City Research Online

City, University of London Institutional Repository

Using Information Behaviors to Evaluate the Functionality and Usability of Electronic Resources: From Ellis's Model to Evaluation

Stephann Makri, Ann Blandford, and Anna L. Cox
UCL Interaction Centre, University College London, 8th Floor, Malet Place Engineering Building, Gower Street, London WC1E 6BT, United Kingdom. E-mail: {s.makri, a.blandford, anna.cox}@ucl.ac.uk

Information behavior (IB) research involves examining how people look for and use information, often with the sole purpose of gaining insights into the behavior displayed. However, it is also possible to examine IB with the purpose of using the insights gained to design new tools or improve the design of existing tools to support information seeking and use. This approach is advocated by David Ellis who, over two decades ago, presented a model of information seeking behaviors and made suggestions for how electronic tools might be designed to support these behaviors. Ellis also recognized that IBs might be used as the basis for *evaluating* as well as designing electronic resources. In this article, we present the IB evaluation methods. These two novel methods, based on an extension of Ellis's model, use the empirically observed IBs of lawyers as a framework for structuring user-centered evaluations of the functionality and usability of electronic resources. In this article, we present the IB methods and illustrate their use through the discussion of two examples. We also discuss benefits and limitations, grounded in specific features of the methods.

Introduction

In this article, we present the information behavior (IB) methods—two novel, specialized methods for evaluating electronic resources ranging from Internet search engines to digital libraries and indexes/citators. The IB methods are novel as they are based on the observed IB of lawyers and are theoretically underpinned by an extension of Ellis's behavioral model of information seeking. The methods are specialized in the sense that they are intended to be used to evaluate *electronic resources* as opposed to other types of interactive system. The IB methods were developed based on the premise that observed IB can provide a useful structure for evaluating electronic resources by using the observed behaviors as theoretical "lenses" to evaluate the *functionality* of electronic resources (i.e., the features provided by the resource aimed at supporting users) and the *usability* of these resources (i.e., how easy to use the resource is).

Although both the IB functionality and usability methods are underpinned by the same IB theory, they can be regarded as two separate evaluation methods in their own right. An IB functionality evaluation aims to provide data on the range of functionality supported by a particular electronic resource (and provides the basis for discussing whether this range is appropriate). An IB usability evaluation aims to provide data relating to the difficulties that users face when using a particular resource and how severe and easy to address the evaluator considers these issues to be.

We continue this article by discussing the rationale behind the IB methods and the theoretical basis of the methods. This includes reference to the empirically observed IBs observed among academic lawyers, which forms part of the basis of the methods (see Makri, Blandford, & Cox, 2008). This discussion also includes reference to the IBs identified in previous studies by Ellis and his colleagues (see Ellis 1989; Ellis, Cox, & Hall, 1993; Ellis & Haugan, 1997) and by Meho and Tibbo (2003). We then provide an overview of both of the methods, including a discussion of the attributes that make the methods novel and useful. Next, we illustrate the use of the IB methods through the discussion of two detailed examples of how the methods can be used to evaluate both the functionality and usability of the current public version of LexisNexis Butterworths (LNB), a widely used electronic legal resource available to academic institutions and law firms worldwide. Finally, we conclude by discussing the benefits and limitations of using the methods.

Page numbers differ from published version

Background: The Rationale Behind the IB Methods

IB research has led to the development of several models that describe aspects of IB. These have ranged from models that view information seeking as a series of stages, such as Marchionini's (1995) model, to those that view it as a cognitive or affective process (e.g., Sutcliffe & Ennis, 1998; Kuhlthau, 1988) to those that regard information seeking as a set of interrelated behaviors (e.g., Ellis, 1989; Ellis, Cox, & Hall, 1993; Ellis & Haugan, 1997). Other models and approaches have been devised that allow us to conceptualize information seeking in more ways still, such as a problem-solving activity (Wilson, 1999), an evolutionary process (Bates, 1989), a "foraging" activity (Pirolli & Card, 1995), a gap in knowledge to be filled (Belkin, 1980), and a sense-making activity (Dervin, 1983). Each of these models and approaches provide a different insight into IB. However, most of these models and approaches provide scope only for gaining insights into people's behavior and do not provide support for examining IB with the purpose of using the insights gained to design new tools or improve the design of existing tools to support information seeking and use. Indeed, as highlighted by Colbert, Peltason, Fricke, R., and Sanderson (1997), "their fitness for purpose for design has rarely been assessed" (p. 73). One exception is the work of David Ellis who, over two decades ago, presented a model of information seeking behaviors and made suggestions for how electronic tools might be designed to support these behaviors. He suggests that "the general principle of using the behavioral aspects of users' information seeking activities to inform the design of electronic resources . . . could play a more prominent role in the design of computer based information retrieval systems than, at present, it does" (Ellis 1989, p. 202). Ellis also recognized that his behavioral model might be used as the basis for *evaluating* electronic resources, "as an evaluatory tool to identify the existence and ease of implementation of features of the model in existing systems" (Ellis, 1987, p. 242).

Our IB method, based on an extension of Ellis's model, uses the empirically observed IBs of lawyers as a framework for structuring user-centred evaluations of the functionality and usability of electronic resources. Part of the empirical work that forms the basis of the methods has been published in Makri et al. (2008) and concerns the IB displayed by *academic* lawyers (i.e., law students and academic/research staff). The remainder of the empirical work that forms the basis of the methods concerns the behavior displayed by *practicing* lawyers working for the London branch of a multinational law firm. Both academic and practicing lawyers performed similar IBs and we refer to both sets of findings in this article.

The rationale behind the IB methods is closely related to Norman's (1986) notion of "bridging the gulf" between user and system (see Figure 1). The IB *functionality* method aims to support people with a stake in a resource to "push" the resource closer to its users by supporting them in examining whether the resource supports an appropriate range of user behavior. The IB *usability* method aims to push electronic resources closer to users by supporting stakeholders to identify usability issues that, if addressed properly, should lead to improved support for user behavior.

This is with the broad aim of improving the design of interactive systems so that they "speak the user's language" as opposed to forcing users to learn the system's language. Through the IB evaluation methods and the resultant functionality and usability insights that they provide, we seek to help bridge the gulfs between a system and its users by making the system more closely match user goals and therefore become more user-centred.

Background: The Theoretical Basis of the IB Methods

The IB methods are based primarily on the empirical observation of a vertical slice of 32 academic and 24 practicing lawyers, who were asked to find information that they currently need for their work (or if they did not have a current information need, to step through a recent information seeking episode). The academic lawyers ranged from first year undergraduate law students to PhD level, and also included research and academic staff up to professor level. The practicing lawyers ranged from paralegals and trainees to associate level and worked for two London-based departments of a multinational law firm. The first department (Dispute Resolution, otherwise known as litigation), dealt mostly with contentious legal issues where there were two legal parties involved. The second department (Tax) dealt mostly with non-contentious issues involving a single client. During the observation, the lawyers were asked to think aloud (i.e., verbalize their thoughts, actions, and feelings) while using the electronic resource or resources of their choice to find the required information. During the think-aloud session, the lawyers were asked short but probing questions surrounding their actions. A more detailed methodology and the findings of this study in relation to academic lawyers are presented in Makri et al. (2008).

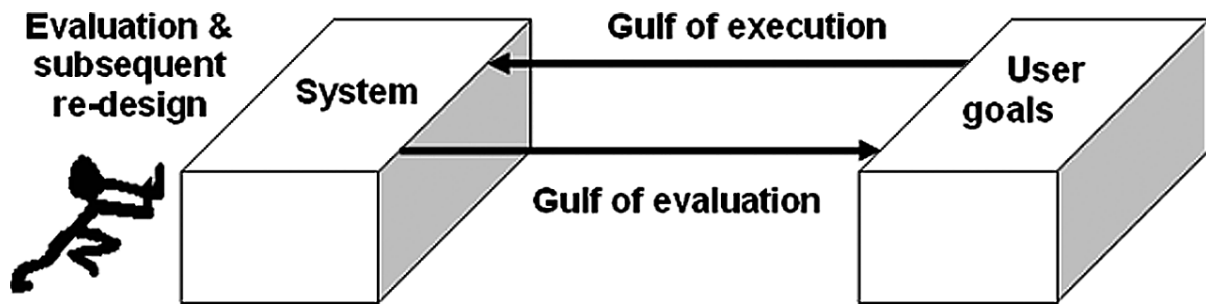*Page numbers differ from published version*

FIG. 1. How evaluation and subsequent re-design can bridge the gulfs between system and users by 'pushing' systems closer to users and their goals. Diagram adapted from Norman (1986).

The observations led to the identification of a number of IBs displayed by the lawyers, many of which were similar to those identified by Ellis and his colleagues (see Ellis, 1989; Ellis et al., 1993; Ellis & Haugan, 1997). These observations served to validate Ellis's model in the new domain of law and also served to extend Ellis's model to include additional behaviors pertinent to legal information seeking (updating and history tracking behaviors, which are discussed later), broaden the scope of the model through the identification of information *use* as well as information seeking behaviors, and enhance the potential analytical detail of the model through the identification of several *levels* at which many of the behaviors were found to operate. The model of IBs resulting from these observations, which incorporates the theoretical enhancements described above, forms the theoretical basis of the IB methods. The IBs at the heart of the model are presented, along with definitions, in Appendix A. We also discuss these behaviors, in the context of previous related work, below. This not only includes work by Ellis and his colleagues, but also by Meho and Tibbo (2003), who revisited Ellis's findings to see whether they were still applicable now that electronic information seeking has become more widespread. This is followed by a discussion of the concept of *levels* at which these behaviors can operate.

The Information Behaviors that form the Basis of the IB Methods Discussed in Relation to Previous Work

Ellis's (1989) model is based on interviews with academics and commercial researchers from a number of scientific disciplines: the social sciences (Ellis, 1989), physical sciences (Cox, 1991; Hall, 1991—both reported in Ellis et al., 1993) and engineers and research scientists (Ellis&Haugan, 1997). The latter study served to validate the model outside an academic domain, while interview data by Smith (1988) was analysed by David Ellis to validate the model in the non-scientific discipline of English literature. All of these studies identified similar behavioral characteristics (although reports of the studies of physicists and English literature students in particular used different terminology to describe what Ellis et al. (1993) consider to be similar behaviors). Ellis (1993) attributes this difference in terminology to the fact that "the later studies were not simple verification studies as, in each case, the models were developed from the data and then compared to the original model" (p. 483). Ellis also notes that the models in each of the studies differed in detail somewhat, partly due to minor differences in focus between studies and partly from subject differences between groups. Overall, Ellis and Haugan (1997) deem the model to be "quite robust in relation to the information seeking patterns of scientists, engineers and social scientists in both an academic and an industrial research environment and over a period of time which has seen accelerating changes in the information environment itself" (p. 402).

Ellis's model comprises a number of broad information seeking behaviors identified from the interviews that, according to Ellis (1989), are non-sequential. It is also possible to display more than one characteristic at any given time. A number of these behaviors were also identified in our empirical observations of lawyers (see Makri, Blandford, & Cox, 2008). The behaviors identified by both Ellis and ourselves are as follows:

- *Starting/Surveying*. According to Ellis et al. (1993), starting involves "activities characteristic of the initial search for information" (p. 359). Ellis and Haugan (1997) elaborate on this definition, suggesting that this behavior (which they renamed "surveying") is "characteristic of the initial search for information to obtain an overview of the literature within a new subject field, or to locate key people operating in this field" (p. 395). Ellis highlights that this behavior can be supported in

3   *Page numbers differ from published version*

electronic resources by helping users to identify review-type or heavily-cited materials and by providing an indication of the sources that publish material in the required area. These design recommendations are sometimes, but not often, supported in current electronic resources.

- *Chaining*. "Following chains of citations or other forms of referential connections between material" (Ellis, 1989, p. 179).

Ellis (1989) highlights that there are two types of chaining: forwards chaining (which involves identifying and accessing documents that have subsequently cited the current document) and backwards chaining (which involves following references to documents that have been cited in the current document). Ellis suggests that electronic resources should support both types of chaining, and they can provide enhanced chaining support by indicating material in the database by a particular author, by providing a facility for identifying all types of (forward and backwards) referential connections to and from the material of interest, and by supporting more advanced forms of citation chaining (such as bibliographic coupling—identifying common works that particular documents cite— and cocitation searching—identifying pairs of highly cited papers). Basic support for forwards and backwards chaining is commonly supported by current electronic resources, however the type of advanced chaining support described by Ellis is less commonly supported.

- *Browsing*. "Semi-directed searching in an area of potential interest" (Ellis, 1989, p. 179). Ellis suggests that support for browsing all types of information held on an electronic resource should be provided by infrared radiation (IR) systems. Support for browsing material in electronic resources is increasing, although we are yet to see many electronic resources that allow all of their content to be browsed.
- *Differentiating*. "An activity which uses differences between sources as a filter on the nature and quality of the material examined" (Ellis et al., 1993, p. 179). Ellis (1989) explains that "differentiating is effected by the researcher identifying different sets of sources in terms of the differing probability of their containing useful material" (p. 190). Ellis suggests that electronic resources can support differentiating by allowing users to specify preferences for sources that they think are most likely to contain material of interest and then using these preferences to restrict the search, to exclude certain sources or types of source from the search, or to display material in order of source or source type. This, Ellis (1989) asserts, "is far easier than attempting to deal with differences of approach and level of treatment directly" (p. 193).
- *Filtering*. The "use of certain criteria or mechanisms when searching for information to make the information as relevant and as precise as possible" (p. 399). Current electronic resources support filtering in a number of ways, all allowing users to refine or reformulate their searches and some allowing users to search within the current search results, perform field-restricted searches (including date restriction).
- *Monitoring*. "Maintaining awareness of developments and technologies in a field through regularly following particular sources" (Ellis & Haugan, 1997, p. 396). Ellis (1989) suggests that "monitoring" can be supported by allowing users to specify sources to monitor and automatically searching these sources when the user next logs in to the system or when the sources are next updated. Ellis also suggests that monitoring may be enhanced through the provision of an "alerting" function, where sources of interest are brought to the attention of the user. E-mail alerts are becoming increasingly common in current day electronic resources, with the systems providing the facility for users to save searches and have them automatically and periodically run by the system on their behalf (and any new results sent to them in an e-mail). Only a limited amount of monitoring behavior was displayed by the lawyers in our study, most likely due to the fact thatthe broad task they were set involved them finding or stepping through how they went about finding information as opposed to maintaining awareness of developments.
- *Extracting*. "Systematically working though a particular source to identify material of interest" (Ellis et al., 1993, p. 364). Ellis (1989) explains that for social scientists ,"the source may consist of a run of a periodical, a set of conference proceedings, a series of monographs, the contents of an archive, a collection of publishers' catalogues, or bibliographies, indexes or abstracts" (p. 198). Therefore, *material of interest* in this instance can be regarded as a relevant document, reference, or abstract as opposed to the textual content within it. Ellis suggests that extracting can be supported by facilitating "continuous movement through different source streams," (p. 200) ensuring that material in the electronic resource is "recomposed more or less in their original form for searching purposes" (p. 200). Many current day electronic resources store and display documents under hierarchies that mirror the paper-based form of the information (for example all

Page numbers differ from published version

the papers from a particular conference proceedings or journal issue are stored together to facilitate browsing as well as searching).

A couple of behaviors were identified by Ellis and his colleagues in other disciplines but were not found among our lawyers. These behaviors, which do not feature by default in the enhanced behavioral model that forms the basis of the IB methods, are *verifying* and *ending*. Verifying involves "checking the information and sources found for accuracy and errors" (Ellis et al., 1993, p. 364). Verifying was identified in Ellis et al.'s study of physical scientists, in Ellis and Haugan's (1997) study of engineers and research scientists, and Meho and Tibbo's (2003) study of social scientists. As the legal documents that lawyers found were considered to be accurate and free of errors simply because they had become law, this behavior was not identified among lawyers in our observations. Ending involves "the assembly and dissemination of information or the drawing together of material for publication" (Ellis et al, 1993, p. 365) and was also identified in the studies by Ellis et al., 1993 and Ellis and Haugan, 1997. similar behavior, *assembly* and *dissemination*, was also identified in the study of English literature researchers by Smith (1988), reported in Ellis et al. (1993). These studies give examples of ending behavior that involve *finding information* as part of the process of assembly and dissemination, which suggests that the scope of this behavior as defined by Ellis is limited to activities associated with finding information towards the end of the writing process. The behavioural model that forms the basis of the IB methods does not include ending behavior as we do not distinguish between the activities involved in finding information at different stages of the information seeking process.

When revisiting Ellis's findings among social scientists Meho and Tibbo (2003) identified three new behaviors: accessing, information managing, and networking. Although we did not observe much networking among academic and practicing lawyers (practicing lawyers tended to use internal knowledge management sources to obtain information published by their colleagues, possibly due to the large size of the law firm), accessing was found to be an important IB. Although Meho and Tibbo discussed accessing solely related to physical problems accessing paper-based information, such as long travel distances and difficulty getting hold of published materials within certain countries, accessing behavior as identified in *our* empirical study of lawyers' IB (see Makri et al., 2008) involved *electronic* as opposed to physical access to resources. Current day electronic resources support accessing through two main methods; some offer automatic (and invisible) Internet protocol (IP) recognition to recognize the company or institution from which the user is accessing the system and grant them access to the system or materials within the system. Others rely on users logging in with a username and password. Others still rely on a mixture of the two, mainly in order to provide personalisation services to users.

Meho and Tibbo (2003) found that many social scientists spoke about "filing, archiving, and organizing information collected or used in facilitating their research" (p. 582) and presented a behavior that they called information managing. This is an example of an information *use* as opposed to information seeking behavior. In our study of lawyers' IB, we identified several information use behaviors. The first of these behaviors (and the most similar to Meho and Tibbo's information managing behavior) is *recording*, which can also involve filing, archiving, and organizing information. However, although information managing is a rather broad term that can subsume several types of lower-level behavior, we believe recording is a more precise description of the behavior we observed. Recording involves making a record of information.

This information might be resources or sources used, documents or content found, or the query terms used or results returned in a search. Current electronic resources usually support users in recording documents but rarely support keeping a record of sources used. Many electronic legal resources also keep an automatic record of users' searches. Other information use behaviors identified by Makri et al. (2008) that form the basis of the IB methods are as follows:

- *Analysing*. This behavior was identified but not greatly elaborated on by Meho and Tibbo (2003). Adopting a definition based on the Oxford English Dictionary, analysing involves examining in detail the elements or structure of the content found during information seeking. Electronic resources rarely provide explicit support for analysing. This support might include the facility to make notes about particular documents or sources used and to arrange these notes under different categories, topics or headings.

Page numbers differ from published version

- *Collating* and *Editing.* Collating involves drawing together documents and/or content for later use. Editing involves preparing and arranging documents or content for later use by making revisions or adaptations. Some electronic legal resources support the downloading, printing or e-mailing of batches of documents (thereby collating them). However, electronic resources rarely support editing behavior (although some resources allow documents to be downloaded into an editable Word processing format).

Another information use behavior, *distributing*, was identified only among practicing lawyers. Distributing involves handing or sharing out documents, contents of documents or search queries/results to others. Some electronic legal resources provide the facility to e-mail documents to colleagues. However, in general, support for distributing behavior is rare.

A number of additional behaviors were also identified by Makri et al. (2008):

- *Searching* involves formulating a query in order to locate information. Searching behavior is not discussed at all by Ellis and his colleagues, perhaps due to the paper-based nature of information seeking in these studies. This behavior was, however, discussed briefly by Meho and Tibbo (2003). Searching is supported by electronic resources in a number of ways. One such way is through the provision of segmented search fields that restrict the search to particular parts of documents or document meta-data.
- *Selecting* involves carefully choosing resources, sources, or documents as being potentially useful for the information task at hand (definition adapted from Oxford English Dictionary). Selecting shares conceptual similarity to distinguishing, filtering, and extracting behaviors but it subtly different to these behaviors. It is different to distinguishing because it does not involve ranking resources based on perceived importance. It is also different to filtering because although various criteria are used when selecting which resource to use, these criteria are not used as precise, explicit filters to help decide between which document or source to select. Selecting is supported by electronic resources through the provision of document metadata (such as the title or year of publication of a document) in search result or browsable lists.
- *Updating* behavior, which we believe is particularly pertinent to information seeking in the legal domain, involves ensuring a current understanding of amendments or changes to a particular legal document. Updating with regard to a particular legal case or piece of legislation, for example, involves finding out whether the case or legislation is currently *good law.* Electronic legal resources support updating and history tracking by providing dedicated case and legislation citators which allow users to view the history of a case or piece of legislation.

The final IB that forms the basis of the IB methods is *history tracking*, which is similar to updating behavior but involves ensuring an *historical* as opposed to a current understanding of amendments or changes to a particular document (i.e., an understanding of the treatment a particular case or piece of legislation has received over time). Refer to Appendix A for a summary list of the behaviors that form the basis of the IB methods. In general, users' IB has been found to be similar across domains (see Ellis, 1989; Ellis, 1993; Ellis et al., 1993;

Ellis & Haugan, 1997) and has not been found to have changed substantially since Ellis's studies in the 1980's and 1990's even though electronic information seeking is now far more common (see Meho & Tibbo, 2003). Discrepancies between sets of findings are minor and have often been due to differences in terminology (as highlighted by Ellis et al., 1993) or differences in the scope of the research carried out (e.g., Meho and Tibbo, 2003 identified *information managing* behavior and Makri et al., 2008 identified additional *information use* behaviors in addition to a number of information seeking behaviors). Some discrepancies have, however, been due to differences in behaviors across domains. For example, verifying behavior was found among some disciplines but not others. This suggests that IBs in a new domain are likely to be similar but not necessarily identical to those identified among the lawyers in our study.

Although the methods were developed based on empirical data of *lawyers'* IB, we hypothesize that the IB methods can be used to evaluate both legal and non-legal resources. This hypothesis is based on evidence that IB has been found to be similar across domains (see Ellis, 1989; Ellis, 1993; Ellis et al., 1993; Ellis & Haugan, 1997) and has not been found to have changed substantially since Ellis's studies in the 1980s and 1990s even though electronic information seeking is now far more common (see Meho and Tibbo, 2003).

Indeed, the IB methods are extensible and customisable and can therefore be tailored to include any additional or alternative behaviors relevant to a new, non-legal, domain. For example, currently the behaviors used as the basis of the IB evaluation methods do not include behaviors identified in other empirical studies but not in our own (such as verifying and networking). However, it is possible to incorporate additional relevant behaviors into both an IB functionality and IB usability evaluation. For example, when evaluating the functionality of electronic resources designed to support physical scientists' information work, evaluators might wish to assess functionality support for verifying behavior, i.e., support for "checking the information and sources found for accuracy and errors" (Ellis et al., 1993, p. 364). Similarly, when conducting an IB usability evaluation of an electronic resource designed for physical scientists, it is possible to set one or more tasks focused on verifying behavior (e.g., check one of the documents you have found for accuracy and errors). It is therefore possible to customize the IB methods to new domains by including or excluding certain behaviors. In effect, this means that when applying the IB methods to a new domain, it is only necessary to change the *theory base* of the methods (i.e., the IBs to be used to frame the evaluation). It is not necessary to change the methods themselves. Indeed, we hypothesize that the IB methods can be applied to a wide range of domains with only slight modification to the theory base. As some identified IBs share conceptual similarities (e.g., selecting and distinguishing), care must be taken when customising the methods to ensure that all behaviors used as part of the theory base are clearly and precisely defined.

This should help minimize the chance of users of the methods misunderstanding what the behaviors entail (and therefore making functionality suggestions or defining usability tasks related to an incorrect understanding of a particular behavior).

The Behavioral 'Levels' that Form the Basis of the IB Methods

The IB methods are strongly underpinned by the IBs discussed above. They are also underpinned by the concept of levels at which these IBs can operate. Makri et al. (2008) identified five levels at which IBs can operate. These include the *resource* level (i.e., the level of the electronic resource itself), the *source* level (i.e., the level of an information source or sources within a particular electronic resource), the *document* level (i.e., the level of a document or documents within a particular information source), the *content* level (i.e., the level of content within a particular document), and the *search query/result* level. Four of these levels are illustrated in Figure 2, which highlights that an electronic resource can contain several sources that, in turn, can contain several documents, each with content within them. For example, LNB, a widely used electronic legal resource, contains many sources ranging from different series of legal case reports and legal journal articles to collections of different types of legislation such as Acts of Parliament and Statutory Instruments. Within each source are a number of documents (individual case reports, articles, pieces of legislation etc.), each with their own content. We do not illustrate the fifth level (the search query/result level) in Figure 2, but it can be regarded as the means of bridging each of the other levels (i.e., searching for content that is held in a particular electronic resource). Note that, particularly in the digital library community, the word resource is often used to describe both digital libraries themselves and electronic sources within a library (e.g., a particular journal series available within). In this article, we refer to resources and sources as separate and distinct entities.
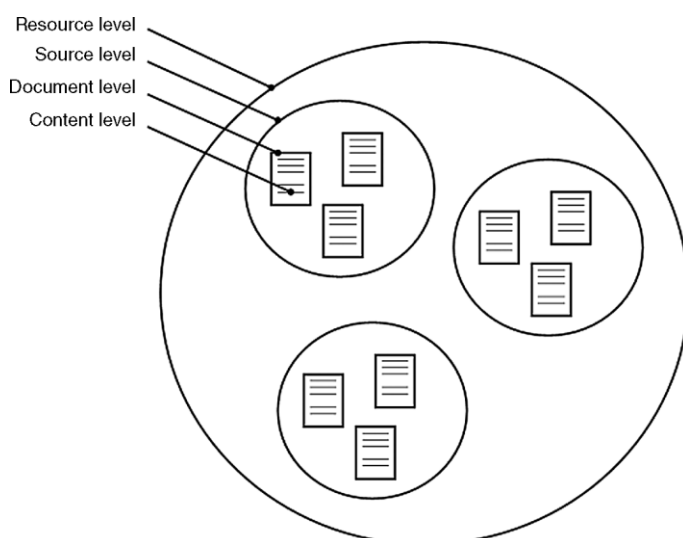

FIG. 2. Diagram to illustrate four of the levels at which many of the information behaviors can operate.

Page numbers differ from published version

The empirical findings presented in Makri et al. (2008) demonstrate that many of the behaviors identified by Ellis and his colleagues can be performed at multiple levels. For example, consider recording behavior (which involves making a record of information). Recording might be performed at the resource level by bookmarking or otherwise saving details of electronic resources. Recording might be performed at the source level by saving a list of favorite or frequently used sources within a particular resource. At the document level, recording might be performed by downloading, printing, or saving a particular document (or even e-mailing it to one's self). At the content level, parts or sections of documents can also be downloaded, printed, saved, or e-mailed. Finally, at the search query/result level, a manual or automatic record can be made of the search queries entered during a particular information seeking episode and possibly the results returned from those queries.

Some behaviors, such as surveying and monitoring, operate at a combined document and content level. This level is used when it is difficult, impossible, or undesirable to separate whether a particular behavior is performed on the document, or the content within it. For example, surveying and monitoring involve gaining an overview of and maintaining awareness of developments in a particular research area. This involves looking at both documents and the content of those documents. However, the resulting observable behavior at the document level is likely to be the same as the observable behavior at the content level and therefore only one combined level is used to describe this behavior. Although these levels may be applicable to paper-based information seeking (for example, it is possible to regard a paper volume of journal titles as a resource, an individual issue as a source, an article within an issue as a document, and the textual content of the article as content), this has not been empirically tested as the focus of Makri et al.'s study was on *electronic* not paper-based IB.

The concept of levels is important for structuring both IB functionality and usability evaluations. An IB functionality evaluation involves using the IBs discussed in this section to frame a functionality evaluation of an electronic resource by examining whether and in which ways each behavior is supported by the resource *at each applicable* level. Applicable levels are those at which the behavior was found to operate by Makri et al. (2008) and by subsequent research. An IB usability evaluation involves setting behavior-focused tasks to intended or actual users of an electronic resource, asking them to think aloud while performing the tasks (i.e., to verbalise their thoughts, actions and feelings), and analysing the resultant think-aloud data. Depending on the focus of the usability evaluation, these tasks may be aimed at performing a particular behavior *at a certain level.* For example, one of the possible tasks that can be set as part of a custom IB usability evaluation is to browse to see whether a particular source is available in the electronic resource (for example, by browsing through a list of sources that the resource contains). This task aims to encourage users to perform browsing at the source level as opposed to the more common document level (where the user browses to locate particular documents, not sources).

The IB Methods

We now present the IB methods. We begin with an introduction to usability and functionality evaluation in general and highlight the niche in the evaluation methods market that the IB methods address. We then briefly discuss the development and evaluation of the methods, followed by an overview of both the functionality and usability methods.

Introduction to Functionality and Usability Evaluation
Functionality and *usability* are two aspects of an interactive system that can be evaluated with the aim of suggesting design improvements. Evaluating the functionality of a resource involves in some way examining the features provided by the resource aimed at supporting users. Evaluating the usability of a resource essentially involves examining how easy it is to use it (the International Organization for Standardization defines it as the effectiveness, efficiency, and satisfaction with which specified users achieve specified goals in particular environments [ISO# 9241, www.iso.org]).

We were not surprised to learn that little has been written about functionality evaluation methods. Mack and Nielsen (1994) highlight that one type of functionality evaluation, known as a feature inspection, focuses "on the function delivered in a software system: for example, whether the function as designed meets the needs of intended end users" (p. 6). Mack and Nielsen (1994) also highlight that feature inspections can include design as well as evaluation of features within a system.

Page numbers differ from published version

However, there is almost no literature detailing existing feature inspection methods. This is with the exception of work by Bell (1992), who details a method for designing programming languages that are easy to write that is aimed at evaluating both the facility provided by a programming environment (which Bell describes as "the ability to solve problems easily" [p. 7]) and the expressiveness of the environment (which he describes as "the ability to state solutions to hard problems simply" [p. 7]).

The IB Methods' Place Alongside Other Evaluation Methods
Existing methods in the human–computer interaction (HCI) domain provide different ways to examine the usability of interactive systems and many of these methods have a specific focus which helps to make them useful in a specific way. For example, cognitive walkthrough (Polson, Lewis, Rieman, &Wharton, 1992; Wharton, Reiman, Lewis, & Polson, 1994) allows those using the method to assess the learnability of an interface, with a focus on the user's cognitive processes and perception. Similarly concept-based analysis of surface and structural misfits (CASSM; see Blandford, Green, Furniss, & Makri, 2008) is a method that can be used to highlight mismatches between how users conceptualize aspects of an interactive system and how the system supports user concepts. Some of these user evaluation methods have been applied to a digital library and electronic resource context. For example, Blandford, Keith, Connell, and Edwards (2004) used a range of evaluation methods (heuristic evaluation, cognitive walkthrough, claims analysis, and CASSM) to evaluate various digital libraries. Similarly, Blandford et al. (2008) present a worked example of how CASSM can be applied to think-aloud data of postgraduate HCI and library and information studies students using a range of electronic resources. Blandford and her colleagues have also tailored claims analysis to a digital libraries context (see Blandford, Keith, & Fields, 2006; Blandford, Keith, Fields, & Furniss, 2007).

As highlighted by Blandford and Green (2008), evaluation methods can be broadly classed along three dimensions:

- whether they are carried out with the active involvement of users (those that involve users are known as *empirical* methods and those that do not as *analytical* methods)
- whether they are carried out with a running system
- whether they are carried out in a realistic context of use

According to Nielsen (1993), usability evaluation with *real users* "is the most fundamental usability method and is in some sense irreplaceable, since it provides direct information about how people use computers and what their exact problems are with the concrete interface being tested" (p. 165). Similarly Landauer (1995) has described user testing as the gold standard for evaluation. However, user testing can be resource intensive and this has resulted in the use of evaluation methods that do not require user involvement (such as those described in the previous paragraph).The IB functionality method is not carried out with active end-user involvement per se; however, evaluators may seek to use the method armed with usage data of their resources in order to help them reason about increasing or even reducing functionality (both are discussed in detail later in this article). Therefore, this method can be classed as an analytical method. The IB usability method, on the other hand, is primarily an empirical method as it involves analysing think-aloud data of *users* performing certain IB-focused tasks. The IB usability method is also partly analytical, as it involves *evaluators* identifying usability issues from the think-aloud data and deciding on how severe and easy to address they are. The mixed nature of the IB usability method allows it to benefit from the use of rich user data (where it would otherwise be difficult for evaluators to predict how users are likely to behave with an electronic resource) and from a theoretical underpinning to drive task setting and analysis (where it would otherwise be difficult to analyse this rich data in a structured way). As the IB usability method involves observing real users performing behavior-focused tasks, it requires the use of a *running system*. An IB functionality evaluation can be supported by (but does not require) a running system, particularly as evaluators may be familiar with most of the functionality provided by the resource under evaluation. Although based on empirically observed behavior, an IB functionality evaluation does not involve a *realistic context of use* (as it does not actively involve users). On the other hand, although the think-aloud tasks that are part of an IB usability evaluation might be performed in an artificial setting, the tasks themselves (also based on empirically observed behavior) *do* aim to ensure a realistic context of use.

We also believe the IB methods address a clear niche in the market of evaluation methods. Although there are other methods that are underpinned by theory (such as Polson, Lewis, Rieman,&Wharton's [1992] cognitive walkthrough), we are unaware of any other evaluation methods with roots grounded in *information* theory. In addition, the IB methods are specialized in the sense that they aim to

evaluate the functionality and usability of *electronic resources* and not interactive systems in general. Although Blandford and her colleagues have applied and tailored various evaluation methods to an electronic resource context, we are unaware of any other evaluation methods *developed especially* to evaluate electronic resources (see Blandford, Keith, & Fields, 2006; Blandford, Green, Furniss, & Makri, 2008).We also believe that the IB methods are novel because they are empirically grounded (i.e., they are based on the electronic IB displayed by the academic and practicing lawyers in our study). The IB methods aim to provide a bridge between the domains of information science and HCI, by providing users with the opportunity to conduct functionality and usability evaluations that are highly structured but are also flexible in the sense that they can be tailored to particular domains or foci. Part of this structure is provided by the theoretical underpinning of the IBs and levels that provide a framework for assessing electronic resource functionality (as part of an IB functionality evaluation) and for setting behavior-focused tasks (as part of an IB usability evaluation). Part of the structure is also provided by supporting forms, which can be used to record the output of the evaluations (see Appendices 1 and 2 for examples of these forms). Flexibility is provided by the extensibility and customisability of the methods, which can be tailored to facilitate the evaluation of particular IBs or levels of interest.

Introduction to the IB methods

The IBs that underpin the IB methods feed in to the functionality and usability evaluations in different ways. They are used in an IB *functionality* evaluation as a framework for assessing the functionality provided by electronic resources. An IB functionality evaluation involves users of the method (i.e., evaluators) discussing whether and in what ways an electronic resource currently supports the IBs at the five levels described earlier. Where the resource under evaluation is being evaluated by people with a stake in it, an IB functionality evaluation also involves evaluators discussing whether it might be possible to support user behaviors/levels in additional ways and considering the arguments for and against supporting particular behaviors/levels.

The IBs are used in the IB *usability* method as the foundation of think-aloud tasks that are set to intended or actual users of the electronic resource (see Table 3 for a list of behavior-focused tasks derived from our empirical data).

In the usability component of the method, evaluators set a number of behavior-focused tasks to users who are asked to perform the tasks while thinking aloud. This involves the users verbalizing their thoughts, actions, and feelings while performing the tasks using the specified electronic resource (just as in a conventional think-aloud session). The evaluators then identify usability issues from the resultant think-aloud data and make summary judgements on how severe they consider the issues to be (i.e., whether they need immediate attention) and the amount of effort they consider to be required to address the issues. The process of identifying and making summary judgements on usability issues that are identified from think-aloud data is not unique to the IB usability method. Indeed, this is often a standard part of user testing. It is the theory-based, task-setting element that makes the IB usability method unique as the behavior-focused tasks are used to frame the think-aloud session – providing a structure to the tasks that aims to encourage the display of a broad range of IBs (or particular behaviors/levels of interest). The IB methods can be used by anyone with an interest or stake in an electronic resource. However, we recommend the methods are used by people with a basic grounding in usability evaluation and without a strong bias towards the existing design of the electronic resource under evaluation. This means that while the methods can, in theory, be used by the developers of a particular resource themselves, this is not advisable as it is likely to be difficult to avoid attachment to particular system functionality or other related issues (such as power relationships within the firm or the evaluation team itself). This is the case with many HCI evaluation methods. The process of planning and conducting an IB evaluation is framed by Blandford et al.'s (2008) PRET A Rapporter Framework (PRETAR)—a framework for structuring user centred evaluation studies, including evaluation studies of information retrieval systems. The broad process of conducting a functionality or usability IB evaluation are described below:

1. Defining the *purpose* and *boundaries* of the IB evaluation specifically involves deciding which (a) electronic resource to evaluate, (b) type of evaluation to carry out (i.e., a functionality evaluation, a usability evaluation or both), © parts of the resource to evaluate, and (d) behaviors to evaluate the resource in relation to. The size of the resource is likely to influence the decision on which parts to evaluate (it may not be possible to evaluate large resources in their entirety). The decision on which behaviors to evaluate the resource in relation to is likely to be influenced by the *focus* of the evaluation. For example, it may not be within the scope of the resource to support wider

Page numbers differ from published version

information-use behaviors, such as analysing, synthesising, recording, collating, editing, and distributing information. Therefore, in such cases, certain behaviors may be excluded from the evaluation.

2. Deciding on the *practicalities* of the evaluation involves considering issues such as when in the design and evaluation cycle the resource should be evaluated, who should participate in the evaluation, how much time should be devoted to the evaluation, and how the evaluation should be recorded.

3. Considering the *ethical issues* surrounding the evaluation involves considering issues surrounding keeping participant data as anonymous as possible and respecting participants' confidentiality and privacy. It may also involve considering how participant data will be used and, if applicable, disseminated.

4. *Conducting the evaluation itself and recording the output* is discussed in detail below for both the functionality and the usability methods.

5. *Communicating the findings* from the evaluation. There are many varied ways of communicating the findings of an IB evaluation, ranging from using them as the basis of formal reports to using them as a basis of informal presentations or discussions.

The IB Functionality Method

An IB functionality evaluation involves deciding whether the IBs, at each applicable level, are currently supported by the electronic resource and then:

- For levels of a behavior that the resource *currently supports*, determining in which way(s) the resource currently supports and in which additional ways it might support the behavior at this level.
- For levels of a behavior that the resource *does not currently support*, determining in which way(s) the resource might support the behavior at this level.

An IB functionality evaluation also involves asking the following general questions:

- Are there any behaviors/levels that it may *no longer be necessary to support*? For any behaviors/levels that you are considering ceasing support, what are the potential arguments for and against support?
- Are there any *ways that you currently support* any of the behaviors/levels that *may no longer be necessary*? Forways of supporting a particular behavior/level that you are considering ceasing support, what are the potential arguments for and against support?

The functionality evaluation can be supported by reference to or exploration of a running version of the resource under evaluation (whether this be a full running version or a limited functionality prototype). Table 1 lists the behaviors and each applicable level at which resource functionality should be assessed.

It will not always be appropriate to determine ways that an electronic resource *might* support particular behaviors/levels or to discuss the arguments for and against support. This is especially the case when evaluating resources where users of the method only have an *indirect* stake in the resource (for example, when evaluating an electronic resource developed by a competitor firm). In such cases, a "cut down" functionality evaluation can be conducted that involves exploring a running version of the resource to determine *whether it supports* the behaviors/levels in Table 1 and if so, *in which ways* it currently supports them. The form for recording the detailed output of an IB functionality evaluation is presented in Appendix A.

Page numbers differ from published version

TABLE 1. Behaviors and levels to be considered in an IB functionality evaluation.

	Resource level	Source level	Document level	Content level	Search query/result level
Accessing	Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)		
Surveying		Currently supported? (Y/N)	Currently supported? (Y/N)		
Monitoring		Currently supported? (Y/N)	Currently supported? (Y/N)		
Searching		Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)	
Browsing and extracting		Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)	
Chaining		Currently supported? (Y/N)	Currently supported? (Y/N)		
Selecting, distinguishing and filtering		Currently supported? (Y/N)	Currently supported? (Y/N)		
Updating			Currently supported? (Y/N)		
History Tracking			Currently supported? (Y/N)		
Analysing				Currently supported? (Y/N)	
Recording		Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)
Collating			Currently supported? (Y/N)	Currently supported? (Y/N)	
Editing			Currently supported? (Y/N)		
Distributing			Currently supported? (Y/N)	Currently supported? (Y/N)	Currently supported? (Y/N)

Key: Core information-seeking behaviors ☐   Law-specific behaviors ☐   Information use behaviors ☐

The IB usability method.

An IB usability evaluation involves asking intended or actual users of an electronic resource to think aloud while using the resource to perform a number of tasks and analysing the resultant think-aloud data. This essentially involves conducting a conventional think-aloud session where users are asked to verbalize their thoughts, actions, and feelings and their verbal protocols are recorded (with minimal researcher intervention). The general think-aloud process is documented in both HCI textbooks (such as Dumas & Redish, 1999) and in articles (such as Boren & Ramey, 2000). Both present guidelines for conducting think-aloud sessions. An IB usability evaluation differs from a conventional HCI think-aloud study in that it is *solely based on the behavior-focused tasks* that users are asked to perform. There are three sets of tasks—core, recommended, and custom—, and it is possible to ask users to perform them all in an IB usability evaluation. In a *core* IB usability evaluation, participants are asked to perform two tasks related to accessing the electronic resource under evaluation and then the broad (and only somewhat behavior-focused) task of finding information currently or recently needed for their work. This task is the same as the one set to participants in our empirical study of lawyers' IB (see Makri et al, 2008). The broad nature of this task encourages (but does not guarantee) the display of a wide range of IBs. Although a *core* IB usability evaluation is highly naturalistic, it does not, however, encourage the display of *particular* behaviors. We recommend a core IB usability evaluation as a "quick and dirty" way of acquiring user think-aloud data that highlights usability issues. The tasks set as part of a *recommended* evaluation are more behavior-focused than the *core* tasks but not as naturalistic.

They are based on *common ways* that the lawyers in our empirical study performed the full range of IBs that were identified. A recommended IB usability evaluation is a way of acquiring rich and behavior-focused think-aloud data, again with the potential to highlight usability issues. We advise conducting a recommended evaluation in most cases. The tasks set as part of a *custom* evaluation are more prescriptive and less broad than those set in a core or recommended evaluation. This makes them highly focused on particular IBs but only somewhat naturalistic. These tasks are based on ways that the lawyers in our study performed particular behaviors that were less common. Some of the custom tasks also encourage the display of behaviors at particular levels that were less commonly displayed by the lawyers in our study (i.e., the source, content and search query/result levels). Core tasks should always feature in an IB usability evaluation. Recommended tasks should feature alongside the core tasks unless financial and/or resource constraints make this impossible. Custom tasks should be used to tailor IB usability evaluation where there is a need to focus on particular

Page numbers differ from published version

behaviors/levels at which the behaviors can operate. It is therefore possible to mix and match custom tasks that aim to encourage demonstration of particular behaviors/levels of interest.

In a core IB usability evaluation, participants are asked to perform the three information seeking tasks in Table 2:

TABLE 2. The three information-seeking tasks that think-aloud participants are asked to perform as part of a core IB usability evaluation.

'Core' tasks in a IB usability evaluation:
1. Gain access to the electronic resource.
2. Find out which parts or sources within the resource you have access to.
3. Think of some information that you currently need or have recently needed to find for your work and demonstrate, using the electronic resource, how you might go about finding it.

In a recommended evaluation, participants are asked to perform the three core tasks listed above, *plus any of the tasks in Table 3 that are currently supported by the electronic resource*. The recommended tasks are based on the full range of behaviors that were displayed by the lawyers in our empirical study (and were found to be commonly supported by our survey of the functionality of electronic legal resources). These tasks are, therefore, designed to encourage the demonstration of a broad range of behaviors. We use the word "encourage" as it is not always possible to predict exactly how participants will carry out the tasks. Note that the recommended tasks only encourage demonstration of behavior at the *document* level as this was the most common level at which they were displayed in our empirical study. Also note that some of the tasks in Tables 3 and 4 have been customized for evaluating *legal* resources, however similar tasks can be set to evaluate resources from other domains.

TABLE 3. Tasks that think-aloud participants are asked to perform as part of a recommended IB usability evaluation. As well as these tasks, participants are also asked to perform the three core tasks in Table 2.

'Recommended' tasks in a IB usability evaluation:

Gain an overview of an area by:
- Trying to gain a basic understanding of the law relating to a particular legal area (e.g. Breaches of contract).
- Trying to gain an appreciation of the importance of a certain legal journal author's role in a particular legal area.
- Trying to locate a legal journal article written by an author who has published many articles or many highly cited articles.

Gain a current or historical understanding of the importance of a document by:
- Trying to find out whether a particular case is still good law.
- Trying to find out what amendments have been made to a particular piece of legislation over a certain time period.
- Trying to find out whether a particular piece of legislation is currently in force.
- Trying to locate a historical version of a particular piece of legislation (i.e. a previous version that has since been amended).

Maintain awareness of developments in an area by:
- Trying to find out whether there have been any recent developments in a particular legal area (e.g. Discrimination law).
- Trying to set up an alert so that you can be informed every time new documents are added to the system that match particular search terms (e.g. when new documents that match the term 'discrimination' are added).
- Trying to set up an alert so that you can be informed every time there are new developments in a particular legal area.

Return to any one of the tasks where you found useful documents and:
- Determine which sections of a document that you have found are important to you.
- Keep a softcopy (downloaded or saved) record of a document that you have found.
- Keep a hardcopy (printed) record of a document that you have found.
- Download two documents into a single file (i.e. you should end up with one file saved on the computer that includes the text of two separate documents, e.g. two different legal journal articles or two different sections of a particular piece of legislation).
- Keep a softcopy or hardcopy record of *part* of a document that is important to you (e.g. print or download only certain parts of a case report).
- Distribute a document that you have found, by e-mail, to a fictitious colleague from within the electronic resource.
- Store a document on the server of the resource (i.e. save a copy of the document to a personalised area on the electronic resource itself, so you can access it again quicker in future).

Page numbers differ from published version

TABLE 4. Custom tasks related to chaining behavior.

Custom tasks related to 'chaining' behavior:

Try to follow a hyperlink or other form of connection from:
- A legal case to a previous case or a particular piece of legislation mentioned in the case report.
- A piece of legislation to other pieces of legislation mentioned in the text of the Act.
- A legal journal or commentary article to a case or piece of legislation mentioned in the article.

Try to follow a hyperlink or other form of connection from a document to other documents that have been written since this document and have subsequently mentioned it. Specifically, try to:
- Find a particular case, then find out which more recent cases have mentioned it (if any).
- Find a particular piece of legislation, then find out which more recent pieces of legislation have mentioned it (if any).
- Find a particular legal journal article, then find out which more recent articles have mentioned it (if any).

In a custom IB usability evaluation, the choice of tasks that users are asked to perform will vary depending on the focus of the evaluation. These tasks are designed to encourage the demonstration of specific, individual IBs. Table 4 lists possible custom tasks relating to *chaining* behavior (i.e., "following chains of citations or other forms of referential connections between material" Ellis, 1989, p. 179). In a custom evaluation, tasks can also be set to encourage behaviors at a particular level. For example, three tasks relating to surveying at the document level are presented as part of the set of recommended tasks in Table 3 (listed under the "get an overview of an area by" heading). However, it is possible to an additional (or alternative) task related to surveying sources, such as "try to found out which sources contain information about a particular legal area."

In an IB usability evaluation, users are audio and screen recorded as they think-aloud while performing the information tasks and the resultant think-aloud data is reviewed in order to identify usability issues. Although reviewing the think-aloud data, the evaluator(s) keep a record of any user actions, user comments, or personal observations that might suggest a usability issue and makes a note of what they believe to be the underlying usability issues identified from the actions/ comments/ observations. The evaluator(s) also record details of the screen(s)/page(s)/part(s) of the resource that the actions, comments, or observations relate to and makes judgements on the severity and amount of effort required to address the usability issue. The form for recording the detailed output of an IB usability evaluation is presented in Appendix B. Electronic versions of these forms, along with other support material for conducting functionality and usability IB evaluations, are available from our Web site, www.uclic.ucl.ac.uk/people/s.makri/IBMethods.html. This includes a comprehensive tutorial, detailed examples of how a range of electronic legal resources currently support each IB at the applicable levels, and a full list of custom tasks related to a range of behaviors/levels (not just chaining tasks as presented in Table 4).

Development and Evaluation of the Methods
The IB methods were developed and tested iteratively over a period of several months. This process included a set of user pilots aimed at determining the tasks and procedural details to be used during an IB usability evaluation and a pilot think-aloud data analysis session with an electronic resource developer. The process also included an informal pilot of the functionality method and semi-structured interviews with a digital library developer aimed at finding out how usable and useful he deemed aspects of the method to be, how likely he was to use the method in future, and what improvements he would suggest for the method. Also, although not the focus of this article, we have also recently conducted a formative evaluation of both methods with a group of electronic legal resource stakeholders working for LNB in the UK. In general, the stakeholders found both methods to be useful, usable, and easy to learn. Feedback was also generally positive regarding whether they might use the methods in future. The evaluation session also suggested useful ways that the methods might be improved. Most of these suggestions have been incorporated into our discussion of the methods in this article (such as the need to consider functionality reduction at a summary level as opposed to for every applicable behavior/level). We aim to examine the feasibility of the remaining suggestions (such as creating electronic versions of the forms used) in the near future.

Worked Examples of Carrying Out a Functionality and Usability IB Evaluation

We now present two short worked examples to illustrate the use of the functionality and usability methods. The examples involve evaluating LNB, an electronic resource widely used by legal professionals worldwide. The *purpose* of both the functionality and usability evaluations was to illustrate the use of the methods to an audience likely to be unfamiliar with it. As this resource has many customizable options (e.g., dedicated search screens for various legal practice areas), one

Page numbers differ from published version

boundary that we set was to evaluate only those parts of the system available as default in the publicly available, uncustomized, and academic version of the resource, which we accessed in late 2007. As we did not have a direct stake in LNB, there was no need to consider *practicalities* such as the amount of time to devote to the evaluations or the place of the evaluations in the product's development cycle (we only had access to the publicly available version of the resource). We also had access to a single evaluator, the lead author, and to several audio and screen recordings of lawyers using LNB to perform a range of the tasks that feature in an IB usability evaluation.We chose to use a small clip of a trainee solicitor performing updating behavior as this clip illustrated a number of potential usability issues. The *ethical issues* surrounding the evaluation mainly involved gaining permission from LNB to evaluate their resource and to report the evaluation, complete with screenshots, in this article. It was also important to obtain permission from the trainee solicitor to publish the data arising from his think-aloud session. This data is presented in Appendix C. We communicate the findings from the functionality and usability IB evaluations below.

Example IB Functionality Evaluation of an Electronic Legal Resource
To illustrate part of an IB functionality evaluation, we evaluate LNB in relation to browsing and extracting behaviors. Browsing involves 'semi-directed searching for sources, documents or content.' Extracting can often work hand-in hand with browsing and involves 'systematically working through a particular resource to identify sources of interest, a particular source to identify documents of interest and/or a particular document to identify content of interest.' As well as from the definitions, it can be noted in Table 1 that browsing and extracting can operate at three levels: source, document, and content. As a reminder, browsing and extracting behaviors are presented together in Table 1 because they are intended to be analysed together. We now evaluate LNB in relation to browsing and extracting behaviors for each of these three levels. An IB functionality evaluation usually involves asking ourselves in which ways an electronic resource supports a certain behavior or behaviors at a particular level and in which additional ways might it support them. It also usually involves considering whether there are any behaviors, levels, or ways of supporting them that *may no longer be necessary.* We do not discuss potential functionality reduction and arguments for and against support in this example, as we feel we would only be able to do so if we had a direct stake in the resource (and were armed with knowledge about the use of the various parts of the resource's functionality). We do, however, discuss additional ways in which the resource might support each behavior/level (even though this is not normally necessary if the evaluator does not have a direct stake in the resource under evaluation). However, we do so purely to illustrate how the IB functionality method can lead to suggestions for additional functionality support. See Appendix A for a form that outlines the questions to be considered in different circumstances as part of an IB functionality evaluation. This form can also be used to help record the output of the evaluation. We now turn to ask in which ways the resource *currently supports* and in which *additional ways* it might support browsing and extracting behaviors at each of the source, document and content levels.

In which way(s) does the resource *currently support* browsing and extracting at the *source* level? It is possible to browse to locate (and extract) sources in LNB by using the dedicated browse functionality provided by the resource.

As illustrated by the radio buttons labelled *a* in Figure 3, it is possible to browse for and extract sources in a number of ways. The first two are by publication type (this is illustrated in the part of Figure 3 labelled *b*, where the publication type *legal journals* has been selected and the relevant sources displayed) and by area of law (i.e., the legal areas that the source is deemed to cover). The other two allow browsing and extracting of sources that contain business and news related and industry related materials. As illustrated by the drop-down combo boxes in Figure 3 (labelled *c)*, it is also possible to filter the list of sources to include only those that fit specified criteria (e.g., sources covering only UK law) and then browse the filtered list of sources.
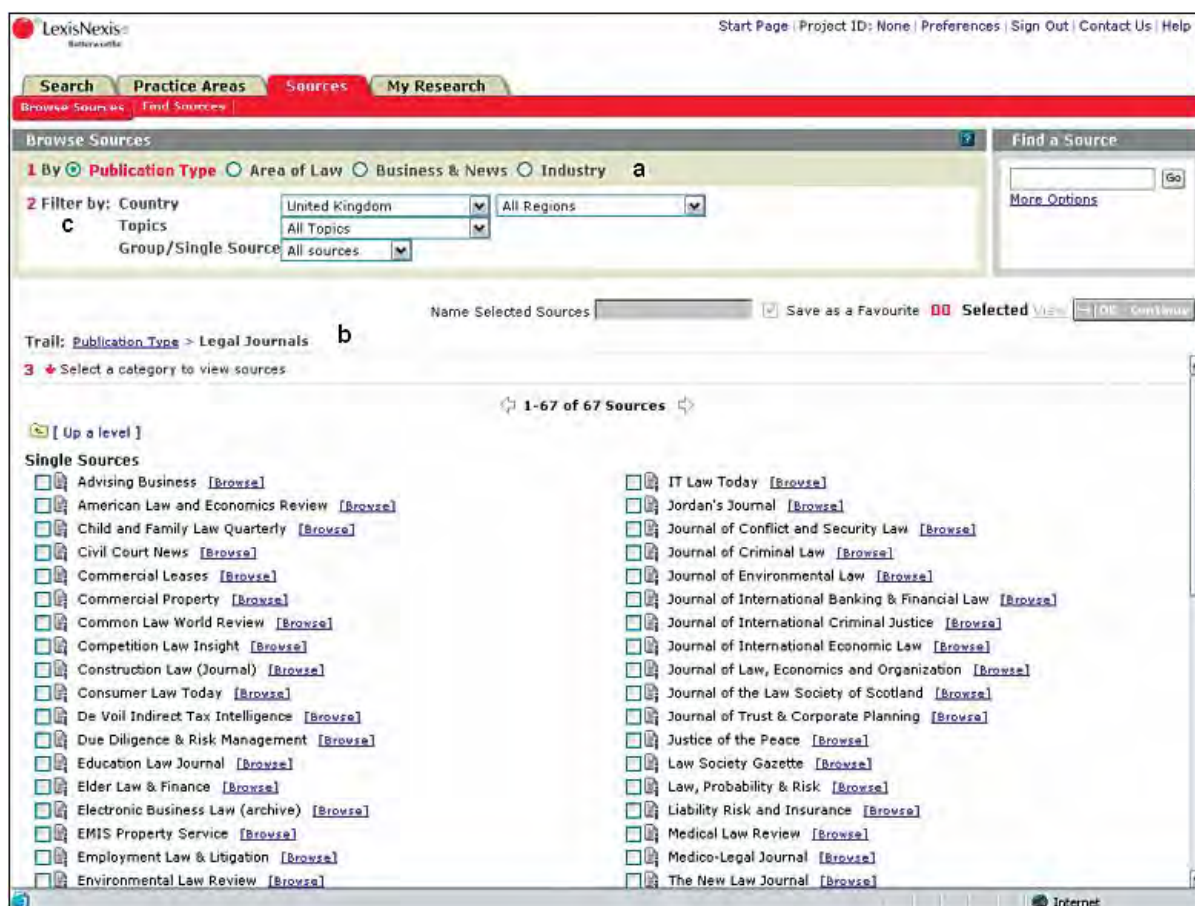
Page numbers differ from published version

LexisNexis
Butterworths

Start Page | Project ID: None | Preferences | Sign Out | Contact Us | Help

Search | Practice Areas | **Sources** | My Research
Browse Sources | Find Sources

Browse Sources   Find a Source

1 By ● Publication Type ○ Area of Law ○ Business & News ○ Industry   a   [Go]
                                                                         More Options
2 Filter by: Country   United Kingdom ▼   All Regions ▼
   c   Topics   All Topics ▼
           Group/Single Source All sources ▼

Name Selected Sources [   ] ☑ Save as a Favourite 00 **Selected** View →| OK Continue

Trail: Publication Type > Legal Journals   b
3 ↓ Select a category to view sources

                         ⇔ 1-67 of 67 Sources ⇒

⤴ [Up a level]
Single Sources
☐ Advising Business [Browse]   ☐ IT Law Today [Browse]
☐ American Law and Economics Review [Browse]   ☐ Jordan's Journal [Browse]
☐ Child and Family Law Quarterly [Browse]   ☐ Journal of Conflict and Security Law [Browse]
☐ Civil Court News [Browse]   ☐ Journal of Criminal Law [Browse]
☐ Commercial Leases [Browse]   ☐ Journal of Environmental Law [Browse]
☐ Commercial Property [Browse]   ☐ Journal of International Banking & Financial Law [Browse]
☐ Common Law World Review [Browse]   ☐ Journal of International Criminal Justice [Browse]
☐ Competition Law Insight [Browse]   ☐ Journal of International Economic Law [Browse]
☐ Construction Law (Journal) [Browse]   ☐ Journal of Law, Economics and Organization [Browse]
☐ Consumer Law Today [Browse]   ☐ Journal of the Law Society of Scotland [Browse]
☐ De Voil Indirect Tax Intelligence [Browse]   ☐ Journal of Trust & Corporate Planning [Browse]
☐ Due Diligence & Risk Management [Browse]   ☐ Justice of the Peace [Browse]
☐ Education Law Journal [Browse]   ☐ Law Society Gazette [Browse]
☐ Elder Law & Finance [Browse]   ☐ Law, Probability & Risk [Browse]
☐ Electronic Business Law (archive) [Browse]   ☐ Liability Risk and Insurance [Browse]
☐ EMIS Property Service [Browse]   ☐ Medical Law Review [Browse]
☐ Employment Law & Litigation [Browse]   ☐ Medico-Legal Journal [Browse]
☐ Environmental Law Review [Browse]   ☐ The New Law Journal [Browse]

                                                                    ● Internet

FIG. 3. Browsable list of sources in LexisNexis Butterworths, listed by 'publication type.'

In which *additional ways* might the resource support browsing and extracting at the *source* level?

Although LNB provides comprehensive support for browsing and extracting at the source level, the resource might support this behavior/ level further by allowing legal professionals to browse by other aspects of source meta-data that they deem to be important. For example, by when the first or most recent document in the source was published (i.e., browsing by source coverage dates).

In which way(s) does the resource *currently support* browsing and extracting at the *document* level?

As well as supporting browsing and extracting at the source level, LNB also provides comprehensive support for browsing sources in order to extract documents. Clicking on the "browse" hyperlink next to any of the listed sources in Figure 3 allows users to view documents contained within the source (albeit after drilling down several more levels). Another way that LNB supports browsing and extracting at the document level is illustrated by the "browse TOC/index" sidebar in Figure 4 (labelled *a*). This sidebar, presented adjacent to the full-text of documents in LNB, facilitates browsing other documents within the currently selected source (Journal of International Economic Law in Figure 4). This is achieved by clicking on one of the article titles listed in the expanded tree. It is also possible to use the "TOC/index" sidebar to browse for and extract journal articles in other issues of the selected journal series.
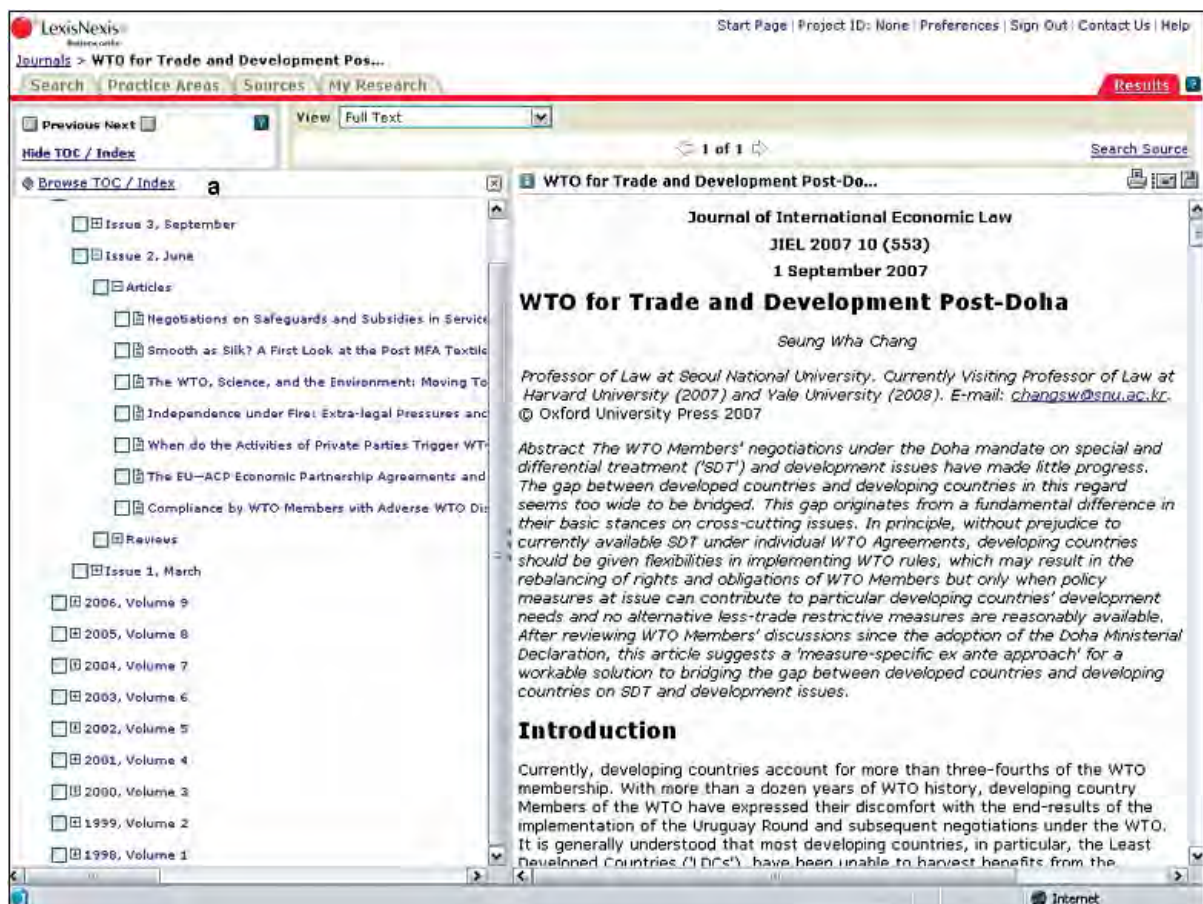
FIG. 4. Table of Contents/Index sidebar in LexisNexis Butterworths, which allows users to browse documents from the current source.

In which *additional ways* might the resource support browsing and extracting at the *document* level?

Despite strong coverage for browsing and extracting at the document level, there are additional ways that these behaviors might be supported at this level. Figure 5 illustrates the keywords listed at the top of a legal case in LNB. The resource currently facilitates browsing to other documents listed as part of the keywords (in the case of Figure 5, the Employment Rights Act 1996). However, there is also scope to allow browsing and extracting by keyword and other items of document metadata.

FIG. 5. Keywords listed at the top of a legal case report in LexisNexis Butterworths.

Page numbers differ from published version

LexisNexis®
Butterworths

Start Page | Project ID: None | Preferences | Sign Out | Contact Us | Help

Legislation - All Subscribed Legi... > /../ > (Secretary of State) (39) > Document (2/39)

Search | Practice Areas | Sources | My Research
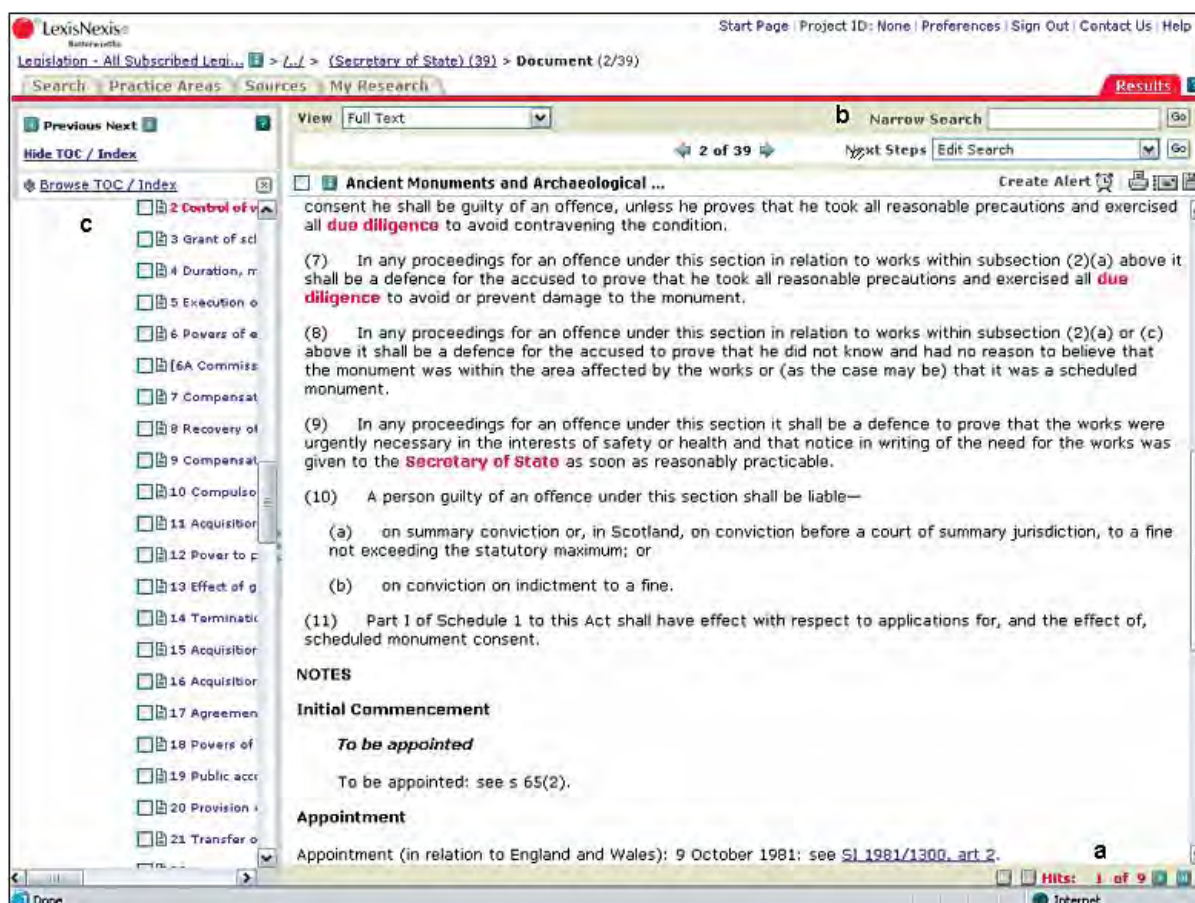
Results

View Full Text

b Narrow Search [ ] Go

2 of 39   Next Steps Edit Search [ ] Go

☐ 📄 Ancient Monuments and Archaeological ...   Create Alert 🔔 🖨 📧 💾

consent he shall be guilty of an offence, unless he proves that he took all reasonable precautions and exercised all **due diligence** to avoid contravening the condition.

(7)   In any proceedings for an offence under this section in relation to works within subsection (2)(a) above it shall be a defence for the accused to prove that he took all reasonable precautions and exercised all **due diligence** to avoid or prevent damage to the monument.

(8)   In any proceedings for an offence under this section in relation to works within subsection (2)(a) or (c) above it shall be a defence for the accused to prove that he did not know and had no reason to believe that the monument was within the area affected by the works or (as the case may be) that it was a scheduled monument.

(9)   In any proceedings for an offence under this section it shall be a defence to prove that the works were urgently necessary in the interests of safety or health and that notice in writing of the need for the works was given to the **Secretary of State** as soon as reasonably practicable.

(10)   A person guilty of an offence under this section shall be liable—

(a)   on summary conviction or, in Scotland, on conviction before a court of summary jurisdiction, to a fine not exceeding the statutory maximum; or

(b)   on conviction on indictment to a fine.

(11)   Part I of Schedule 1 to this Act shall have effect with respect to applications for, and the effect of, scheduled monument consent.

NOTES

Initial Commencement

To be appointed

To be appointed: see s 65(2).

Appointment

Appointment (in relation to England and Wales): 9 October 1981; see SI 1981/1300, art 2.   **a**

Hits: 1 of 9

Done   🌐 Internet

c

☐ 📄 2 Control of w
☐ 📄 3 Grant of scl
☐ 📄 4 Duration, m
☐ 📄 5 Execution o
☐ 📄 6 Powers of e
☐ 📄 [6A Commiss
☐ 📄 7 Compensat
☐ 📄 8 Recovery of
☐ 📄 9 Compensat
☐ 📄 10 Compulso
☐ 📄 11 Acquisitior
☐ 📄 12 Power to p
☐ 📄 13 Effect of g
☐ 📄 14 Terminatio
☐ 📄 15 Acquisitior
☐ 📄 16 Acquisitior
☐ 📄 17 Agreemen
☐ 📄 18 Powers of
☐ 📄 19 Public acc
☐ 📄 20 Provision
☐ 📄 21 Transfer o

Previous Next
Hide TOC / Index
Browse TOC / Index

FIG. 6. Document in LexisNexis Butterworths with search terms automatically highlighted in bold.

For example, it may be possible to allow users to click on a keyword and be presented with other documents that are related to the keyword (whether those documents be cases, or other types of document such as legislation or journal articles). There may also be scope to provide, for legal case reports, the facility to browse to all other cases involving a particular party in the currently displayed case (for example, Secretary of State for Trade and Industry in Figure 5), all other cases with the same judge presiding or all other cases with the same counsel. Similarly, for legal journal articles, there may be scope to find all other articles written by the same author as the currently displayed article. In which way(s) does the resource *currently support* browsing and extracting at the *content* level? As with the other levels, LNB also provides a considerable amount of support for browsing and extracting at the *content* level. As illustrated in Figure 6, users' search terms are automatically highlighted in bold in the full-text of a document, and it is possible to cycle through each occurrence of the search terms by clicking on the small arrows in the bottom-right-hand-corner of the screen (labelled *a)*.

In which *additional ways* might the resource support browsing and extracting at the *content* level? Also as illustrated in Figure 6, it is possible to enter additional search terms in the "narrow search" field (labelled *b)*. This serves to highlight additional words or phrases in the document text. There is scope to provide similar functionality for highlighting particular words or phrases in the current document without having to refine the original search. Similarly Figure 6 illustrates the TOC/index sidebar (labelled *c)*, which allows users to jump to particular documents within the current source (in the case of Figure 6, users can jump to other sections of the currently displayed Act of Parliament). There is scope, however, to extend this functionality to also allow users to jump to particular *section headings within the current document* (for example, to jump to the notes section heading in Figure 6).

As we have previously highlighted, the suggestions for increasing the range of functionality supported by LNB are to be treated as illustrative as they are not grounded in marketing or usability data. Therefore, it follows that developers at LNB, armed with such data, may either decide that the functionality serves a useful purpose and make the resource more usable or adds unnecessary complexity to the resource. It also follows that increasing the range of functionality supported will not

Page numbers differ from published version

necessarily lead to a holistically more usable resource, and care must be taken to ensure that a balance is struck between the functionality supported by the resource (and the resultant complexity arising from this support) and the overall usability of the resource. This provides a case for conducting both functionality and usability evaluations of resources. In the next section, we shift our focus from functionality to usability and present an example of an IB usability evaluation of LNB.

Example IB Usability Evaluation of an Electronic Legal Resource
We now present another example, this time of analysing think-aloud data as part of an IB *usability* evaluation. Note that although the process of conducting an IB usability evaluation also involves setting tasks to users and collecting the data, we discuss only the data *analysis* process in this example. The think-aloud data is based on a trainee solicitor who worked for a small London law firm, using the same electronic resource as in the previous example (LNB). The trainee was asked to use the resource to perform one of the tasks that is set as part of a recommended IB usability evaluation related to updating behavior. This task was to try and find out if a particular case is still good law. In order to complete the updating task, the trainee chose to examine a case called White vs. White.
The trainee had used LNB only a few times before as part of a legal practice course. The trainee's novice status proved useful in yielding usability data that we hypothesize might not have been obtained from a more experienced user.

In this example, we provide a description of how the trainee solicitor went about the updating task. This serves as a summary of the written transcript of the trainee performing the task, presented in Appendix C. Throughout our description, we also discuss the usability issues arising from the actions and comments made by the trainee while performing the task. The usability data is summarized on the IB usability form presented in Appendix B. The usability data on the form can be cross-referenced to highlighted sections of the transcript. Each highlighted section of the transcript in Appendix C is marked with a letter, which corresponds to a row (i.e., a usability issue) on the usability form in Appendix B.

It is important to note that there is an element of subjectivity involved in identifying usability issues related to complex electronic resources such as LNB. Therefore, different evaluators are likely to identify different usability issues (and potentially differ on their subjective ratings of the severity and ease of addressing each issue). This "evaluator effect" is evidenced by Hertzum and Jacobsen (2001), who asked four evaluators to select the ten *most severe* from those they had identified but found that none of the *severe* issues appeared on all four evaluators' list of top-ten issues. It therefore follows that this example evaluation should not be regarded as the sole or most authoritative interpretation of the trainee solicitor's think-aloud data but as *one possible interpretation* of the data.

Description of How the Trainee Solicitor Went About Performing the Updating Task and Discussion of the Resultant Usability Data
The trainee solicitor began by stating that he wanted to find out whether an ancillary relief case, fought by the parties White and White, was still good law. He clicked on the "cases" tab and mentioned that he was about to type the party names in the "enter search terms" field, as opposed to the "case name" field (which he soon realized would have been a more appropriate choice). This suggests a potential usability issue as the case search page did not make it immediately clear to the user in which segmented field to enter their search query. The trainee's comments and actions related to this issue are marked with the letter *a* in the full transcript of the think-aloud session in Appendix C and in Table 5 (which presents only the comments and actions that suggest a usability issue, not the full transcript). Each usability issue is also summarized in a row on the usability form in Appendix B.
As the trainee noticed the correct field within a couple of seconds, this issue is not deemed to be severe. Also, as the trainee quickly noticed the error, this does not suggest that the segmented fields are poorly labelled. As it is not immediately clear what caused the trainee difficulties, this issue would likely take a large amount of effort to address.

The trainee continued by typing "White vW" in the "case name" segmented field, then read a caption underneath the field that instructs users on the correct syntax to use when searching for cases (i.e., *and* instead of *v)*. Although this suggests that the required syntax was not made immediately clear (letter *b* in Appendix B and Table 5), this issue is also non-severe. It could be addressed with little effort by allowing search fields to intelligently accept a variety of syntax. Next, the trainee conducted the search and, upon receiving 1,700 results, filtered the results set to only display legal case reports and again to display only cases from the Family Court Reports source. This brought up the required

White vs. White case, and the trainee began by reading part of the case aloud. The trainee then wondered how [he could] check to see if the case has been updated or not and clicked first on the "view" drop-down combo box, reading the options aloud, and then on the "next steps" combo box (pictured in Figure 7). Again, he chose to read the options aloud and selected the "find related cases" option to see what that brings up. This suggests the actions facilitated by the "next steps" combo may not be as transparent as they could be, an issue that we deem to be quite severe and moderately difficult to address. This relates to letter *c* in Appendix B and Table 5.

TABLE 5. Comments and actions which suggest a usability issue, extracted from the full-transcript in appendix 4 of a trainee solicitor "trying to find out whether a particular case is still good law."

a	I'm gonna click on 'cases' and under 'enter search terms' I'm going to put [pauses]. No, I'm going to put it in the 'case name.'
b	[Participant types in 'White v W,' then pauses]. [Reads caption underneath 'case name' field]. 'To find Smith v Jones, enter Smith and [emphasis placed] Jones.' So I'm going to use the word 'and' rather than 'v.'
c	Here we go. 'Save search,' 'create alert,' 'search the source,' 'find related cases.' Hmm. I'll click on 'find related cases' and see what that brings up.
d	'How do I work with my search results?' I'll click on that as it might help me in some way. [Reads out sub-headings]. No, that doesn't help me. Let's search. [Searches within TotalHelp for 'current cases' and clicks on a result entitled 'results page' then soon after closes the help page]. I've closed that down because it just seems to be frustrating me.
e	Ok, the tutorial seems very slow and cumbersome, so [closes tutorial window]
f/g	[Enters 'white and white' into the 'get a specific document' part of the LexisNexis Butterworths homepage with 'find a case' selected in the combo box. Participant then presses the grey 'go' button next to the field he has filled out, but after a few seconds presses the main red 'search' button. [An information box is displayed, which says 'Please enter a search term'].
h	[Participant types 'White' in 'citation' field]. I'm not looking for a citation. [Participant then enters 'White and White' in the case name field].
i	View [clicks on 'view' combo box again and reads out options]. 'List,' 'expanded list,' 'White v White 2000.' [Selects White v White 2000]. Oops, I clicked on something.

The trainee was then presented with a list of "related cases," which he scrolled through. He then selected one of the cases in the list, "Wood vs. Rost," and noticed a reference and hyperlink to the "White vs. White" case, which he selected. Feeling, as he put it, "a bit stuck," he clicked on "help" but did not find anything that would assist him to perform the task at hand. After a few minutes, he closed the help page as it "[seemed] to be frustrating" him. He then tried viewing a tutorial, but after a few minutes closed the tutorial too, commenting that it "seems very slow and cumbersome." These are both usability issues that we deem to be quite severe if only because they caused the trainee frustration. We deem them to be moderately difficult to address as although updating tasks are important for lawyers, it is a considerable challenge for help systems to provide the required assistance in a useful format without patronising or frustrating users. These issues relate to letters *d* and *e* in Appendix B and Table 5.

After closing the tutorial screen, the trainee decided to try searching again and selected "find a case" from the "get a specific document" combo box (see Figure 8). However, the trainee did not notice the "go" submit button turn grey after he submitted the search and, possibly assuming he had clicked on the wrong submit button, clicked on a "search" button that was not related to the current search field. Figure 8 illustrates the trainee clicking on the unrelated "search" button and receiving a popup error message. This suggests that the relationship between search fields and their associated search buttons may not be as clear as it could be. This usability issue (letter *f*) is, in our opinion, very severe and would require moderate effort to address (perhaps by hiding or greying-out submit buttons that are not relevant to the search fields currently in use). The trainee not noticing the "go" button turning grey can be regarded as a separate usability issue (letter *g)*, associated with the resource providing potentially unclear feedback. We deem this issue to be quite severe and suggest it would require moderate effort to address. The trainee then dismissed the popup box, returned to the "case search" page, and briefly entered a party name in the "citation" field before correcting the error (which was identical to an error made towards the beginning of the task—letter *h*). Like related usability issue *a*, we deem this issue to be non-severe but difficult to address. Next, the trainee ticked "House of Lords" in the "court" selection box, and after conducting the search, filtered the results in the same way as earlier in the task, briefly pausing to tick the tick-box beside the "White vs. White" case and press "view tagged' to check his hypothesis that it would "be the same as if [he] just clicked on it." The trainee's hypothesis was confirmed when the full-text of the White vs. White case was displayed. The trainee then read the "next steps" combo box options aloud for the second time and then selected "White v. White 2000" from the "view" combo, which re-loaded the full-text of the case. This seemed to confuse the trainee, who stated "oops, I clicked on something." This suggests that the effect of selecting the current document from the "view" combo may be unclear. This is a usability issue (letter *l*) that we deem to be non-severe and moderately difficult to address.
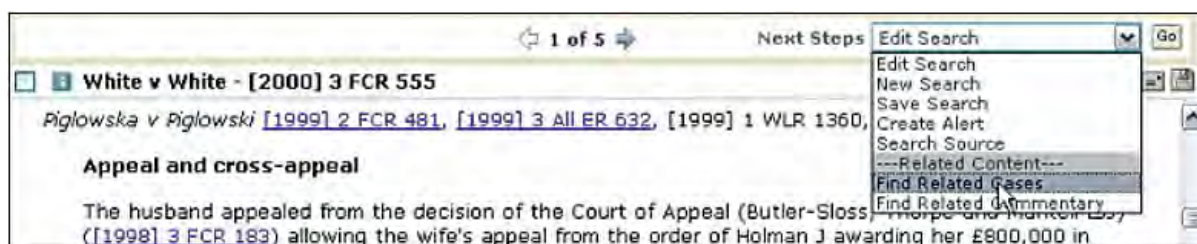
Page numbers differ from published version

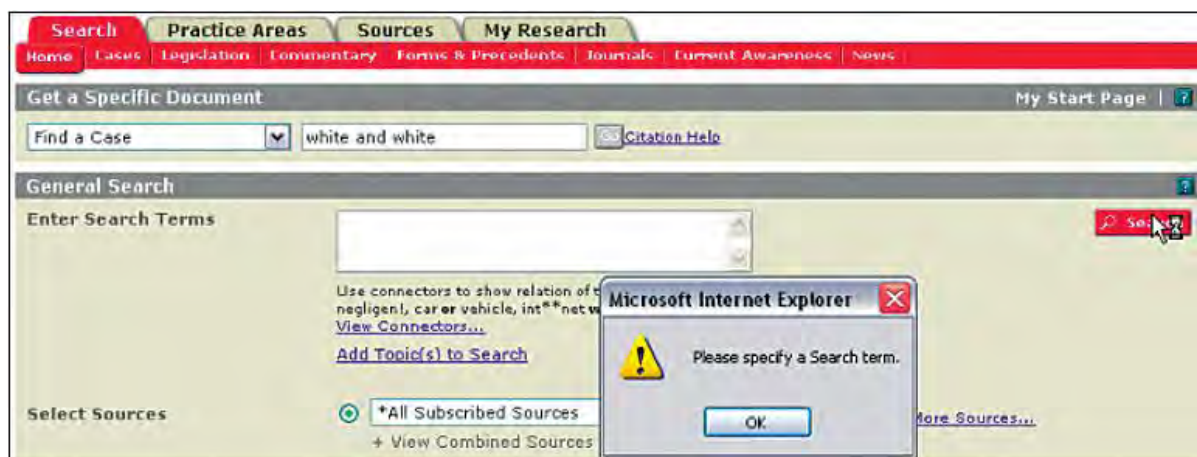FIG. 7. The 'Next Steps' drop-down combo box in LexisNexis Butterworths.

FIG. 8. Popup error message displayed by LexisNexis Butterworths when a search is submitted without anything in the 'enter search terms' field.

Benefits of Using the IB Functionality and Usability Methods

We believe the key benefits of using the IB functionality and usability methods are related to those aspects of the methods that make them novel. For example, because the methods are based on *empirically observed* IBs, they allow evaluators to take a truly user-centred as opposed to system-centred focus on improving usability and providing an appropriate range of functionality. Similarly, because the methods *specialize in facilitating the evaluation of electronic resources* and not other types of interactive system, they allow evaluators to focus on functionality and usability-related issues that are specific to or particularly problematic with these types of systems. For example, issues to do with "accessing" behavior are particularly important for developers of electronic resources to consider as these resources are not always accessed in a straightforward manner like an e-commerce or other Web site might be. Similarly, there are few other types of interactive systems where it might be important to consider providing the functionality to restrict access at a variety of levels (e.g., access to the resource itself, access to particular sources within the resource, access to particular documents). One should note, however, that addressing access issues may be not necessarily be within developers' control (see Bates, 2002; Blandford, Gow, Buchanan, Rimmer, &Warwick, 2007).

Other benefits, which we have already discussed, include the *extensibility* and *flexibility* of the methods. To recap, the IB methods can be tailored in two main ways. First, particular behaviors/levels can be included or excluded from both functionality and usability evaluations depending on the current domain or focus. Second, evaluations can be conducted on entire resources or particular parts of a resource. Modularity, scalability, flexibility, and customizability are all highlighted as important features of methods by Garzotto and Perrone (2007). We have also previously discussed the *highly structured* nature of the methods. This feature of the methods, combined with their extensibility and flexibility, allows users of the methods to strike their own balance between having the necessary structure and guidance to perform a successful evaluation with the flexibility to tailor the evaluations to meet their needs. There is also potential for customization to particular domains to, over time, enrich the behavioral theory base at the heart of the methods. We are unaware of any existing evaluation methods with similar enrichment potential. There are further benefits associated with both the separate methods. Because an IB functionality evaluation allows users of the method to consider ways that particular behaviors *might be supported* at certain levels, the method supports the making of novel as well as incremental design improvements, within certain limits. In addition, because an IB

functionality evaluation allows users of the method to consider the arguments for and against supporting a particular behavior or set of behaviors at each relevant level, it does not propagate the (often incorrect) assumption that supporting a greater range of user behaviors will result in a better, more usable system. Because the behavior-focused tasks used in the IB usability method were generated as a result of empirical data (and by surveying how a range of electronic legal resources support IBs at various applicable levels), the tasks (and therefore the usability method in general) are easily customizable and updatable. In addition, new ways of supporting particular behaviors/levels that are discussed as part of functionality evaluations can feed in to the tasks provided in usability evaluations. Future empirical work or surveys of electronic resources, perhaps in another domain also has the potential to update and customize the methods.

Limitations and Scope of the IB Functionality and Usability Methods
Although there are a number of benefits associated with both the IB functionality and the usability methods, there are also a number of potential limitations associated with them. Many of these limitations are related to the scope of the methods.

Although some solutions to usability observations may become apparent simply by noting that an issue exists, the scope of the IB usability method is restricted to *identifying* as opposed to addressing usability issues, and the scope of the IB functionality method is restricted to examining the *range of support* provided for particular IBs in an electronic resource (as opposed to adding or removing functionality). Although there is an ongoing debate with regard to what extent evaluation should directly inform design, discussed in Wixon (2003), we argue that the IB methodswould become too complex and difficult to use if they were intended to help users of the method decide exactly how to address usability observations or make *binding* decisions on whether and if so in which ways to support a particular behavior at the interface level. Although we do not believe that evaluation and design efforts should be separated completely, our IB methods are pitched primarily as evaluation rather than as design tools. That is not to say that both methods cannot provide useful feed-in to future design discussions. For example, the usability issues identified and the broad ways in which an electronic resource *might* support particular behaviors/levels might be used as a basis for future design discussions. Indeed the outputs of both IB functionality and usability evaluations can be used informally to support design discussions, or more formally in conjunction with design tools such QOC (see MacLean, Young, Bellotti, & Moran, 1991) to help stakeholders make interface-level design decisions.

A sizable challenge for evaluation methods is to support the making of *novel* as well as *incremental* design improvements. The IB usability method indirectly supports users of the method in making incremental design improvements by facilitating the identification of usability issues associated with the current system (which, if addressed effectively, should lead to an incremental improvement in the usability of the electronic resource concerned). The IB functionality method also indirectly supports users in making novel design improvements by allowing them to consider *additional ways* in which the resource might support IBs at particular levels. However, it only provides support for revolutionary design improvements within those strict limits. It does not directly take into account the fact that new ways of performing information tasks might be supported by future electronic resources that change the way that users perform behaviors with these resources (or even change the fundamental behaviors that they perform). Instead, it relies on the assumption that the empirical basis for using the behaviors as part of the method will remain valid (or at least that subsequent research will help to maintain the validity of the method). Therefore, despite evidence from our study of lawyers' IB and from Meho and Tibbo's (2003) study of social scientists that similar behaviors are performed nowadays to those originally identified almost 20 years ago by David Ellis and his colleagues, we cannot be absolutely certain that IBs observed in a particular domain will remain the same over time. In short, the IB functionality method supports novel as well as incremental design but only to a limited extent. The heavy research investment required to ensure the empirical basis of the methods remains valid is also an important limitation of both the IB functionality and usability methods. Finally, the IB functionality method allows users of the method to consider new ways in which user behavior might be supported. However, just as introducing new ways of supporting a certain behavior at a particular level has the potential to lead to improved usability overall (e.g., by providing support for a behavior that has so far been neglected by the resource), it also has the potential to have a *detrimental* impact on the overall usability of the resource (e.g., by helping to create a feature-overloaded resource that remains too complicated to use). This is why the functionality method does not assume that support for a greater range of behaviors/levels will necessarily lead to a more usable electronic resource—an

argument also made by Mack and Nielsen (1994) who point out that several chapters in their handbook of usability evaluation methods "refer to the need for evaluations to focus on the usefulness of interface function, and not simply the usability of the interface, as an implementation of that function" (p. 6). The IB usability method allows users of the method to make usability observations associated with particular behaviors/ levels. However, this also has the potential to have a detrimental impact on the overall usability of the resource. For example, a design intervention aimed at addressing some of the usability observations identified might unknowingly introduce new usability observations, or the design intervention may seem to be an improvement on paper but not improve the usability of the resource in practice. This is why both the functionality and usability methods can be used at various points and at multiple times during the design process (and we encourage iterative use of the methods to complement an iterative design process).

Summary of the IB Methods
The IB methods allow stakeholders in an electronic resource to evaluate its functionality and usability by using a number of empirically identified information behaviors/ levels at which these behaviors can be performed as springboards. These behaviors are used to consider whether and in which ways a particular resource currently or might support each behavior, at each applicable level, aimed at ensuring that the resource supports an appropriate range of functionality. It is also possible to use these behaviors as "lenses" on the usability of the resource by having intended or actual users step through the resource and attempting to perform tasks related to each behavior (and potentially at particular levels). This is with the aim of highlighting usability issues related to each behavior that if addressed effectively, can lead to an improvement in the usability of the resource.

Our evaluation efforts suggest that the IB methods have the potential to yield rich functionality and usability-related data. We believe the most important benefits of the methods are associated with the aspects that make them novel. One aspect that makes the IB methods novel is the specialized nature of the methods, which were developed to facilitate the evaluation of electronic resources as opposed to other types of interactive system. Another aspect is the extensibility and flexibility provided by the methods, which allows evaluators to tailor various aspects of the methods to meet their particular needs while simultaneously providing enough structure to help ensure that rich data is obtained from the resultant evaluations.

The evaluation of the IB methods also suggests, more broadly, that empirically grounded HCI evaluation methods have the potential to be useful, usable, and used in practice. However, a sizable and long-term research challenge still exists in ensuring methods such as ours successfully transfer from theory to practice. We believe our work has taken a number of small but important steps towards addressing this challenge.

Acknowledgments

References

Bates, M. (1989). The design of browsing and berrypicking techniques for the online search interface. Online Review, 13(5), 407–24. Bates, M. (2002). The cascade of interactions in the digital library interface. Information Processing and Management, 38, 381–400.

Belkin, N. (1980). Anomalous states of knowledge for information retrieval. Canadian Journal of Information Science, 5, 133–143.

Bell, B. (1992). Using programming walkthroughs to design a visual language. Unpublished doctoral dissertation, University of Colorado at Boulder.

Blandford, A., Adams, A., Attfield, S., Buchanan, G., Gow, J., Makri, S., Rimmer, J., & Warwick, C. (2008). The PRET A rapporter framework: Evaluating digital libraries from the perspective of information work. Information Processing and Management, 44(1), 4–21.

Blandford, A. & Green, T. (2008). Methodological development. In Cairns, P., & Cox, A. (Eds.), Research methods for human-computer interaction (1st ed.). Cambridge, United Kingdom: Cambridge University Press.

Blandford, A., Green, T., Furniss, D., & Makri, S. (2008). Evaluating system utility and conceptual fit using CASSM. International Journal of Human-Computer Studies, 66(6), 393–409.

Blandford, A., Gow, J., Buchanan, G., Rimmer, J., & Warwick, C. (2007). Creators, composers and consumers: Experiences of designing a digital library. Baranauskas, C. et al. (Eds.). INTERACT 2007, Lecture Notes in Computer Science, Part I, pp. 239–242.

BlandfordA.,Keith, S., Connell, I.,&Edwards, H. (2004). Analytical usability evaluation for digital libraries: A case study. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, pp. 27–36. June 7–11 2004. Tucson, Arizona, USA.

Blandford, A., Keith, S., & Fields, B. (2006). Claims analysis "in the wild:" A case study on digital library development. International Journal of Human-computer Interaction, 21(2), 197–218.

Blandford, A., Keith, S., Fields, B., & Furniss, D. (2007). Disrupting digital library development with scenario informed design. Interacting with Computers, 19(1), 70–82.

Boren, M.,&Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. IEEE Transactions on Professional Communication, 43(3), 261–278. Colbert, M., Peltason, C., Fricke, R., & Sanderson, M. (1997). The application of process models of information-seeking during conceptual design. In Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (pp. 73–81), August 18–20, Amsterdam, the Netherlands.

Dervin, B. (1983). An overview of sense-making research: Concepts, methods, and results to date. Paper presented at the meeting of the International Communication Association, May 1999, Dallas, TX, USA. Dumas, J., & Redish, J. (1999).A practical guide to usability testing. Exeter, United Kingdom: Intellect Books.

Ellis, D. (1987). The derivation of a behavioral model for information retrieval system design. Unpublished doctoral dissertation, University of Sheffield, United Kingdom.

Ellis, D. (1989). A behavioral approach to information retrieval system design. Journal of Documentation, 45(3), 171–212. Ellis, D. (1993). Modeling the information-seeking patterns of academic researchers: A grounded theory approach. Library Quarterly, 63(4), 469–486.

Ellis, D., Cox, D., & Hall, K. (1993). A comparison of the informationseeking patterns of researchers in the physical and social sciences. Journal of Documentation, 49(4), 356–369.

Ellis, D., & Haugan, M. (1997). Modelling the information-seeking patterns of engineers and research scientists in an industrial environment. Journal of Documentation, 53(4), 384–403.

Garzotto, F., & Perrone, V. (2007). Industrial acceptability of web design methods: An empirical study. Journal ofWeb Engineering, 6(1), 73–96.

Hertzum, M., & Jacobsen, N. (2001). The evaluator effect: A chilling fact about usability evaluation methods. International Journal of Human-Computer Interaction, 13(4), 421–443.

Kuhlthau, C. (1988). Developing a model of the library search process: Cognitive and affective aspects. Reference Quarterly, 28(2), 232–242.

Landauer, T. (1995). The trouble with computers: Usefulness, usability and productivity. London, United Kingdom: MIT Press.

Mack, R.,&Nielsen, J. (1994). Executive summary. In J. Nielson & R. Mack (Eds.), Usability inspection methods. NewYork:Wiley.

Makri, S., Blandford, A., & Cox, A.L. (2008). Investigating the information seeking behavior of academic lawyers: From Ellis's model to design. Information Processing and Management, 44(2), 613–634.

Marchionini, G. (1995). Information-seeking in electronic environments. Cambridge, United Kingdom: Cambridge University Press.

MacLean, A., Young, R., Bellotti,V.,& Moran,T. (1991). Questions, options, and criteria: Elements of design space analysis. Human-Computer Interaction, 6(3 & 4), 201–250.

Meho, L., & Tibbo, H. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. Journal of theAmerican Society for Information Science and Technology, 54(6), 570–587.

Page numbers differ from published version

Nielsen, J. (1993). Usability engineering. NewYork: Academic Press.

Norman, D. (1986). Cognitive engineering. In D. Norman& S. Draper (Eds), User centred system design, pp. 31–61. Mahwah, New Jersey: Lawrence Earlbaum Publishers.

Pirolli, P., & Card, S. (1999). Information foraging. Psychological Review, 106(4), 643–675.

Polson, P., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies, 36(5), 741–773.

Smith, K. (1988). An investigation into the information behavior of academics active in the field of English literature. Unpublished master's thesis, University of Sheffield, United Kingdom.

Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. Interacting with Computers, 10, 321–351.

Sutton, S. (1994). The role of attorney mental models of law in case relevance determinations: An exploratory analysis. Journal of the American Society for Information Science, 45(3), 186–200.

Wharton, C., Reiman, J., Lewis, C.,& Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In J. Nielsen & R. Mack (Eds.), Usability evaluation methods. NewYork:Wiley.

Wilson, T. (1999). Models of information behavior research. Journal of Documentation, 55(3), 249–270.

Wixon, D. (2003). Evaluating usability methods: Why the current literature fails the practitioner. Interactions, 10(4), 28–34.

Page numbers differ from published version

Appendix A

Form to be Filled out When Conducting an IB Functionalty Evaluation

Steps 1 and 2 should be carried out (and a separate form filled out) for *each behavior and level* that you have chosen to evaluate the electronic resource in relation to. Step 3 should only be carried out once, at the end of the functionality evaluation.

Step 1: Tick *one* box to indicate the behavior and level you will be answering the questions below about:

	Resource Level	Source Level	Document Level	Content Level	Search query/ result level
Accessing					
Surveying					
Monitoring					
Searching					
Browsing and extracting					
Chaining					
Selecting, distinguishing and filtering					
Updating					
History tracking					
Analysing					
Recording					
Collating					
Editing					
Distributing					

Step 2: Answer the relevant questions below about the behavior and level that you have ticked:

If the behavior you ticked is currently supported at this level by the resource, answer the following questions:

(1) *In which way(s)* does the resource currently support this behavior at this level?

```
[                                                                            ]
```

(2) In which *additional* way(s) might the resource support the behavior at this level? You do not need to answer this question if you *do not have a direct stake in the resource* (e.g., you have no input into the design, improvement or marketing of the resource or it is a competitor resource)

```
[                                                                            ]
```

If the behavior you have ticked *is not* currently supported at this level by the resource, in which way(s) might the resource support the behavior at this level?

```
[                                                                            ]
```

Additional notes from discussion of issues surrounding supporting the behavior(s) at this level:

```
[                                                                            ]
```

Step 3: Answer the following two questions: Only complete this step once, at the end of the evaluation.

(a) Are there any *behaviors/levels* that may *no longer be necessary* to support? For any behaviors/levels for which you are considering ceasing support, what are the potential arguments for and against support?

```
[                                                                            ]
```

(b) Are there any ways that you currently support any of the behaviors/levels that may *no longer be necessary*? For ways of supporting a particular behavior/level for which you are considering ceasing support, what are the potential arguments for and against support?

```
[                                                                            ]
```

Page numbers differ from published version

Appendix B

Excerpts from Form Used to Record the Output of an IB Usability Evaluation

The form below is filled out with usability data related to the think-aloud transcript presented in Appendix 3. The letters next to the "user actions, comments or own observations that might suggest a usability issue" allow you to cross-reference the data on this form with highlighted parts of the transcript. This serves to provide context and further explanation of the summary data on this form.

Fill in these columns whilst watching the think-aloud session.		Fill in these columns either whilst watching or after watching the think-aloud session.			
User actions comments, or own observations that might suggest a usability issue	Approx. time in video actions/ comments/ observations occurred	Usability issue(s) identified from actions/comments/ observations.	Screen(s)/ page(s) / part(s) of resource that actions/ comments/ observations relate to	Severity of issue (Not severe - does not need immediate attention, Quite severe – needs attention in the future, Very severe – needs immediate attention)	Amount of effort required to address issue (Small/ Medium/ Large)
a, g) User about to type/types search query into incorrect segmented field	1 min 0s, 11min 10s	Not immediately clear which segmented field to enter data in	'Cases' search page	N	L
b) User initially enters incorrect search syntax	1 min 10s	Required syntax not immediately clear	'Cases' search tab and all other search pages	N	S
c) User clicks on 'next steps' combo box hoping to check whether case has been updated	4 min 50s	Actions facilitated by 'next steps' combo not transparent or clear to user	All full-text document pages	Q	M
d) Help pages 'seem to be frustrating' user	9 min 0s	Help pages did not provide the required help	Total Help pages	Q	L
e) 'Slow and cumbersome' tutorial	10 min 20s	Tutorial pages did not help	Tutorial pages	Q	M
f) User unsure which 'submit' button to use	10 min 40s	Relation of submit buttons to search fields potentially unclear	'Home' search page	V	M
h) User unaware that search was processing even though 'go' button turned grey upon submit	10 min 40s	Unclear feedback provided relating to search progress	'Home' search page	Q	M
i) User unsure of what happened when selecting case name from 'view' combo box	12 min 50s	Effect of selecting current case from 'view' combo potentially unclear	All full-text document pages	N	M

Fill in this column after watching the session.
Reflections on usability issues identified (e.g. broad themes arising from issues identified, issues for future discussion etc.)
As a novice user, the trainee solicitor was able to perform basic searches, but it was not always clear how to achieve particular goals using the resource, or what effect performing certain interface actions might have. The help and tutorial screens frustrated the user rather than aiding the updating task.

Appendix C

Think-Aloud Transcript of a Trainee Solicitor Using Lexisnexis Butterworths to "Find out Whether a Particular Case is Still Good Law"

The transcript below is a verbatim account of the comments and actions made by a trainee solicitor who was asked to use the LexisNexis Butterworths electronic legal resource to "find whether a particular case is still good law." The trainee solicitor was asked to think aloud while performing the task, verbalizing his thoughts, actions, and feelings.

The trainee's actions are presented in square brackets [ ]. The highlighted parts of the transcript denote sections that are referred to, by the letter alongside them, on the form presented in Appendix 2.

There's one case called White and White, which is used in ancillary relief cases and I want to find out about that. I want to find out if that case is still valid as good law, you know, or whether any amendments have been made to the case over a period of time. So what
a I'm gonna do is I'm back at the LexisNexis home screen. I'm gonna click on "cases" and under "enter search terms" I'm going to put [pauses]. No, I'm going to put it in the "case name." So White and [Participant types in "White v W," then pauses]. [Reads caption
b underneath "case name" field]. "To find Smith v Jones, enter Smith and [emphasis placed] Jones." So I'm going to use the word "and" rather than "v." I don't have a citation unfortunately and I'm going to click "search" and see what that brings back.

1700 cases, which may make it a bit difficult to find. [Clicks on "cases" results group]. So I've clicked on 'cases' to expand it [scrolls through list of cases]. It's going to be quite difficult to find. I'm just going to look on the left-hand side at where the case has been reported [scrolls through side-bar detailing result groups that comprise mostly report series and levels of court]. This is a family case, so I'm just going to have a quick browse. If not, I'll have to go back and narrow the search down slightly. Aah, "Family Court Reports" [reads report series aloud and clicks on hyperlink]. I'm hoping that that may [emphasis placed] bring up the case. Right, White and White 2000 and it's a House of Lords case. I think that may be it. So I'll click on that. Right, [reads case name aloud] White and White. [Begins to read keywords aloud]. "divorce," "financial provision, ancillary relief," "available assets exceeding parties' financial needs for housing and income." Aah, there we go, this is the case. "Principles to be applied." [Reads first paragraph of full-text of case]. "The parties' marriage broke down after 30 years." "The overall net worth of their assets was approximately £4.6 million." It's a big money case. Ok, this is the case. [Scrolls through text]. Now, I just want to see if it's been updated in any way.

Hmm. I've just scrolled to the bottom. That's the law lords' dictum. [Scrolls back up through text]. Ok. I'm looking at the cases referred to in the opinions. I can't see anything there. I'm just wondering how I can check to see if the case has been updated or not. [Clicks on "view" combo box]. I'm clicking on "view," [reads options aloud]. "Expanded list," "list." These are presumably lists. Next steps
c [clicks on "next steps" combo box and reads some of the options aloud]. "Edit search." Here we go. "Save search," "create alert," "search the source," "find related cases." Hmm. I'll click on "find related cases" and see what that brings up. [Scrolls through results]. And this has brought up 48 further cases. I'm just going to click on the top one, which is dated the 18th June 2007 [clicks on Wood vs. Rost case] which is called Wood and Rost [reads keywords of case aloud]. None of that's got anything to do with the split. [Scrolls through text of case]. Aah. Here we go. Paragraph 7. [Reads out text in paragraph 7 of the case report]. "Mr Mostyn, fortified by the decision of the House of Lords in White v White." And it's got the references. And it's got one of the references in red. I'm going to click on that in a minute. [Continues to read paragraph 7]. Let's click on the red link. Presumably that will take me back to what I was looking at before. Yeah, it has done.

[Scrolls through text of White and White case again]. As to see whether or not it is still good law, I think I'm a bit stuck. So I think what I'm going to have to do is press "help." [Reads title in TotalHelp screen]. I'll click on "home" because that doesn't really help me. [Hovers cursor over headings, then clicks on "tutorials" hyperlink]. Tutorials. Well, I know how to search. Home. [Clicks on "home" hyperlink]. It's taken me back to where I was. I'll click on "search for information." Nope. [Clicks on "locate sources" heading]. Nope.
d [Clicks on "review search results" heading and reads out some of the sub-headings]. "How do I work with my search results?" I'll click on that as it might help me in some way. [Reads out sub-headings]. No, that doesn't help me. Let's search. [Searches within TotalHelp for "current cases" and clicks on a result entitled "results page" then soon after closes the help page]. I've closed that down because it just seems to be frustrating me.

[Participant returns to White and White case and scrolls through text again]. What I am [emphasis placed] going to do is go back to the Lexis homepage as there are often tutorials on the right-hand-side. [Reads headings in 'related links' side box]. "View tutorials," "overview," "selecting sources," "working with results." [Reads other headings aloud, then scrolls down and back up the LexisNexis Butterworths homepage]. No. [Clicks on "overview" under 'view tutorials' heading]. I'll click on that. Maybe it will bring up something different. It has linked to a tutorial. [Click on "working with results" tutorial option]. Working with results. I'm waiting for it to start. [Clicks
e mouse two or three times]. Ok, the tutorial seems very slow and cumbersome, so [closes tutorial window], I'm going to try searching again. [Enters "white and white" in the "get a specific document" part of the LexisNexis Butterworths homepage with "find a case"
f/g selected in the combo box. Participant then presses the grey "go" button next to the field he has filled out, but after a few seconds presses the main red "search" button. [An information box is displayed, which says "Please enter a search term"]. [Participant then clicks on the "cases" tab on the LexisNexis Butterworths homepage].

h [Participant types "White" in "citation" field]. I'm not looking for a citation. [Participant then enters "White and White" in the case name field]. I know it's a House of Lords case [ticks "House of Lords" in the "Court" selection box] so that will narrow everything down a bit [conducts search]. Cases [clicks on "cases" results group]. Family Court Reports [clicks on Family Court Reports sub-group and selects the top case from the list, White and White]. Actually, I'm just going to go back [clicks browser "back" button and ticks box beside the White and White case that participant had just previously clicked on]. [Clicks on "sort" combo box and "view" combo box, then clicks on "view tagged" button. View tagged. I'm assuming that will be the same as if I just clicked on it. Yeah. I don't think I can [pauses]. [Clicks on "next steps" combo box and reads some of the options aloud again]. "Find related cases," "find
i related commentary" [reads other options]. View [clicks on "view" combo box again and reads out options]. "List," "expanded list," "White v White 2000." [Selects White v White 2000]. Oops, I clicked on something. [Scrolls through text of White and White 2000 case]. Unfortunately I think I'm stuck and can't go any further on this one.