Abstract

This thesis compares an expanded set of usability principles, whose usage and derivation is made more explicit, with the set of shorter and more general heuristics used in Nielsen's heuristic evaluation. One theme investigated whether such principles can be used by either novices or experienced evaluators to identify more usability problems than with Nielsen's heuristics alone. A second examined the claim that in a heuristic evaluation only a small number of evaluators are required to uncover most of the problems with an interface.

First, it was shown that both groups of evaluators could make use of the author's expanded principles set (in various forms) to report or accurately predict more usability problems than similar subjects using the heuristics. However, for experienced subjects this was so only for problems assigned high severity ratings (on an 'open-ended' task), and for novices only when using subsets of the evaluation materials (on a 'closed', complex task).

Second, these experiments offer little support for Nielsen's claim. The numbers of experienced and novice subjects required to identify 75% of respective problem totals were much greater than the 3 to 5 claimed for experienced evaluators and (in most cases) greater than the 14 claimed for novices. For experienced subjects, this was also true for high-severity problems. Cumulative plots of problem count showed little agreement with those on which the '3 to 5' claim is based, exhibiting lower detection rates than the minimum (of 33%) required by the claim. Further lack of agreement was shown when counting correct predictions ('hits') rather than reported (or 'detected') problems, the only exception being for very brief and contrived tasks. It was demonstrated that comparable detection rates could be achieved by assigning problems to higher categories (or problem types) rather than counting by unique problem tokens (UPTs) or problem instances.

Preface