Click to verify

Severity ratings can be used to allocate the most resources to fix the most serious problems and can also provide a rough estimate of the need for additional usability efforts. If the severity ratings indicate that several disastrous usability problems remain in an interface, it will probably be unadvisable to release it. But one might decide to go ahead with the release of a system with several usability problems if they are all judged as being cosmetic in nature.The severity of a usability problem is a combination of three factors:The frequency with which the problem occurs: Is it common or rare?The impact of the problem if it occurs: Will it be easy or difficult for the users to overcome?The persistence of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?Finally, of course, one needs to assess the market impact of the problem since certain usability problems can have a devastating effect on the popularity of a product, even if they are "objectively" quite easy to overcome. Even though severity has several components, it is common to combine all aspects of severity in a single severity rating as an overall assessment of each usability problem in order to facilitate prioritizing and decision-making.The following 0 to 4 rating scale can be used to rate the severity of usability problems:0 = I don't agree that this is a usability problem at all1 = Cosmetic problem only: need not be fixed unless extra time is available on project2 = Minor usability problem: fixing this should be given low priority3 = Major usability problem: important to fix, so should be given high priority4 = Usability catastrophe: imperative to fix this before product can be releasedSeverity Ratings in Heuristic EvaluationIt is difficult to get good severity estimates from the evaluators during a heuristic evaluation session when they are more focused on finding new usability problems. Also, each evaluator will only find a small number of the usability problems, so a set of severity ratings of only the problems found by that evaluator will be incomplete. Instead, severity ratings can be collected by sending a questionnaire to the evaluators after the actual evaluation sessions, listing the complete set of usability problems that have been discovered, and asking them to rate the severity of each problem. Since each evaluator has only identified a subset of the problems included in the list, the problems need to be described in reasonable depth, possibly using screendumps as illustrations. The descriptions can be synthesized by the evaluation observer from the aggregate of comments made by those evaluators who had found each problem (or, if written evaluation reports are used, the descriptions can be synthesized from the descriptions in the reports). These descriptions allow the evaluators to assess the various problems fairly easily even if they have not found them in their own evaluation session. Typically, evaluators need only spend about 30 minutes to provide their severity ratings. It is important to note that each evaluator should provide individual severity ratings independently of the other evaluators.Often, the evaluators will not have access to the actual system while they are considering the severity of the various usability problems. It is possible that the evaluators can gain additional insights by revisiting parts of the running interface rather than relying on their memory and the written problem descriptions. At the same time, there is no doubt that the evaluators will be slower at arriving at the severity ratings if they are given the option of interacting further with the system. Also, scheduling problems will sometimes make it difficult to provide everybody with computer access at convenient times if special computer resources are needed to run a prototype system or if software distribution is limited due to confidentiality considerations.My experience indicates that severity ratings from a single evaluator are too unreliable to be trusted. As more evaluators are asked to judge the severity of usability problems, the quality of the mean severity rating increases rapidly, and using the mean of a set of ratings from three evaluators is satisfactory for many practical purposes. Design creates stories, and stories create memorable experiences, and great experiences have this innate ability to change the way in which we view our world. Christian SaylorDelight the users with the best experience by providing exactly what they want is the ultimate goal of any product. Thus we do extensive researches and broad studies to get a better understanding of our target users and their desired needs. By examining the output of these researches and studies, we create certain problems/pain points and then its solutions that act as the baseline of our design process. As a sum of all these process and solutions, we design a clickable prototype on which we can do the usability testings. This is one of the crucial processes before the design goes live. But this is not the point where our prototype meets the real users.We need to do an expert review on the design before its being reviewed by the real users. Because we need to make sure that we have implemented all the solutions we found in the previous UX processes flawlessly. Its not that simple as it sounds. So here we run a process called Heuristic Evaluation.What is Heuristic EvaluationHeuristic Evaluation is a process to perceive the usability issues in our product interface design and solving them based on its severity. In simple words, it is a process to evaluate the usability of a design.As I mentioned above, this is not a process done by the end usersThen who??Usability Experts are the people who are going to review the design in this expert review process. UX designers having years of experience in working for different kind of users, solving complex design problems for different kinds of products or services are the usability experts. They have a keen understanding of the users perspective as they are working more closely to the end users.To get the best result in Heuristic Review, we need to have 35 usability experts to review the design. They are not just evaluating the design based on their own perspective or previous experiences but using a set of predetermined guidelines. These guidelines are the key factor in this Heuristic Review. In most cases, we use Jacobs Neilsons 10 usability heuristics as the base guidelines to do this process. In this publication, he has mentioned about 10 heuristics that can be incorporated into our evaluation process. This publication was released in 1994 but still, its relevant and we use it. We can customize these guidelines based on the nature of our product and users or the activities happening inside our product.How to do Heuristic EvaluationHeuristic Review can be done at any stage of the UX process if we want to, in the beginning, in the middle or at the end. But in my point of view, its better to run a heuristic review once we have a completed clickable prototype whether its a greyscale wire-frame or a colorful visual design to reduce time and budget. But if you run this process at the initial stages, we can find and solve some obvious mistakes that can happen there in the beginning. Solving minor problems would help us to figure out the major issues in our designs/findings. In this blog, I am describing how to do a heuristic review on a clickable prototype.The process behind the Heuristic Review is not that complex if we have the perfect usability experts and guidelines/heuristics. Each usability experts will review the design based on those predetermined guidelines. During the evaluation process, they will point out all the issues on a paper and give an appropriate rating based on its severity. So here comes another very important term in Heuristic Review, the Severity RatingThe severity of usability issues can be determined by three main factors.How frequently does the problem occur: Very often or rarely?The impact of the problem on users: Would this problem be a huge headache for the users or they can easily overcome the situation?Persistence of the problem -Is this something that the user can overcome once they know about it, or would this be a big concern for the user again and again?Severity Rating in Heuristic EvaluationUsability experts will add a Severity Rating to each usability issues based on these above 3 factors. They will rate the issues from 0 to 4 as follows,Using this parameter we can find out major issues and allocate more resource to solve them quickly. Likewise, we can allocate resource and solve issues for the rest of the findings as well. All we need to do is prepare accurate guidelines/heuristics for the review process. Here you can find Jacobs Neilsons 10 usability heuristics with examples, which will give you an understanding about how to choose heuristics.Heuristic Evaluation and Usability TestingThe fact is, the Heuristic evaluation itself is a usability testing process. When it is done by usability experts against a set of guidelines, its called heuristic review. In the other scenario, End users are the people who are testing our app/prototype. We need to do both these usability testings to get the best result in the UX/UI design. So it is completely meaningless to compare both of these processes to see which one is better/powerful. Both of these are useful in their own ways.We will find merits and demerits of both these process if we are comparing one with another. If you want to complete the whole UX process in a short span of time, you can avoid Usability testing with end users, since it takes more time. But if you want the best result or you have enough time to invest, then go ahead with both these processes. Jeff RubinIn Jeffs influential 1994 book, he outlined the following scale for problem severity:4: Unusable: The user is not able to or will not want to use a particular part of the product because of the way that the product has been designed and implemented.3: Severe: The user will probably use or attempt to use the product here, but will be severely limited in his or her ability to do so.2: Moderate: The user will be able to use the product in most cases, but will have to undertake some moderate effort in getting around the problem.1: Irritant: The problem occurs only intermittently, can be circumvented easily, or is dependent on a standard that is outside the products boundaries. Could also be a cosmetic problem.Dumas and RedishJoe Dumas and Ginny Redish, in their seminal book, A Practical Guide to Usability Testing, offer a similar categorization as Rubin and Nielsen but add a global versus local dimension to the problems. The idea is that if a problem affects the global navigation of a website, it becomes more critical than a local problem only affecting, say, one page.Level 1: Prevents Task CompletionLevel 2: Creates significant delay and frustrationLevel 3: Problems have a minor effect on usabilityLevel 4: Subtle and possible enhancements/suggestionsChauncey WilsonChauncey Wilson suggests that usability severity scales should match the severity rating of bug-tracking systems in a company. He offers a five-point scale with the following levels. Earlier, hes used a similar four-point variant[pdf].Level 1: Catastrophic error causing irrevocable loss of data or damage to the hardware or software. The problem could result in large-scale failures that prevent many people from doing their work. Performance is so bad that the system cannot accomplish business goals.Level 2: Severe problem, causing possible loss of data. User has no workaround to the problem. Performance is so poor that the system is universally regarded as pitiful.Level 3: Moderate problem causing no permanent loss of data, but wasted time. There is a workaround to the problem. Internal inconsistencies result in increased learning or error rates. An important function or feature does not work as expected.Level 4: Minor but irritating problem. Generally, it causes loss of data, but the problem slows users down slightly. There are minimal violations of guidelines that affect appearance or perception, and mistakes that are recoverable.Level 5: Minimal error. The problem is rare and causes no data loss or major loss of time. Minor cosmetic or consistency issue.The Wilson and Dumas & Redish scales have the more severe problem with lower numbers. That is because in the early days of computing, severe bugs were called level 1 bugs and those had to be fixed before product release (Dumas, Personal Communication 2013). In this scale, the problems are defined in terms of data loss rather than their impact on users performance or emotional state.Molich & JeffriesRolf Molich is famous for his series of comparative usability evaluations (CUE). Hes also famous for reviewing and writing (often critically) about the quality of usability reports. He and Robin Jeffries offered a three-point scale.1. Minor: delays user briefly.2. Serious: delays user significantly but eventually allows them to complete the task.3. Catastrophic: prevents user from completing their task.This three-point approach is simpler than others but tends to rely heavily on how the problem impacts time on task.Our ApproachOriginally we started with a 7-point rating scale where evaluators assigned the problem severity a value from cosmetic (1) to catastrophic (7) but we found it was difficult to distinguish easily between levels 2 and 6. We reduced this to a four-point scale similar to Rubin, Nielsen and Dumas/Redish above and treated them more as categories than a continuum.While there was much less ambiguity with four points, we still found a murky distinction between the two middle levels in both assigning the severity and reporting the levels of problems to clients.So we reduced our severity scale to just three levels, along with one for insights, user suggestions or positive attributes.1. Minor: Causes some hesitation or slight irritation.2. Moderate: Causes occasional task failure for some users; causes delays and moderate irritation.3. Critical: Leads to task failure. Causes user extreme irritation.Insight/Suggestion/Positive: Users mention an idea or observation that does or could enhance the overall experience.SummaryIve put abbreviated versions of these scales below in the table to show the similarities in some of the terms and levels. Ive also aligned the scales so higher numbers indicate more severe problems.LevelNielsenRubinDumasWilsonMolich & JeffriesSauro0Not a ProblemInsight/ Suggestion/ Positive1CosmeticIrritantSubtle & possible enhancements/ suggestionsMinor cosmetic or consistency issueMinor (delays user briefly)Minor : Some hesitation or slight irritation2MinorModerateProblems have a minor effect on usabilityMinor but irritating problem3MajorSevereCreates significant delay and frustrationModerate problemSerious (delays user significantly but eventually)Moderate: Causes occasional task failure for some users; causes delays and moderate irritation4UnusablePrevents Task CompletionSevere problemCritical: Leads to task failure. Causes user extreme irritation.5CatastropheCatastrophic errorCatastrophic (prevents user from completing their task)Some lessons from these problem severity levels:Dont obsess over finding the right number of categories or labels: Three categories is probably sufficient, but merging scales with bug tracking levels or having more levels to generate more internal buy-in are both legitimate reasons to have more points. Once you pick a system, try and stick with it to allow comparison over time.There will still be inter-rater disagreement and judgment calls: These are rough guides, not precise instruments. Different evaluators will disagree, despite the clarity of the severity levels. One of the best approaches is to have multiple evaluators rate the severity independently, calculate the agreement, and then average the ratings.The numbers assigned to each level are somewhat arbitrary: Dont obsess too much over whether higher severity problems should have higher numbers or lower ones. I prefer the latter, but its the order that has meaning. While the intervals between severities of 1, 2 and 3 are likely different, the ranks can be used for additional analysis when comparing different evaluators or problems severity and frequency.Dont forget the positives: Dumas, Molich & Jeffries wrote a persuasive article talking about the need for pointing out positive findings. While a usability test is usually meant to uncover problems, understanding the positives encourages the developers and doesnt make you or your team come across as the constant harbingers of bad news.Treat frequency separately from severity: We report the frequency of an issue along with its severity. When possible, we have a separate analyst rate the severity of a problem without knowing its frequencya topic for a future blog.Thanks to Joe Dumas for commenting on an earlier draft of this article. Heuristic Evaluation Chapter 9 Mohamad Eid Heuristic Evaluation Introduction to Heuristic Evaluation Phases of Heuristic Evaluation How to perform the Heuristic Evaluation Heuristics Examples Severity Rating Debriefing Summary Mohamad EidHeuristic Evaluation Developed by Jakob Nielsen Helps find usability problems in a UI design Small set (3-5) of evaluators examine UI independently check for compliance with usability principles (heuristics) different evaluators will find different problems evaluators only communicate afterwards findings are then aggregated Can perform on working UI or on sketches Mohamad EidWhy Multiple Evaluators? Every evaluator doesnt find every problem Good evaluators find both easy & hard ones Mohamad EidHeuristic Evaluation Evaluators goes through UI several times inspects various dialogue elements compares with list of usability principles consider any additional principles or results that come to mind Usability principles Nielsens heuristics supplementary list of category-specific heuristics competitive analysis & user testing of existing products Use violations to redesign/fix problems Mohamad EidPhases of Heuristic Evaluation 1) Pre-evaluation training give evaluators needed domain knowledge and information on the scenarios 2) Evaluation individuals evaluate and then aggregate results 3) Severity rating determine how severe each problem is (priority) 4) Debriefing discuss the outcome with design team Mohamad EidHow to Perform Evaluation At least two passes for each evaluator first to get feel for flow and scope of system second to focus on specific elements If system is walk-up-and-use or evaluators are domain experts, then no assistance needed otherwise might supply evaluators with scenarios Each evaluator produces list of problems explain why with reference to heuristic or other information be specific and list each problem separately Mohamad EidHeuristics H1-1: Simple and natural dialog H1-2: Speak the users language H1-3: Minimize users memory load H1-4: Consistency H1-5: Feedback H1-6: Clearly marked exits H1-7: Shortcuts H1-8: Precise and constructive error messages H1-9: Prevent errors H1-10: Help and documentation Mohamad EidHeuristics H2-1: Visibility of system status keep users informed about what is going on example: pay attention to response time 0.1 sec: no special indicators needed 1.0 sec: user tends to lose track of data 10 sec: max. duration if user to stay focused on 1 action for longer delays, use percent-done progress bars Mohamad EidHeuristics H2-2: Match between system and real world speak the users language follow real world conventions Mac desktop Dragging disk to trash should delete it, not eject it Mohamad EidHeuristics H2-3: User control and freedom exits for mistaken choices, undo, redo dont force down fixed paths Wizards must respond to Q1 before going to next for infrequent tasks modem config. not for common tasks Good for beginners have 2 versions WinZip Mohamad EidHeuristics H2-4: Consistency & standards Mohamad EidHeuristics H2-5: Error prevention H2-6: Recognition rather than recall make objects, actions, options, and directions visible or easily retrievable MS Web Pub. Wiz. Before dialing asks for id & password When connecting asks again for id & pw Mohamad EidHeuristics H2-7: Flexibility and efficiency of use accelerators for experts (e.g., gestures, kb shortcuts) allow users to tailor frequent actions (e.g., macros) H2-8: Aesthetic and minimalist design no irrelevant information in dialogues Mohamad EidHeuristics H2-9: Help users recognize, diagnose, and recover from errors error messages in plain language precisely indicate the problem constructively suggest a solution Mohamad EidHeuristics H2-10: Help and documentation easy to search focused on the users task list concrete steps to carry out not too large Mohamad EidExamples Cant copy info from one window to another violates Minimize the users memory load (H1-3) fix: allow copying Typography uses mix of upper/lower case formats and fonts violates Consistency and standards (H2-4) slows users down probably wouldnt be found by user testing fix: pick a single format for entire interface Mohamad EidSeverity Rating Used to allocate resources to fix problems Estimates of need for more usability efforts Combination of frequency impact persistence (one time or repeating) Should be calculated after all evaluations are in Should be done independently by all judges Mohamad EidSeverity Ratings (cont.) 0 - dont agree that this is a usability problem 1 - cosmetic problem 2 - minor usability problem 3 - major usability problem; important to fix 4 - usability catastrophe; imperative to fix Mohamad EidSeverity Ratings Example [H1-4 Consistency] [Severity 3][Fix 0] The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function. Mohamad EidDebriefing Conduct with evaluators, observers, and development team members Discuss general characteristics of UI Suggest potential improvements to address major usability problems Add ratings on how hard things are to fix Make it a brainstorming session little criticism until end of session Mohamad EidSummary Heuristic evaluation is a discount method Have evaluators go through the UI twice Ask them to see if it complies with heuristics note where it doesnt and say why Combine the findings from 3 to 5 evaluators Have evaluators independently rate severity Discuss problems with design team Alternate with user testing Mohamad Eid DMnvwd Dankie go raibh maith agaibh WAD MAHAD SAN TAHAY GADDA GUEY Asante Urakoze Mohamad Eid Identifying usability problems through a heuristic evaluation is the first step towards eliminating problems and improving the interface. Once this step has been taken, severity ratings should be fabricated for each problem. Ranking of usability problems by severity helps to determine those that should be addressed, given that not all problems can be fixed due to constraints on the design life cycle (e.g. budget, schedule, etc.,). The ratings also help in the allocation of resources for addressing the user interface problems [5]. Before a usability problem can be rated according to severity, a definition of severity must be understood. According to Nielsen [13], severity is considered to be a combination of three factors: frequency, impact, and persistence. Frequency ranges from common problems to rare ones. Impact defines the ease or difficulty with which a user can overcome a problem. Finally, persistence varies from the one-time problem that can be overcome to the problem that continuously repeats itself becoming tiresome to the user. Naturally, the severity of a usability problem increases as the level of these factors increase. To facilitate the severity rating process, these three factors are combined into one single severity rating which is an overall assessment of each usability problem. As previously mentioned, Nielsen [13] states that evaluators have difficulty formulating severity ratings for each usability problem during the evaluation process because they are more focused on finding new usability problems. Another disadvantage of producing severity ratings during the evaluation process is that each evaluator will not find all the usability problems in the system. Therefore, the severity ratings will only reflect those problems found by the evaluator and will be incomplete. To solve this dilemma, severity ratings for all the usability problems can be found by sending a questionnaire to each inspector once the evaluations process has been completed [13]. There is concern that there may be some bias on the part of the evaluators performing the severity ratings. It might be expected that evaluators will rank the problems they found as more serious. However, Nielsen [13] found that any given evaluator's severity rating of a usability problem was essentially independent of whether that evaluator had found that problem. There was a positive correlation between the evaluators' ratings and the number of evaluators having found each problem [13]. This correlation is not to bias in the severity assessment since individual evaluators do not know how many other evaluators had found each problem. This correlation is due to the fact that the more severe usability problems are found more frequently by heuristic evaluations [13]. Again, the reliability of the severity ratings depends upon the number of evaluators used with the method. Ratings from three to four evaluators would seem to be satisfactory for many practical purposes. Severity ratings from a single evaluator are naturally considered bias and too unreliable to be trusted [13]. But one must remember that given certain circumstances, one evaluator may be all the resources available to perform a heuristic evaluation. Return to Home Page

What is severity rating. Heuristieken. Heuristic severity rating. What is a good severity rate. Question 7 what is a severity rating in a heuristic evaluation.