves as an Associate Editor for Case Studies in the International Journal of Information Security and Privacy.

1 Introduction

As internet usage continues to increase, e-government websites are undertaking to offer improved performance in meeting citizens’ needs. E-government websites now serve a variety of purposes. These include simple information listed online, downloadable application forms, online requests for government records, contact with government entities, transactions such as license renewal and payment of government bills or fees due (Cook, 2000; Norris and Moon, 2005). Previous studies have categorised the offerings of e-government by the kind of service offered; for example, Fang (2002) identifies three major categories:

- **Information**: Information requests of a firm or a citizen regarding taxes, business licenses, registers, laws, political programmes, administrative responsibilities, etc.

- **Communication online**: Information requests and discussions regarding administrative processes and products and communication with politicians and authorities, etc.

- **Transaction**: Online delivery of service and posting of results, electronic voting, providing solutions online, participation online, etc.

Fang’s (2002) categories, while critical for many e-government services, do not capture the IT artefact studied here: E-government Online Software (EOS) tools for decision support and analysis directly on the website. These pages provide input/control interfaces to software used for various numerical analyses. Such pages help users do financial planning for taxes, pensions, retirement needs and related financial topics. The results of such analyses help users project their future income, needs and financial obligations under various conditions. EOS tools provided by the government are likely to provide many benefits, such as a higher level of trust, cost savings and greater awareness of financial position and considerations. The importance of effective diffusion of such services is even more pronounced considering the impact of the digital divide on access to e-government services (Brown, 2002; Sipior and Ward, 2005; Belanger and Carter, 2006) and because the less privileged actually need these services much more. Ndou (2004) identified improving the quality of citizens’ decision-making as one of seven e-government opportunities.

The principal motivation for this study is the lack of established methods suitable for evaluating EOS. The goal of our research is, thus, to improve e-government services by studying EOS tools. These tools, like the rest of e-government, need to be accessible, usable, effective and adopted by citizens. Poorly designed e-government services will
A heuristic evaluation instrument for e-government online software

generate unpleasant experiences in users and lead to increased resistance to adoption (Weerakkody and Choudrie, 2005; Wangpipatwong et al., 2008). Therefore, their full potential may not be met (West, 2002). If these criteria are met, then they can serve a valuable purpose, giving citizens better control over their relationships with government (Reffat, 2003). High quality EOS can help citizens by making the government’s impact on their money more transparent and improving the quality of financial planning. To the best of our knowledge, no metric is available in the literature that specifically evaluates EOS tools.

As users encounter these EOS tools, it is important to have well-founded ways of evaluating them (Mofleh and Wanous, 2009; Chircu, 2008). E-government websites should, for example, maximise usability for the widest possible range of potential users (Schneiderman, 1999). Ergonomic quality should also be taken into account (Yucel and Ozok, 2010). While some principles of evaluating websites generally are appropriate, such as layout and readability (Becker and Nowak, 2003), other evaluation methods intended specifically for e-government websites also apply. For example, the users of e-government may be persons of moderate levels of computer efficacy, people who need to access the internet through shared computers (e.g., workstations in public libraries or community centres), or people who have limited mathematical, analytical and legal skills.

The gap in research which this study addresses is lack of established methods suitable for evaluating EOS. This paper constitutes the first effort at such a method. While e-government services have been studied in other contexts (e.g., Gotoh, 2009), none to date have integrated evaluation of EOS tools with previously established methods of evaluating e-government websites. Nor have previous studies of software evaluation taken into account the particular challenging problems of e-government provision of services. Our study undertakes to fill this gap.

This study is built on two principles. First, we adopt the point of view that evaluation, specifically of EOS tools, must address existing known issues for e-government website services. Second, we take the approach that evaluation of EOS tools should also incorporate established principles for evaluating software from the users’ points of view. This study develops an instrument by integrating both principles.

To be as useful as possible, a set of guidelines for heuristic evaluation for EOS needs to be applicable in a wide variety of contexts: differing cultures, differing regions and differing tasks. The guidelines developed in this paper thus endeavour to cover aspects considered important in a wide variety of contexts, largely independent of the particular governments or cultures involved. This study demonstrates the effectiveness of the developed heuristic evaluation guidelines by applying them to EOS tools from five different nations in four different regions of the globe. We address three different tasks which are typical of EOS tools: taxes, pensions and retirement planning. The results of this scan are used to draw preliminary conclusions about the current state of development of EOS tools.

The remainder of the paper is structured as follows. The next section covers literature relevant to the evaluation of both e-government services interfaces and software interfaces. Section 3 develops an evaluation methodology which integrates requirements of both e-government service design and software interface design. Section 4 applies the methodology directly to different EOS tools in five different countries. Section 5 discusses the results and the final section presents conclusions and implications.
2 Literature

Literature in the e-government research stream reveals universal concern for quality of e-government services from the users’ viewpoint. For example, Kunstelj and Vintar’s (2004) review of methods of e-government development highlighted the importance of an integrated approach which provides services designed around user needs. A comprehensive review of e-government services done by Wu et al. (2009) emphasised usability and the importance of good interface design. Halaris et al. (2007) reviewed a variety of means to evaluate overall e-government website and services quality, including measures of customer satisfaction. Chatzopoulos and Economides (2009) implemented an extensive website evaluation framework known as ‘GovQual’ and identified problems in Greek municipalities’ websites in terms of the sites’ interactivity and feedback.

While e-government as a whole is concerned with electronic delivery of government services, not necessarily restricted to the internet (Fang, 2002), our study also emphasises quality from the users’ viewpoint, but by looking at one particular type of service: EOS tools for numerical analysis. In so doing, we incorporate not only recognised issues of e-government website interfaces, but also those of software interfaces. We turn first to the literature relevant to e-government service interfaces.

2.1 Online interfaces in e-government service frameworks and models

The importance of e-government interfaces is well-established in the e-government literature. For EOS tools, the literature places such interfaces at a high level of user interactivity. Gupta and Jana (2003), for example, developed a framework which incorporated cost-benefit analysis measures to support effective decisions in funding e-government initiatives. Under this framework, the present study involves Gupta and Janas’ ‘Level 5’ concerns (‘value of information’) such as usefulness and accuracy, as well as ‘Level 6’ concerns (‘system characteristics’) which address areas such as interface inputs and outputs, flexibility, simplicity and user satisfaction. Improvement is needed: Choudrie et al. (2004) characterised e-government online as improving, but lagging significantly behind equivalent services offered by businesses. Among the shortfalls they identified were the areas of quality, accessibility and privacy; as we shall see, each of these has important ramifications for consideration of EOS tools as well.

Andersen and Henriksen (2006), like others, call for a focus on end users in their study developing a five-phase model of e-government maturity. In this model, phase one is simply publishing information online, with improved interactivity as e-government matures through successive phases. Andersen and Henriksen’s fourth phase refers to completing transactions such as payment for services, while phase five refers to integrated service delivery. If EOS tools online were integrated with users’ own information already stored in online databases (as apparently planned for the future by the United States’ Social Security department, based on information provided with their tool), then such eventual software services would constitute a further type of phase five e-government service.

For now, EOS tools rely upon users to provide their own data. The current methods of provision of online software for e-government users thus constitute interactivity of high degree, but of a type not captured in existing studies of e-government services.
2.2 Interfaces in software evaluation frameworks and models

To evaluate EOS tools, it is thus necessary to look as well outside the field of e-government, to research on the quality and effectiveness of user interfaces for software. Among the relevant factors are those studied in the human-computer interface (HCI) discipline. Bevan (2001) reviews such HCI standards, citing multiple definitions of usability, such as this one developed for the ISO 9241 international quality standard:

“Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”. (Bevan, 2001, p.536)

While goals of effectiveness, efficiency and satisfaction are appropriate, it must be noted that e-government service providers do not have the luxury of designing the application toward specific users. All citizens are targeted users of e-government; this constitutes a special challenge for software interface design, one not typically undertaken for business applications. From this point of view, it becomes important in the present context to examine the interface for its ability to provide desired results under widely varying conditions of both end user technology and end user abilities; essentially, it should be user-friendly in its design (Escarfullet et al., 2010) and amenable to use by anyone who potentially needs it (Emmanouilidou and Kreps, 2010). This sense of usability is developed by Bevan (2001, p.541) in a list developed under ISO 9241, which includes ‘self-descriptive’ and ‘error tolerant’. Other considerations under Bevan which we take for granted for ESO tools are ‘suitable for the task’ and ‘controllable’; these are simple software tools compared to other more complex applications and are designed for and labelled as tools for specific and well-defined purposes, such as estimating a pension benefit.

3 Theory

The methodology used in this study is known as ‘heuristic evaluation’, which is recognised as appropriate for judging websites (Agarwal and Venkatesh, 2002). This was chosen because the services provided by EOS are made available to citizens over the internet, embedded in web pages on e-government websites (see examples below). The unique contribution of this study is to go beyond issues of standard website design, by

1 applying e-government website heuristic measures, specifically to EOS
2 incorporating issues of software usability in evaluation guidelines.

The result is an evaluation instrument aimed specifically at EOS tools. We begin with those features important for e-government services overall.

3.1 E-government services quality considerations

Accessibility: Users with visual impairment may find online services difficult or impossible to use. To overcome these barriers to use, previous researchers (Huang, 2003; West, 2007) have referred to design guidelines developed by the World Wide Web Consortium (W3C, 2002). While too extensive to fully review here, these guidelines aim at enhancing users’ ability to access information in the ways which are most effective for
them. Since current online tools employ text to explain their use, label input fields, display users’ inputs and provide outputs, users need to be able to change the text size, or access website text through browser complements which provide speech synthesis or Braille rendering capability. Size changes must not negatively impact the layout of interface elements.

Language: Although not an issue for all governments, many nations have diverse populations which may need service in multiple languages; e-government services provided to such populations should provide multiple versions of content (Gant and Gant, 2002; Barnard et al., 2004).

Security: Information security is an important factor in e-government services (Reddick, 2009; West, 2007; Henriksson et al., 2007). Security risk is incurred whenever personal information is transmitted online and users perceive certain likelihood of their information being stolen. Since the tools examined here are associated with private financial information, strong security protections are appropriate.

We now address e-government website usability features (Wang et al., 2005; Withrow et al., 2000). The question of usability for e-government websites has significant overlap with usability for websites generally. Since the evaluation methodology for this study is focused on interfaces for relatively simple software tools, the relevant aspects include basic website design considerations, as follows.

Layout: The importance of clear and logical layout for user satisfaction with e-government websites has been established (Freed, 2008). This involves placement of elements on the page, distinguished by borders and possibly colours, the placement and appearance of blocks of text and how individual elements lead users through the areas of the screen (Lynch and Horton, 1999).

Readability: The appearance of text: font typeface, size, colour and style can all impact readability for e-government website users (Becker and Nowak, 2003). Equally important is the use of text appearance controls to distinguish types of information on the page: labels, directions, input, output, etc. A well-designed interface helps users control the software by identifying each function with a unique text appearance (Lynch and Horton, 1999).

Look and feel: In e-government website design, this aspect of appearance refers to consistency for the sake of usability; implementing a consistent appearance for similar elements helps orient the user in terms of navigation and function (Vassilakis et al., 2003). For simple software interfaces, the cognitive difficulty is less severe than on, for example, large and complex websites, but the basic principle still applies and its importance for e-government has been noted (Wimmer and Tambouris, 2002). We now turn to issues which are specific to the quality of software.

3.2 E-government software quality considerations

The effectiveness of software interfaces involves elements not accounted for in the sections above and in the literature on e-government generally (Henriksson et al., 2007). The following paragraphs address elements relevant to the effectiveness specifically of online software interfaces for e-government analytical tools online. We argue that these elements are important in the evaluation of EOS tools because citizens making use of
such services come to them in the role of software users. The problem of citizen as online software user, which has not been addressed in prior research in e-government services, raises issues from the discipline of software development, in which developers must strive to make their products not only useful, but also easy to use. In this case, the special problem to be addressed is the extremely wide range of computer skills in the pool of intended users of EOS. Our unique contribution, thus, is the development of an evaluation instrument which incorporates principles for software usability and applies them to EOS tools.

*Instructions/documentation:* E-government researchers have noted that transaction-oriented functions in e-government should include clear instructions (Gant and Gant, 2002). Likewise, all software requires documentation: instructions and guidelines for how to use the analytical tool. Unlike applications such as office productivity software or websites such as social media providers, citizens are unlikely to use EOS every day. For the purposes of financial planning tools online, developers should include explanations not only of the controls, but also of the individual financial elements included, their meaning and significance.

*Input validation:* Standard software design includes input validation: preventing errors by checking to see that user inputs are of the type required by the underlying code (Wenham and Zaphiris, 2003). Online software should check to see that inputs for such items as annual income, years of service, etc., include only numeric characters and that such inputs are reasonable (no users claiming 210 years of service, for example). While these are important for individual fields, software should also check to see that users’ inputs are do not entail contradictions among multiple fields (a person who began to work in 1988 and retired in 2008 could not have 25 years of work, for example). Erroneous output may lead to significant adverse impacts, especially if users take action based on such outputs.

*Error recovery:* When users do make mistakes in their use of the tool, such as entering incorrect data, the software interface should halt the analysis, warn the user of an error, explain what it is and how it occurred and provide instructions on how to rectify it (Wenham and Zaphiris, 2003). Ideally, the software will automatically return the user to the point at which the error was made, ready for the data to be re-entered correctly, without resetting all the values and asking users to start the process over from the beginning.

*Save/store results:* Once users have completed an analysis, the software should provide users with a means of preserving their results for that particular run. One important use in particular of financial planning tools would be the performance of what-if analyses – users change one or more parameters among the inputs to see the effect on the results. A ‘save’ feature would make subsequent comparisons relevant to important financial decisions easy to carry out.

Software as opposed to web content involves a few more considerations, relevant to e-government, which impact its usability.

*Platform:* While implementation of information presentation and transaction functions in e-government can be carried out with typical web code components such as text, images and forms, analytical software tools must be implemented in actual software code. This can be made available to users by a variety of means: Javascript, Java applets, Java server pages, Active Server Pages and ActiveX controls are a few examples. While many users
are capable of employing these and more technologies, not all will necessarily have the ability to run Java applets, or the willingness to download and run ActiveX controls.

**Learnability:** Nielsen (1993) identifies learnability as an important element in system usability, referring to the cognitive load placed on persons trying to use it. Software tools vary enormously in complexity, depending on the tasks undertaken and the users targeted. For the case of e-government, the very broad range of skills in the citizen population forces developers to limit the complexity of the tool, to make it available in practical terms to as many citizens as possible (Wenham and Zaphiris, 2003). Related to this is the notion of data requirements.

**Data requirements:** One inherent difficulty with online software tools is the need, for analytical purposes, for users to provide data themselves. This difficulty can be lessened in some cases by allowing users to select from a limited menu of options. However, for financial planning, there is an unavoidable need for users to provide relevant numbers which apply to themselves – projections of future dates, income, needs, etc., which could not be provided by government-hosted databases in any case, even if live links were available for other purposes. The need to type in long strings of figures raises the likelihood of errors and may provide a disincentive for users to adopt the tool.

**Revisions:** Users can make good use of financial planning tools, as already noted above, by performing what-if analyses: “What if I retire at age 68 instead of 67?”, “What if my income increases by 3% next year?” and so forth. For users to carry out such multiple analyses, usable software tools will allow users to return to form data they have already filled in and change a few factors, without having to re-enter everything again from scratch. This sort of functionality helps fulfil a requirement that e-government interactivity be easily controlled by users, including undo and redo actions (Wenham and Zaphiris, 2003).

### 4 Methodology

The HCI research discipline has developed multiple methods of evaluating software usability. Some major categories in this area include heuristic evaluation, cognitive walkthrough and action analysis (Holzinger, 2005). These methods are all carried out by inspection of the interface; others involve testing software performance under various conditions. To address the problems and challenges relevant to EOS, aspects of both approaches are included in the analysis methodology.

This paper develops a heuristic evaluation method (Agarwal and Venkatesh, 2002; Holzinger, 2005; Wenham and Zaphiris, 2003; Jeffries et al., 1991). For this approach, evaluators inspect the interface for a list of features, characteristics and capabilities judged to be important for the particular software in question, to produce a set of guidelines for good design. Wenham and Zaphiris (2003) produced a list of twelve such guidelines for the evaluation of online banking websites for the purposes of comparison. While such banking websites are not directly analogous to e-government financial tools, a number of features they list are indeed relevant to online analytical software, such as security, error prevention and error recovery. Therefore, we provide an extended exercise in the development and use of a heuristic evaluation framework for e-government financial planning online software tools which are the subject of the present study.
Table 1  Heuristic evaluation guidelines – website features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Poor – 0 points</th>
<th>Fair – 1 point</th>
<th>Good – 2 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>No controls are provided to address problem of accessibility for the visually impaired. Browser controls to resize negatively impact layout</td>
<td>No controls are provided to address problem. Browser controls to resize do not negatively impact layout</td>
<td>Controls are provided to address problem. Resizing does not negatively impact layout</td>
</tr>
<tr>
<td>Language</td>
<td>No provision is made for citizens speaking other languages</td>
<td>Provides information about alternative resources for citizens speaking other languages</td>
<td>Provided in multiple languages</td>
</tr>
<tr>
<td>Security</td>
<td>Personal data is transmitted in the clear, with no security protections</td>
<td>Personal data is collected but analysed locally on the client’s machine</td>
<td>Security protocols such as SSL are implemented on data collection and analysis pages of the tool</td>
</tr>
</tbody>
</table>

Note: Features: online software tool exhibits characteristics identified by e-government literature as important in providing services.

Table 2  Heuristic evaluation guidelines – website usability

<table>
<thead>
<tr>
<th>Feature</th>
<th>Poor – 0 points</th>
<th>Fair – 1 point</th>
<th>Good – 2 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layout</td>
<td>Fails to use graphic cues such as colours and borders to produce clear and logical separation of the functional areas of the tool</td>
<td>Uses basic graphic cues such as colours and borders to separate the functional areas of the tool</td>
<td>Provides highly effective use of graphic cues such as colours and borders to separate the functional areas of the tool</td>
</tr>
<tr>
<td>Look and feel</td>
<td>Appearance of page is unrelated to the rest of the e-government website. Design is unattractive and difficult to use</td>
<td>Appearance of page is related to the rest of the website. Design is presentable but not attractive; functional but not simple and easy</td>
<td>Appearance of the page is consistent with overall website design. Design is attractive, simple and easy to use</td>
</tr>
<tr>
<td>Readability</td>
<td>Fails to make use of font features such as typeface, bold, colour and size to categorise text elements by meaning and function</td>
<td>Makes minimal use of font features such as typeface, bold, colour and size to categorise text elements by meaning and function</td>
<td>Makes highly effective use of font features such as typeface, bold, colour and size to categorise text elements by meaning and function</td>
</tr>
</tbody>
</table>

Note: Usability: online software tool exhibits characteristics identified by e-government literature as important for website usability.

To operationalise the elements discussed in Section 3 above, a set of guidelines for heuristic evaluation (Wenham and Zaphiris, 2003) was developed. For each of the fourteen features and capabilities, the theoretical considerations above were used as a basis for developing definitions of poor, fair or good levels of success in the online software tool. These levels were scored as 0, 1, or 2 points respectively. The resulting guidelines for heuristic evaluation of software online tools for the first six elements of website features and website usability, which have been identified as important for e-government websites, are shown in Tables 1 and 2. The use of these guidelines in this
study constitutes the first application of heuristic evaluation methodology specifically to EOS. The resulting guidelines for heuristic evaluation of a further seven considerations which are important to the effectiveness and usability of software appear in Tables 3 and 4 below. Together, these constitute 14 separate categories of evaluation; with a possible rating of 0, 1 or 2 on each, the aggregate scores for any single EOS tool may range from 0 (worst) to 28 (best).

Table 3  Heuristic evaluation guidelines – software effectiveness

<table>
<thead>
<tr>
<th>Poor – 0 points</th>
<th>Fair – 1 point</th>
<th>Good – 2 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructions</td>
<td>Few or no instructions are provided</td>
<td>Basic instructions are provided</td>
</tr>
<tr>
<td>Error detection</td>
<td>Fails to detect errors in the data entered by users</td>
<td>Detects some types of errors in the data entered by users</td>
</tr>
<tr>
<td>Error recovery</td>
<td>Fails to provide any functionality in correcting errors</td>
<td>Provides notification of errors. Provides little or no guidance on how to correct</td>
</tr>
<tr>
<td>Save results</td>
<td>No means of saving the analysis is provided</td>
<td>Allows user to print out the completed analysis, but no electronic save is provided</td>
</tr>
</tbody>
</table>

Note: Software effectiveness: online software tool exhibits characteristics important in providing highly effective software.

Table 4  Heuristic evaluation guidelines – software usability

<table>
<thead>
<tr>
<th>Poor – 0 points</th>
<th>Fair – 1 point</th>
<th>Good – 2 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Uses a technology which may require additional downloads or configuration changes or permissions to be made by the user. Requires a specific browser</td>
<td>Uses a technology which may require additional downloads or configuration changes or permissions to be made by the user. Does not require a specific browser</td>
</tr>
<tr>
<td>Learnability</td>
<td>Complex tool to understand and learn. Labels unclear and/or unexplained</td>
<td>Moderately difficult to understand and learn. Labels are adequate.</td>
</tr>
<tr>
<td>Data</td>
<td>Requires user to do considerable work collecting and preparing data for use in the tool</td>
<td>Requires user to do some work collecting data, but no work to prepare data for use in the tool</td>
</tr>
<tr>
<td>Revisions</td>
<td>Provides no means for revising the analysis short of starting over</td>
<td>Provides means of revising the analysis without starting over</td>
</tr>
</tbody>
</table>

Note: Software: usability: online software tool exhibits characteristics important for software usability.
5 Data and analysis

To demonstrate the use of this heuristic evaluation method, five different online software tools were examined. These were drawn from five different nations in four different regions of the globe, addressing three of the most typically provided types: tax estimators, pension calculation and retirement planning. The trial sites evaluated in April of 2011 are shown in Table 5.

Table 5 Online software tools examined

<table>
<thead>
<tr>
<th>Tool and country</th>
<th>URL</th>
<th>Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>US income tax withholding estimation tool</td>
<td><a href="http://www.irs.gov/individuals/page/0,,id=14806,00.html">http://www.irs.gov/individuals/page/0,,id=14806,00.html</a></td>
<td>CGI server-side platform</td>
</tr>
<tr>
<td>India pension calculation tool</td>
<td><a href="http://www.pensionersportal.gov.in/PensionCalculators/PensionCalculator.asp">http://www.pensionersportal.gov.in/PensionCalculators/PensionCalculator.asp</a></td>
<td>Javascript client-side platform</td>
</tr>
</tbody>
</table>

These sites were chosen to demonstrate the tool in a broad variety of contexts: five different nations are represented, from four different parts of the globe, with quite different cultures, as well as differing approaches to the problem of providing EOS for various tasks. Table 6 shows the overall results, using aggregated scores for each site (Gupta and Jana, 2003).

Table 6 Heuristic evaluation score totals and averages

<table>
<thead>
<tr>
<th></th>
<th>US tax</th>
<th>India pension</th>
<th>UK pension</th>
<th>Australia retirement</th>
<th>Singapore tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviewer 1</td>
<td>16</td>
<td>13</td>
<td>17</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td>Reviewer 2</td>
<td>15</td>
<td>11</td>
<td>15</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>Reviewer 3</td>
<td>19</td>
<td>13</td>
<td>22</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>37</td>
<td>54</td>
<td>70</td>
<td>57</td>
</tr>
<tr>
<td>Avg (unbalanced)</td>
<td>16.67</td>
<td>12.33</td>
<td>18</td>
<td>23.33</td>
<td>19.00</td>
</tr>
<tr>
<td>Avg. (balanced)</td>
<td>16.33</td>
<td>12.44</td>
<td>17.79</td>
<td>23.04</td>
<td>18.67</td>
</tr>
</tbody>
</table>

Note: Maximum possible score (single reviewer): 28.

Because the tool uses three measures each for e-government features and e-government usability, but four measures each for evaluating e-government software effectiveness and usability, the latter two categories have greater impact on the overall raw score. Were each of these categories to be considered equally important, it would be necessary to apply a weighting system to balance the contributions of each category of evaluation (Gupta and Jana, 2003). In Table 6, these balanced scores were obtained by multiplying the weight of each of the first two categories (e-government features, e-government
usability) by 7/6 and the weight of the last two categories (software effectiveness and usability) by 7/8; for either balanced or unbalanced scores, the resulting maximum and minimum possible score range is the same, 0 to 28.

The data produced by this measure is ranked ordinal data, in which a ‘0’ vs. a ‘2’ rating reflects a greater level of disagreement than ‘0’ vs. ‘1’ or ‘1’ vs. ‘2’. To take these characteristics into account, inter-rater reliability was calculated with weighted Cohen’s kappa for each pair of raters, using quadratic weighting (Sim and Wright, 2005). The results of the tests for inter-rater reliability appear below in Table 7. The percentage agreement values are also provided for comparison.

Table 7  Inter-rater reliability

<table>
<thead>
<tr>
<th>Inter-rater measure</th>
<th>Raters</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 and 2</td>
<td>1 and 3</td>
</tr>
<tr>
<td>Weighted Cohen’s Kappa (Quadratic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td>0.571</td>
<td>0.622</td>
</tr>
<tr>
<td>Z</td>
<td>4.82</td>
<td>5.49</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Percentage agreement</td>
<td>% agree</td>
<td>61.4</td>
</tr>
</tbody>
</table>

These reliability results fall into the higher side of the range designated as ‘moderate agreement’ (0.4–0.6) by (Landis and Koch, 1977). Given the very broad range of EOS tools evaluated here, considered by variety of countries, tasks and technical approaches to
the interface (including forms, wizards and dashboards), we consider this an encouraging result.

To provide a more granular view of the results, Figure 1 provides the total scores from all reviewers of each site, aggregated and weighted by the four major categories for cross-category comparison: e-government features, e-government usability, software effectiveness and software usability. The values are shown in Table 8. The maximum possible score for any single cell is 21. Here, we can see that among the example websites evaluated, the components comprising the e-government features, namely, Accessibility, Languages and Security had the poorest results. All of the examples had lower scores in this category than in any of the other three.

Table 8 Category scores of EOS software

<table>
<thead>
<tr>
<th>Country</th>
<th>E-gov features</th>
<th>E-gov usability</th>
<th>SW effectiveness</th>
<th>SW usability</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>4.67</td>
<td>16.33</td>
<td>14.88</td>
<td>13.13</td>
</tr>
<tr>
<td>India</td>
<td>4.67</td>
<td>15.17</td>
<td>6.13</td>
<td>11.38</td>
</tr>
<tr>
<td>UK</td>
<td>10.50</td>
<td>14.00</td>
<td>12.25</td>
<td>16.63</td>
</tr>
<tr>
<td>Australia</td>
<td>10.50</td>
<td>21.00</td>
<td>21.00</td>
<td>16.63</td>
</tr>
<tr>
<td>Singapore</td>
<td>9.33</td>
<td>15.17</td>
<td>14.00</td>
<td>17.50</td>
</tr>
</tbody>
</table>

6 Discussion, implications and conclusions

This paper sets out to develop and use a heuristic evaluation tool specifically for EOS applied across five different EOS tools from five different nations in four different regions of the globe, for three different types of tasks and four different technology implementations, the results indicated both strengths and weaknesses in current EOS implementations. We examine first the apparent strengths of these EOS tools.

6.1 EOS strengths

The website usability features of e-government usability – layout, look and feel and readability – received at least adequate attention on the examples considered here. However, since the user is actually carrying out analyses with software instead of simply passively receiving government information, the importance of these elements is heightened. Clarity of presentation in this context (software) becomes a matter not merely of comprehension and user satisfaction, but of being able to carry out and correctly complete an analysis at all.

The results in terms of software effectiveness indicated that attention is being paid to the need for user instructions, with all but one example receiving more than half the available maximum number of possible points (21). However, it should be noted, since the context is analytical software, that the need for excellent instructions and documentation is heightened here, as compared to non-software e-government services, again because lack of adequate instruction could prevent users from being able to obtain any useful result at all.

Error detection and recovery (two of the components of ‘software effectiveness’) showed some room for improvement, with only one tool, the Australian retirement
planner, receiving maximum scores in both respects. For this particular context, financial planning, the opportunities for error are considerable and the results, should incorrect analyses be used as the basis for real-world financial actions, could cause substantial problems for users.

The lack of means for users to save the results of their analyses and particularly in some form which would easily lend itself to side-by-side comparison of alternatives, is a notable problem for four of the examples examined here, with Australia again the only tool yielding top scores. While some additional programming investment would be needed to provide such a feature, its usefulness to citizens could be expected to considerably improve user adoption and satisfaction.

These example sites deployed their EOS tools via a variety of technology platforms: server-side processing using CGI (USA), client-side Javascript (India, UK), Adobe Flash-based dashboard (Australia) and Active Server Pages (Singapore). Each of the examples represents a technology which could reasonably be expected to be available on nearly all users’ machines. In this case, the work of cross-compatibility has been done by the creators of browser software and internet technologies and the program code is so limited that no users are in danger of having insufficient RAM or CPU power, as might be the case with installed software. Indeed, in terms of universal technological usability, this quality alone constitutes a powerful argument for e-government service providers to exclusively pursue only online software solutions similar in terms of platform to the ones seen here. Also likely because of the limited ambitions of the software – these are simple tools, for the most part, compared to what software generally is capable of – none of the examples posed severe problems for learnability.

On the other hand, the data requirements for financial planning can be considerable; this problem was particularly noticeable in the US Social Security tool. Faced with a page full of blank fields for annual income to be filled in one by one, by hand, how many users are likely to do the considerable work of digging up all the numbers and laboriously typing them in? For this particular case, financial planning, a Phase V level of e-government service quality (Andersen and Henriksen, 2006) that provides integrated transaction support, which in this case would consist of linking existing government database information directly to the tool, would be highly desirable.

The EOS example websites performed relatively better on e-government usability, software effectiveness and software usability than on e-government features, with just two exceptions (both from the Australian wizard-based example). However, with most scores falling considerably short of the maximum possible (21 points), room for improvement clearly remains. We turn now to weaknesses of EOS tools in the example websites.

6.2 EOS weaknesses

The sample websites evaluated all had their lowest scores on the measures of features deemed important specifically for e-government. In terms of accessibility, all of the sites would nominally require adequate visual acuity to use the tool and all rely upon users with diminished visual capabilities to themselves make any adjustments required. In some cases, those adjustments, such as user-adjusted text size in the browser, trigger further problems with layout of the elements of the page. Of course, making analytical
software tools easily usable by the visually impaired would be very challenging in any circumstances; it is noted here as a consequence of the considerable attention which e-government literature has already dedicated to the special requirements of government-provided services, which promise access to all citizens, including those with physical disabilities.

Our examples did not provide language alternatives; while having users undertake to translate the text of the software interfaces themselves, this falls short of the guidelines for good e-government service (Barnard et al., 2004; Gant and Gant, 2002). Security is another area which appears to have received inadequate attention. The India and UK tools do avoid one possible class of problems (transmitted personal data) by implementing their tools as Javascripts which run on the clients’ side and, therefore, involve no data transmission. For maximum security and protection against eavesdropping, it would be preferable for such private data to rely on security protocols such as Secure Sockets Layer or an equivalent.

6.3 Implications and conclusions

This paper presents a first effort in the development of a heuristic evaluation procedure for EOS. It fills gaps in both the research literature and current practice by addressing a prominent and widely provided e-government service, online software tools, for which no specific evaluation method had previously been created. The development process used to create an evaluation method employed extant and well-established principles drawn from the literature for addressing both the needs of e-government websites and those of software, focusing on the development of guidelines based on factors considered most relevant to e-government users.

Implications for practitioners from the example analyses are that EOS tools are not yet optimal from the standpoint of either e-government or software development issues. Specifically, our indications are that EOS services could benefit from improvements in many areas: accessibility, security and languages are particularly in need of closer attention, while error detection and recovery as well as other aspects of usability are weaker than they could be.

Implications for research include further refinement of the heuristic evaluation procedure, as well as rigorous testing of the evaluation procedure in a variety of contexts: topics other than taxes and pensions, or financial planning software generally; online software provided by different government departments, or levels of government; collecting evaluation responses from different segments of the citizen population; and testing of the possible impacts of cultural differences on online software evaluation methods’ content and approaches. Furthermore, alternative methods for evaluating both e-government services and software exist; these may also be usefully applied to the improvement of our understanding of EOS.

As e-government services generally continue to move towards greater interactivity, integration and sophistication, the means by which they are evaluated should grow in sophistication as well. It is hoped that the product of this study can, thus, usefully serve as the basis for further development in this area of e-government services.
References
A heuristic evaluation instrument for e-government online software


