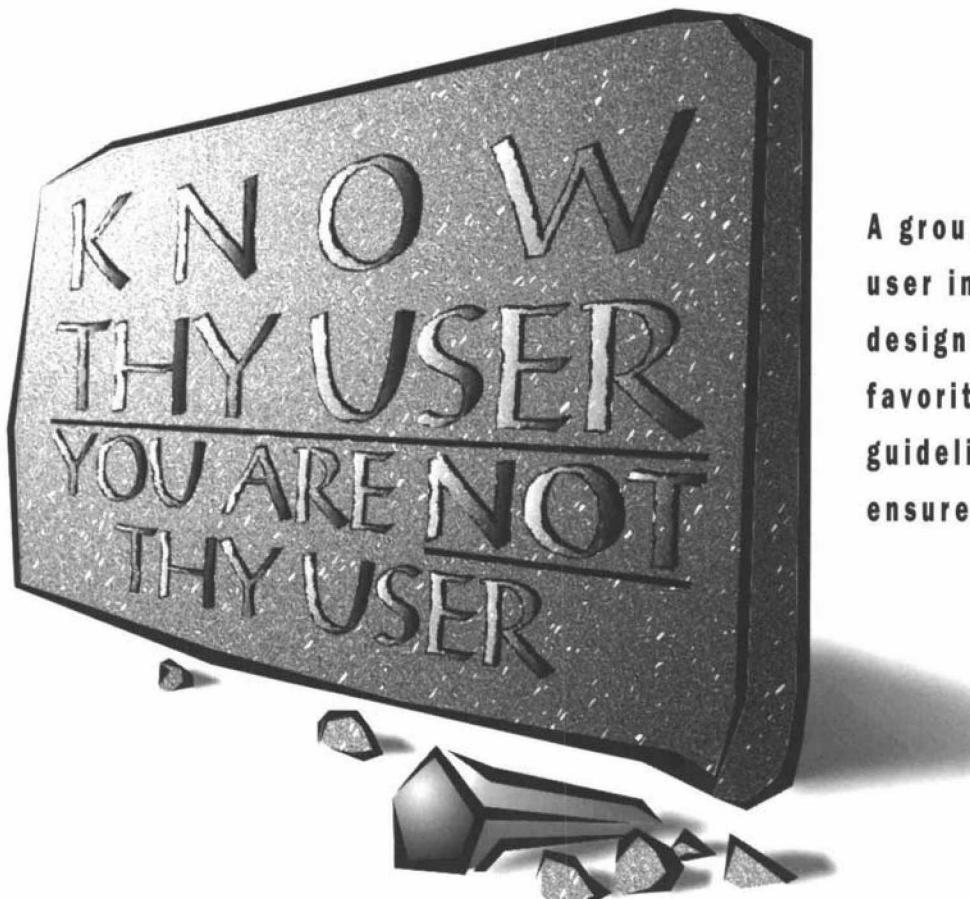


Expert Ratings of Usability Maxims



A group of graphical user interface designers list their favorite brief guidelines to ensure usability.

ILLUSTRATION BY LESLIE V. CARLSON

ONE OF THE PROBLEMS FACING human-computer interface (HCI) designers and those evaluating designs is that there are too many guidelines for interface design. Collections such as those by Brown (1988), Engel and Granda (1975), Smith and Mosier (1986), Boff and Lincoln (1988), and Williges and Williges (1981) are used widely. The International Standards Organization (ISO) is completing work on several sections of its software design standards (ISO 9241), and the Human Factors and Ergonomics Society is sponsoring work to extend and update the standards (ANSI/HFES 200). Jean Vanderdonckt (personal communication, September 25, 1995) participated in a project that resulted in the collection of 3700 guidelines

for the Belgian government and published a bibliography of more than 1200 guidelines references (Vanderdonckt, 1994).

It is unlikely that anyone can keep a significant number of guidelines in mind while reviewing a design, even if there was clear agreement on exactly how to apply them to elements of the design. Perhaps it isn't too surprising, then, that many designers have developed their own lists of general maxims that they keep in working memory. A designer, for example, might make sure that everything in the design that should be consistent *is* consistent. These maxims may represent part of the expertise of the experienced professional (Glaser & Chi, 1988), capturing the categories of design characteristics that the professional has found to be meaningful.

BY
ARNOLD M. LUND



In this article, I present a collection of maxims used by experienced design and usability professionals, and order them by the value that professionals find in the maxims.

Are There "Best" Maxims?

Recent books and articles on heuristic analysis techniques list handfuls of maxims that the authors believe yield significant usability benefits (Nielsen, 1993, 1994), and there is a rich tradition of providing lists of design principles (for example, Cooper, 1995; Hill, 1987; Shneiderman, 1987; Williges, Williges, & Elkerton, 1987). These heuristics and other usability design principles are terse statements of useful principles with wide application. They are not intended to be strictly accurate; and applied without testing, they do not guarantee usability. The maxims can provide a language for discussing usability issues, and they serve as a reminder of important elements of design and techniques that have proved useful for improving design usability.

Little has been written about which maxims experienced practitioners actually use. It might be assumed that these professionals have seen some of the principles represented in the maxims violated more than others by developers and designers. Some of the principles are probably more practical to implement than others. Some are more likely to have a larger impact on usability

Maxims serve as a reminder of important elements of design and techniques that have proved useful for improving design usability.

than others. Over time, a given practitioner might select a set of maxims representing the intersection of these factors that he or she uses implicitly or explicitly from project to project. The experienced professional may have found the small set of maxims that address common but significant usability problems and that he or she is comfortable applying.

Is there a core set of maxims that represent design principles that most practitioners are using, presumably reflecting at least some of the evaluation expertise that is accumulated through experience? If there is a core set, these maxims might be the focus of training, and software tools to support

their application might be developed that would enable their implementation by people entering the field. To answer this question, I conducted a study to identify a list of maxims that seem to reflect good practice and then asked practitioners with expertise in the field to rate their value.

Listing and Rating the Maxims

I asked colleagues in the HCI design field to suggest maxims that they have found particularly useful during the design process. Editing eliminated obvious duplications and refined wording, and then the HCI designers reviewed the edited list to see whether any maxims were missing. I also submitted the list to members of the ANSI/HFES 200 standard committee to solicit their suggestions. After their suggestions were incorporated, the list contained 34 maxims. Some appeared to overlap in terms of the problems they could address, though I believed that they might be differentially useful. Unless they were synonymous, therefore, I retained similar but nonidentical maxims.

I then randomized the list and sent it to a group of professionals with whose work I was familiar (including the ANSI/HFES 200 standard committee). Members of this group are active designers and usability testers, many of whom have specified design guidelines for style guides and standards bodies. I asked them to rate each of the maxims based on their estimate of the impact of the principle represented by the maxim on the usability of designs on which they work. They were asked to think about how, in their experience, violations of each principle affect usability, the frequency with which a given principle tends to be violated in designs, and which principles they try to keep in mind when working on a design. Using a forced-choice technique, I asked them to assign a 5 (for the greatest impact on usability) to about a fifth of the maxims, a 4 (for the next-greatest impact) to about a fifth, and so on down to a 1 (for the least impact relative to the others) for the final fifth. I received responses from 16 people in this group, and they will be referred to as Respondent Set 1.

The list was randomized again and sent to several listservers used by people working in the HCI area. They were instructed to identify their assessment of the potential

of the maxims to affect usability. The same instructions and forced-choice technique used with the group of known experts were used with this group. I received responses from 18 people in Respondent Set 2.

Some respondents predicted that there would be very little interrater reliability in the survey. They probably assumed that if designers have identified maxims that they believe are useful, they will have identified very different maxims (presumably based on individual evaluator differences and application domain differences), or they are interpreting the meaning of the maxims in very different ways. The correlation between the two sets of respondent ratings was .82, and the ratings for the top-rated maxims were consistent, so for the following analyses, the data for the two sets are combined (except where noted).

The Experts' Ratings

Each of the top nine maxims was rated as having a strong effect on usability (ratings of 4 or 5) relative to the others by at least 50% of the raters, with an average of 66% across the set of nine. Only 13% of the raters rated them as having a slight effect (ratings of 1 or 2). The top seven of these maxims are as follows:

#	<i>Maxim</i>	<i>Mean Rating</i>
1	Know thy user, and YOU are not thy user.	4.1
2	Things that look the same should act the same.	4.0
3	The information for the decision needs to be there when the decision is needed.	4.0
4	Error messages should actually mean something to the user and tell the user how to fix the problem.	3.9
5	Every action should have a reaction.	3.8
6	Everyone makes mistakes, so every mistake should be fixable.	3.8
7	Don't overload the user's buffers.	3.7
8	Consistency, consistency, consistency.	3.6
9	Minimize the need for a mighty memory.	3.6

The next two maxims appear to be alternative versions of higher-rated maxims (2 and 7, respectively).

For the next 11, people rating the maxims with a 4 or 5 outnumbered those rating them with 1 or 2. In this group, members of Respondent Set 1 (with known experience in standards work) gave average ratings of 3.6 to Maxims 10 and 15, whereas members of Respondent Set 2 gave the edge to Maxims 11 and 12 (with average ratings of 3.8 and 3.6, respectively). At least one way to interpret Maxims 10 and 15 is that they refer to an efficiency attribute, although Maxim 10



Is there a core set of maxims that represent design principles that most practitioners are using?

also may be interpreted as referring to removing "clutter" (information that is not relevant to the task) from a design or to breaking complex tasks down into simple steps. Maxim 11 appears to be related to Maxim 5, though they are clearly not synonymous. If every action has a reaction, that contributes to the user's knowing that at least something is happening and that it relates to what he or she has done.

#	<i>Maxim</i>	<i>Mean Rating</i>
10	Keep it simple.	3.5
11	The user should always know what is happening.	3.4
12	The more you do something, the easier it should be to do.	3.4
13	The user should control the system. The system shouldn't control the user. The user is the boss, and the system should show it.	3.4
14	Eliminate unnecessary decisions, and illuminate the rest.	3.3
15	The best journey is the one with the fewest steps. Shorten the distance between the user and the goal.	3.3
16	The user should be able to do what the user wants to do.	3.2
17	If I made an error, let me know about it before I get into REAL trouble.	3.2
18	You should always know how to find out what to do next.	3.2
19	The idea is to empower the user, not speed up the system.	3.2
20	Things that look different should act different.	3.1

Although the next eight maxims were rated as important by several professionals (an average of 32%), there was little consensus about their value; 48% of the respondents rated them at the lower end of the scale.

#	<i>Maxim</i>	<i>Mean Rating</i>
21	Don't let people accidentally shoot themselves.	3.0
22	Even experts are novices at some point. Provide help.	2.9
23	Keep it neat. Keep it organized.	2.9
24	Provide a way to bail out and start over.	2.9
25	Design for regular people and the real world.	2.9
26	The fault is not in thyself, but in thy system.	2.6
27	Color is information.	2.5
28	If it is not needed, it's not needed.	2.5

There was more consensus that the following maxims were not useful, or at least were clearly less useful than the more highly rated maxims.

#	<i>Maxim</i>	<i>Mean Rating</i>
29	Everything in its place, and a place for everything.	2.3
30	The user should be in a good mood when done.	2.2
31	If I made an error, at least let me finish my thought before I have to fix it.	1.9
32	Let people shape the system to themselves and paint it with their own personality.	1.8
33	"Cute" is not a good adjective for systems.	1.7
34	To know the system is to love it.	2.3

During the study, some of the early participants submitted proposals for maxims, and a few additional maxims were drawn from published lists of heuristics. To screen the maxims, I asked later participants in the survey to complete an optional second rating task. They scored the maxims based on how practical they would be to implement. I asked them to imagine that the forced-choice technique used during the main part of the study defined a scale of usability and to assign scores to these additional maxims using that scale. The two sets of data can't be compared directly, but several of the

TOP-RATED MAXIMS

- 1 Make objects, actions, and options visible.
- 2 Minimize the amount of information a user must maintain in short-term memory.
- 3 Provide good error messages that are expressed in plain language, precisely indicate the problem, and constructively suggest a solution.
- 4 People should not have to remember information across a dialog.
- 5 Provide the user with feedback and error correction capabilities.
- 6 Testing, testing, testing.
- 7 Support undo and redo.
- 8 No, you can't just explain it in the manual.

supplemental maxims were interesting and would be worth including in future research.

Top-Rated Maxims

In the box above are the maxims that were rated as the most implementable and as the most important for usability. Of these, only two appear to be unrelated to the earlier maxims tested. One is "Testing, testing, testing." Although it is not a maxim for evaluating a design heuristically, it is a reminder that the most complete evaluation of a design involves testing it with users to ensure that the most important guideline, "Know thy user, and *you* are not thy user," was correctly followed. It might be characterized as a kind of process heuristic. Gould and Lewis (1985) found that the principles of understanding the user, involving the user in the design process, and iterative prototyping and testing were not obvious to systems planners, designers, and developers. An appreciation of and a commitment to implement these principles may be part of the value that HCI professionals bring to the project team.

The other unrelated maxim is, "No, you can't just explain it in the manual." Again, although this is not an evaluation heuristic,

it is another kind of process heuristic. It is a reminder to the project team that the manual is not an alternative to correcting design errors found through heuristic analyses and testing. The related principle used by practitioners is that users won't read the manual. If you can't rely on a manual, the designer and developer have to come up with alternatives for training and guiding the user and for handling design errors.

How the Ratings Compare with Nielsen

Nielsen (1994) identified seven published lists of usability heuristics that represent a wide variety of perspectives on usability. He then applied these heuristics to a database he had created of 249 usability problems and rated them in terms of their explanatory value. With two possible exceptions, the top 10 unique maxims identified in this study (1-7 and 10-12) were similar to heuristics in Nielsen's lists. One of the exceptions is "Know thy user, and *you* are not thy user." Although some of the recommendations in the lists imply task and user analyses, "Know thy user" is the more general concept. These recommendations include the following:

- Use a natural dialogue.
- Speak the user's language (familiar terms and natural language).
- Use conceptual models familiar to the user.
- Understand the user's language.

The other possible exception is the maxim, "The information for the decision needs to be there when the decision is needed." The lists include recommendations such as these:

- Instructions should be visible or easily retrievable.
- Error messages should be clear.
- Provide lists of choices and allow users to pick.
- Provide default values for input fields.

These are compatible with the goal of providing information when it is needed, but they do not focus the design on the fundamental principle of ensuring that everything is aimed at supporting the user's efforts to complete a task. Without the

maxim identified by the practitioners, a software application could be developed that met the various lists of heuristics but was not useful or usable for its original purpose. On the other hand, the specific heuristics that are implied in the maxim may represent how it is operationalized during evaluation. Future work should try to clarify this more detailed structure and examine individual differences in its use.

Nielsen ordered the top 10 heuristics from the 7 lists he collected in terms of



A commitment to implement these principles may be part of the value that HCI professionals bring to the project team.

their ability to explain all the usability problems in his database. He ranked the top 10 heuristics by the ability of all the heuristics to explain the serious usability problems. Using my mapping of the top 10 maxims to the heuristics that seem to be implied in the maxims, I predict that the maxims that are most useful to experts are probably capable of identifying 85% of the usability problems in Nielsen's database and 89% of the serious usability problems in his database. Of the heuristics Nielsen identified as explaining serious usability problems, only "Prevent errors from occurring" and "Modeless interaction" were not clearly included in the maxims.

Nielsen's final list of nine heuristics includes visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, and helping users recognize, diagnose, and recover from errors. Of these, only error prevention does not rate high on experts' lists of maxims. This may be because the overall goal of usability itself is to design in such a way that errors are less likely to occur and, if they do occur, that they are easy to correct.

The relationship between Nielsen's results and this study is suggestive but leaves many questions unanswered. Although experienced professionals seem to feel that the top-rated maxims have a relatively large impact on usability, no one knows how experts actually apply them. Are the maxims applied

in the same way by experts, or is part of the virtue of the maxims that they represent core principles that each expert applies uniquely to reach a given level of usability?

No one knows the extent to which the maxims are subject to abuse. Some respondents argued that they either found all maxims useless (an interesting topic for individual differences research in itself) or that they had seen examples in which people without expertise applied maxims inappropriately and the overall design was worse than if no maxims at all had been applied. This is probably similar to the case of a novice chess player slavishly applying an expert's strategy even when the circumstances of the game have changed.

Conclusion

For the maxims to be most useful, similar maxims should be combined, and the maxims should be expressed in a way that conveys maximum value for designers and evaluators. The maxims should capture a wide range of design problems, and yet it should be obvious to an experienced professional how the maxims should be applied to specific cases. The maxims should be crisply worded and memorable. It should be possible to build tools based on the maxims that can support designers and developers during software development (e.g., Tullis, 1983), and it should be possible to train

- Engel, S. E., & Granda, R. E. (1975). *Guidelines for man/display interfaces* (Tech. Report TR 00.2720). Poughkeepsie, NY: IBM Corp.
- Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Erlbaum.
- Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM*, 26, 295-308.
- Hill, D. R. (1987). Interacting with future computers. *Future Computing Systems*, 2(1), 83-124.
- Nielsen, J. (1993). *Usability engineering*. Chestnut Hill, MA: AP Professional.
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In *Proceedings of the ACM CHI '94 Conference* (pp. 152-158). New York: Association for Computing Machinery.
- Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human-computer interaction*. Reading, MA: Addison-Wesley.
- Smith, S. L., & Mosier, J. N. (1986). *Guidelines for designing interface software*. Bedford, MA: MITRE Corp.
- Tullis, T. S. (1983). The formatting of alphanumeric displays: A review and analysis. *Human Factors*, 25, 657-682.
- Vanderdonckt, J. (1994). *Tools for working with guidelines* (bibliography). Belgium: University of Namur.
- Williges, B., & Williges, R. (1981). *Technical report prepared for engineering psychology programs* (ONR Contract No. N00014-810-k-0143). Washington DC: Office of Naval Research.
- Williges, R. C., Williges, B. H., & Elkerton, J. (1987). Software interface design. In G. Salvendy (Ed.), *Handbook of human factors*. New York: Wiley.

Arnold M. Lund is a distinguished member of technical staff at US West Advanced Technologies, 4001 Discovery Dr., Boulder, CO 80303, alund@acm.org. 



"Testing, testing, testing" is a reminder that the most complete evaluation of a design involves testing it with users to ensure that "Know thy user" was correctly followed.

people in the application of the maxims. Effective maxims may provide insights for a theoretical structure for usability that can be empirically extended to provide practical and powerful engineering guidance.

References

- Boff, K. R., & Lincoln, J. E. (Eds.). (1988). *Engineering data compendium: Human perception and performance*. Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.
- Brown, C. M. (1988). *Human-computer interface design guidelines*. Norwood, NJ: Ablex.
- Cooper, A. (1995). *About face*. Foster City, CA: IDG Books.