A Comparison of the Effectiveness of a Dichotomous Key and a Multi-Access Key to Woodlice

D.R. Morse & G.M. Tardivel

Computing Laboratory
University of Kent
Canterbury
Kent CT2 7NF

Email:
D.R.Morse@ukc.ac.uk
G.M.Tardivel@ukc.ac.uk

J. Spicer

Department of Animal and Plant Sciences
University of Sheffield
PO Box 601
Sheffield S10 2UQ

Email:
J.I.Spicer@sheffield.ac.uk

Introduction

Species identification, particularly by non-experts is a difficult and often frustrating task. As well as the dichotomous key familiar to most ecologists, taxonomists have developed a range of tools and techniques aimed at easing the task of identification. Increasingly, these tools have used computers either in the production of the key or in the key itself (see Edwards & Morse (1995) and Pankhurst (1991) for reviews).

While the theory and implementation of computer-based species identification tools is well developed in a number of areas, much less is known about how well these tools perform in practice (Edwards & Morse 1995). Exceptions to this include Stucky (1984) and Wright *et al.* (1995) who described evaluation experiments where volunteers were asked to identify species using two identification tools. Both studies compared a novel identification tool (a hypertext key in the study of Wright *et al.* and a polyclave in Stucky's) to a dichotomous key. They both found that the novel identification tool was preferred over the dichotomous key and that there was no difference in the frequency of correct identifications. Most worrying is that both studies reported a high frequency of misidentifications, reaching 30% in Stucky's (1984) study. While this figure lumps all forms of misidentification, it is high and suggests that there is much work to be done in making the identification process easier and more accurate.

One of the most popular forms of computer-based identification tool is the multi-access key (Legg 1992a,b and Pankhurst 1991). It uses a data matrix of species × character combinations, where rows of the matrix represent species and the columns represent characters (see Edwards & Morse (1995) for further details and the availability of multi-access key packages). The most significant feature of the multi-access key is that the user enters the character states in the order they wish rather than the key requesting information about specific character states, as happens in a dichotomous key. This means that multi-access keys potentially have at least two advantages over a dichotomous key:

1. The most obvious characters will tend to be selected first and the user is more likely to be able to identify such characters correctly.

2. The user is not required to 'guess' the states of uncertain characters.

The chief disadvantage of allowing the user to control the identification is that, while the user may select characters that are easily observed, these may not make a significant contribution to the identification by discriminating between the remaining species.

The Project

The aim of the project was to conduct an empirical comparison of the efficiency, accuracy and usability of a multi-access key and a dichotomous key to the same animal or plant group. To act as a control for the multi-access key which was both computer-based and had a different underlying theory to the paper dichotomous key, a hypertext version of the dichotomous key was also used in the evaluation experiments. Previous work had shown that volunteers were no less effective when using the hypertext version of the key than they were when using the paper version of the key, although they were slower in achieving an identification (Wright *et al.* 1995).

The Keys

Three keys were used in the evaluation experiment. They were all based on Hopkin's (1991) *A Key to the Woodlice of Britain and Ireland* which is one of the Field Studies Council's AIDGAP keys. There are several reasons why an AIDGAP key and Hopkin's key in particular were chosen for the evaluation experiment:

- Hopkin's key was recently developed and the text of the key was available in machine-readable form.

- Woodlice are one of the more 'user-friendly' groups of invertebrates. They are fairly hardy, can be cultured in the laboratory and several species are very abundant. Indeed, the 'big-five' woodlice species are known as such because of their abundance and wide distribution (Hopkin 1991).

- The AIDGAP (Aids for the Identification of Difficult Groups of Animals and Plants) project aims to identify those groups for which the difficulty in identification is due to the absence of a simple and accurate key rather than being due to insuperable taxonomic problems and to produce simple, well-written aids to identification (Tilling 1987).

Three versions of this key were developed:

1. The original key (Hopkin 1991) published as an AIDGAP guide.

2. A hypertext version of the key, developed using the PC Guide hypertext system. This contained the same text, diagrams and photographs as Hopkin's key. The main difference between this version and the first was in the manner of presentation: hypertext rather than paper.

3. A multi-access key developed using Legg's (1992a) key. This package was chosen rather than some of the other commercially available multi-access keys (Edwards & Morse 1995) because it is one of the cheaper packages available, it has modest hardware requirements and it is an open system. That is, users of the package are encouraged to develop their own keys and programs are provided to help them do this.

The woodlice multi-access key differed from most other multi-access keys in one important respect, namely that the character × species matrix had many more unknown elements than is usual or desirable in a multi-access key. These unknown elements arose because of the way in which the matrix was developed. It was developed from the information contained in Hopkin's key which was not augmented in any way. Normally a multi-access key would be developed from a number of sources including original material in order to fill in as many entries in the matrix as possible.

Had the matrix been augmented from other sources then the three keys would not have contained comparable amounts of taxonomic information. It is true that the multi-access key is sub-optimal because of all the unknowns in the matrix. However, if all these gaps were filled in from other sources, it could no longer be compared to the other keys because it would then differ in at least two respects: the different user interface and philosophy of use, and secondly, the amount of taxonomic information it contained.

Methods

It was originally intended to carry out evaluation experiments at a number of institutions because we had, in the past, experienced difficulty in obtaining large numbers of volunteers from any one institution. However, it proved possible to integrate an evaluation experiment into the University of Sheffield's Zoology course. This experiment formed one of four practicals being taken by second year students following the Animal Diversity course. The practicals were run on a 'round-robin' basis with students being assigned to one of eight groups. One group participated in the evaluation experiment in each practical class.

Each student performed the same task with one of the three keys, to which they were randomly assigned. They were first asked to identify a specimen of *Ligia oceanica* to familiarise themselves with the key they were using. This was intended to mitigate against the unfamiliarity of the two computer-based keys, which was one reason why the hypertext key had been found to be slower than the paper key in earlier trials (Wright *et al.* 1995). Then, the students were asked to identify an unknown live woodlouse and to note how long the identification took them. They were also asked to fill in a short questionnaire detailing their prior experience of identifying woodlice and how they had found various aspects of using the key. We also asked them to indicate how confident they were in their identification.

Three species of woodlice were chosen: *Armadillidium vulgare*, *Philoscia muscorum* and *Porcellio scaber*. The former two species 'keyed out' at approximately the same level in the identification tree for the hypertext and paper keys (Figure 1), while the latter species keyed out at a much lower level, requiring many more choices to be made before it could be identified correctly.

As three species of woodlice were chosen, and there were three keys, each trial involved nine students. We were extremely fortunate in that seventy-two students took the Animal Diversity course. There were eight practical classes giving us nine students per session, allowing us to run one complete trial per practical class.

Results

All 72 students achieved an identification and only one student had had prior experience of identifying woodlice. Table 1 shows that overall, 53 students achieved a correct identification, corresponding to an accuracy of identification of 53/72 = 74%. Table 1 also shows that two species: *A. vulgare* and *P. muscorum* were identified much more readily than *P. scaber*. This could be due to the distinctiveness of the two former species compared to the latter, or to the fact that with the paper and hypertext keys many more choices had to be made before a correct identification was achieved for the latter species compared to the former (Table 1 and Figure 1).

Turning to the different identification media, Table 2 shows that the two computer-based methods were slower than the paper key. The multi-access key was slightly more accurate than the paper key and the hypertext key was the least accurate. Overall, the accuracy of identification across all three keys was 66%. Most intriguing about the data in Table 2 was how long it took to achieve a correct or incorrect identification. For each of the three media, it took longer to achieve a correct identification than it did an incorrect identification, by at least 3 minutes 18 seconds.

Species	Number of couplets	Identification		Total
		Correct	Incorrect	
Armadillidium vulgare	7	19	5	24
Philoscia muscorum	6	20	4	24
Porcellio scaber	11	14	10	24
Total		53	19	72

Table 1: Frequency of correct identifications broken down by species but pooling across the methods used to identify the specimens. The number of couplets column shows how many choices have to be made before a correct identification is achieved in the paper and hypertext keys (see Figure 1).

Method	Average time taken			Identification		Total
	Correct	Incorrect	Difference	Correct	Incorrect	
Hypertext	18:01	14:43	3:18	16	8	24
Multi-access	22:28	18:21	4:07	19	5	24
Paper	15:31	10:33	4:58	18	6	24
Total	18:46	14:22	4:24	53	19	72

Table 2: Frequency of correct identifications broken down by the method used to identify the specimen. Also shown are the average times to achieve an identification and the difference between the time taken to achieve a correct and an incorrect identification. Times are given in minutes:seconds.

Students were asked to enter their confidence in their identification in one of five categories: ‘Definitely correct’, ‘Probably correct’, ‘Don’t know’, ‘Probably incorrect’ and ‘Definitely incorrect’. One student didn’t fill in this part of the questionnaire, hence the totals in Tables 3 and 4 are 71 and not 72. None of the students thought their identification was either ‘Probably’ or ‘Definitely’ incorrect so these categories were not used.

Looking at the students’ confidence in their identification by species (Table 3) it can be seen that the students are more confident in their identification of the two species where they obtained a higher frequency of correct identifications (*A. vulgare* and *P. muscorum*). They were less confident in identifying *P. scaber* (3 Definitely and 15 Probably correct, compared to 9 Definitely and 13 Probably correct with the other two species) which had many more incorrect identifications.

How confident were the students in their identification with respect to the different media? They had equal confidence in the identifications obtained using the paper and multi-access keys, although with the paper key there were more ‘Definitely correct’ identifications which were in fact wrong (Table 4). There was a wider spread of confidence in the hypertext key, although it appeared that students were in general more confident of their identification with the hypertext key than the other two keys (9 Definitely and 10 Probably correct, compared to 6 Definitely and 15 or 16 Probably correct). However, it was noted above that the hypertext key was the least accurate of the three keys, although it was only a little worse than the other two keys.

Finally, Table 5 summarises the students’ impressions of the keys. Many of their comments concerned usability and navigation issues. In particular, students commented on the frequently encountered problem of finding their way round the paper key. Easing this task was one of the original motivations for developing hypertext keys (Wright *et al.* 1995). In general, students found both computer-based keys easy to use but the diagrams and colour plates in both keys were criticised, either because they weren’t there (the multi-access key) or because of the poor quality of some of the images and text (the hypertext key).

Species	Identification	Confidence in identification			Total
		Definitely correct	Probably correct	Don’t know	
A. vulgare	Right	7	11	1	19
	Wrong	2	2	1	5
	Total	9	13	2	24
P. muscorum	Right	8	12	0	20
	Wrong	1	1	2	4
	Total	9	13	2	24
P. scaber	Right	3	10	1	14
	Wrong	0	5	4	9
	Total	3	15	5	23
Total		21	41	9	71

Table 3: Student assessments of how confident they are in their identification scoring each student’s assessment according to whether their identification was by species, pooling across the methods used to identify the specimens, and right or wrong. The two categories ‘Probably incorrect’ and ‘Definitely incorrect’ were not used by the students. One student did not fill in this part of their questionnaire hence the total of 71 responses.

Discussion

An overall frequency of correct identifications of 74% is high when it is considered that only one student had had any prior experience of identifying woodlice. This compares favourably with Stucky (1984) who found a misidentification frequency as high as 30% (corresponding to 70% correct identifications) in a dichotomous key and polyclave to weed seedlings. It is also an improvement on an earlier study (Tardivel and Morse unpublished) where Year 12 and 13 pupils were asked to identify woodlice using the paper and hypertext keys. In that experiment the frequency of correct identifications fell as low as 60%, although those subjects had little experience of using keys and virtually no experience of identifying woodlice. How many errors experts (either in the taxonomic group or in the use of the key) would make is not known.

Media	Identification	Confidence in identification			Total
		Definitely correct	Probably correct	Don't know	
Hypertext	Right	8	8	0	16
	Wrong	1	2	4	7
	Total	9	10	4	23
Multi-access	Right	6	12	1	19
	Wrong	0	3	2	5
	Total	6	15	3	24
Paper	Right	4	13	1	18
	Wrong	2	3	1	6
	Total	6	16	2	24
Total		21	41	9	71

Table 4: Student assessments of how confident they are in their identification according to the method used to identify the specimen. See the legend to Table 3 for further notes on the data.

This study confirms the findings of a previous experiment (Wright *et al.* 1995) that the hypertext key was both slower and less accurate than the paper key. The difference in times between the two keys was less in this experiment because of the familiarization phase which each student undertook before commencing the identification proper. The higher misidentification frequency could be due to students becoming 'mouse-button happy' and selecting one or other couplet even when they are not sure which one is correct. Edwards & Morse (1995) proposed that in such a situation users of the paper key would be more likely to backtrack or start again than in the hypertext key. The observation that students who obtained the correct identification took about four minutes longer than those who did not, regardless of the media they used to identify the specimen, is intriguing. It could be that they were more careful during the identification and hence they were slower than their counterparts who did not achieve a correct identification. Alternatively, they could have spent the four minutes checking the species description and photographs. Or both. Only close observation of people when they are identifying specimens and experiments like this will confirm the effect if it exists and reveal the difference between the two groups of people.

	Likes	Dislikes	Improvements
Hypertext	Navigation. Availability of information. Fun, easy.	Scanned diagrams. Labelling on diagrams. Only one diagram available at a time.	Improve diagrams. Increase size of glossary.
Multi-access	Easy to use. Easy navigation. Ability to skip characters. Next best character. Probabilities.	Swapping between screen and book. Abbreviations of character descriptions. Difficult to remove characters.	More built-in help. Diagrams on screen.
Paper	Availability of information. Ability to see overall structure of the key.	Moving between different parts of the key (e.g. glossary, colour plates etc.). Numbers to follow paths through key.	Improve movement between different parts of the key.

Table 5: A summary of the features of the keys which students liked, those they disliked and their suggestions for how they would improve the key they used.

This is probably the first experiment in which volunteers were asked to estimate the confidence which they place in their identification. In retrospect, the scale on which volunteers were asked to judge their confidence was too crude. This group of students appeared unlikely to admit that they thought their identification was Probably incorrect. It is more likely that they would try again until they achieved an identification in which they had some confidence. While more data is needed to confirm this conclusion, the students were reasonable judges of whether their identification was correct or not. However, only nine people admitted that they didn't know, of which seven were incorrect and two had the correct identification (Table 3). An overall nineteen students were incorrect in their identification (Table 1).

Another of looking at the experiment is that it was a comparative evaluation of three different user interfaces to the same taxonomic information. In general, the quantitative differences between the three keys are small, although in some cases they may well be important differences, such as the accuracy of the identification and the confidence which the students had in their identification. On the other hand, the students' subjective impressions of the keys (summarized in Table 5), could be more important in determining the future development of taxonomic keys. For example, it is encouraging that students found both computer-based keys easy to use and in the case of the hypertext key, fun! There is clearly room for improvement in all three keys as the list of suggested improvements in Table 5 shows. In the paper key, the students dislikes and their suggested improvements stem from the linear nature of paper documents (Nielsen 1986). Hundreds of years of dichotomous key development (Pankhurst 1991) has not yet overcome these restrictions. In contrast, it was the quality rather than the accessibility of the ancillary information which was commented on in the two computer-based keys.

How much further these keys, and other keys like them can be improved is not clear. However, it is likely that computer-based keys can probably be developed further than can paper keys, partly because the former are a much newer technology than the latter and also that technologies such as hypermedia have considerable potential for improving the way in which information is presented.

Key to the Woodlice of Britain and Ireland. 37 species.

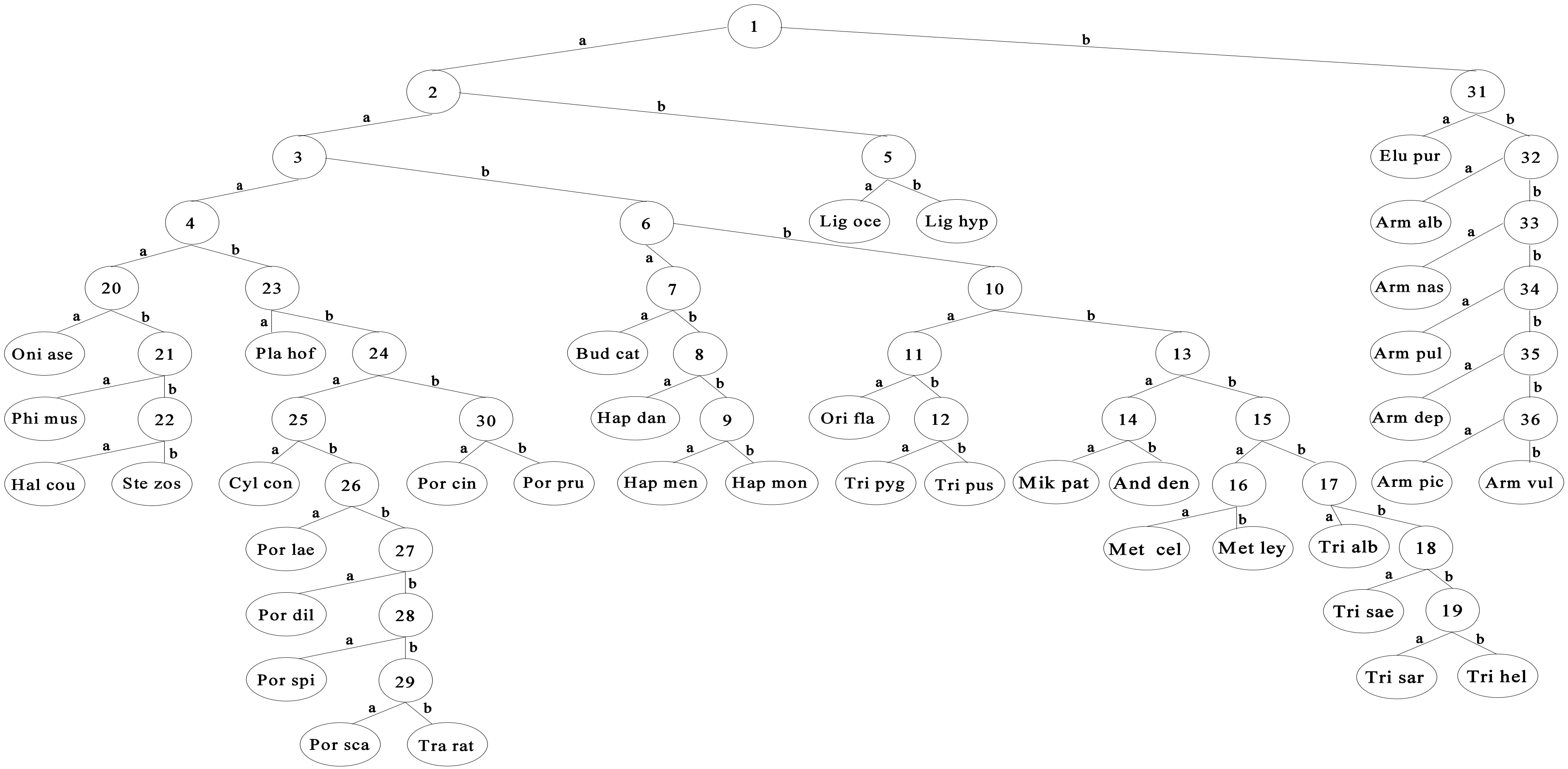

Figure 1. A diagram of the decision tree used in Hopkin's (1991) key. The circles represent the couplets with the couplet numbers as indicated. The two branches of each couplet are labelled 'a' and 'b' on the diagram. Species names have been abbreviated and are shown in the ellipses at the point at which they 'key out' in the decision tree.

Acknowledgements

This work was made possible by a British Ecological Society small ecological projects grant.

Many people have contributed to the success of this evaluation experiment. Steven Hopkin and Steve Tilling gave permission to convert the woodlice key to hypertext form. The British Ecological Society funded the project through a Small Ecological Project Grant. Simon Tardivel and Anne Watson helped to collect woodlice. Phil Warren was a calming influence when things started to go wrong. Last, but by no means least we are very grateful to the 1994/5 cohort of Zoology students at the University of Sheffield who participated in the experiment.

References

Edwards, M., and Morse, D.R. (1995). The potential for computer-aided identification in biodiversity research. *Trends in Ecology and Evolution*, **10**(4): 153-158.

Hopkin, S.P. (1991). A key to the woodlice of Britain and Ireland. *Field Studies*, **7**: 599-650.

Legg, C.J. (1992a). Random-access identification guides for a microcomputer. *Field Studies*, **8**(1): 1-30.

Legg, C.J. (1992b). Random-access guide to sedges of the British Isles using a microcomputer. *Field Studies*, **8**(1): 31-57.

Neilsen, J. (1995).