# Using Heuristics to Evaluate the Playability of Games

**Heather Desurvire**
Behavioristics, Inc.
Playability/Usability Specialist
520 Washington Blvd. #179
Marina del Rey, CA 90292 USA
+ 1 310 823-6543
heather3@gte.net

**Martin Caplan**
Avatar-X
Game Producer/Designer
2466 40th Avenue
San Francisco, CA  94116 USA
+1 415 341-6372
marty@avatar-X.com

**Jozsef A. Toth**
Institute for Defense Analyses
Research Staff Member
4850 Mark Center Drive
Alexandria, VA 22311 USA
+1 703 578-2867
jtoth@ida.org

## Abstract

Heuristics have become an accepted and widely used adjunct method of usability evaluation in Internet and software development.  This report introduces Heuristic Evaluation for Playability (HEP), a comprehensive set of heuristics for playability, based on the literature on productivity and playtesting heuristics that were specifically tailored to evaluate video, computer, and board games. These heuristics were tested on an evolving game design to assess their face validity and evaluation effectiveness compared to more standard user testing methodologies. The results suggest that HEP identified qualitative similarities and differences with user testing and that HEP is best suited for evaluating general issues in the early development phases with a prototype or mock-up.  Combined with user studies, HEP offers a new method for the HCI game community that can result in a more usable and playable game.

## Categories and Subject Descriptors:

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems–*evaluation/methodology*

**General Terms:** Measurement, Human Factors.

## Keywords:

Heuristics, playability, playtesting, design guidelines, video games, computer games, games, evaluation, usability, user testing.

## INTRODUCTION

Heuristics are design guidelines which serve as a useful evaluation tool for both product designers and usability professionals.  In the software productivity industry, heuristics have typically been used to evaluate the usability of interfaces.  The goals of software productivity are to make the software interface easy to learn, use, and master, and somewhat oppose design goals for games, usually characterized as "easy to learn, difficult to master" [6].  In the realm of game playability, there is a need to go beyond basic interface game usability evaluation to assess additional properties of the game experience including game play, story, and mechanics.

## Definitions of Game Heuristic Categories

The four game heuristic categories are defined as the following: *game play* is the set of problems and challenges a user must face to win a game; *game story* includes all plot and character development; *game mechanics* involve the programming that provides the structure by which units interact with the environment; and *game usability* addresses the interface and encompasses the elements the user utilizes to interact with the game (e.g. mouse, keyboard, controller, game shell, heads-up display).

## History of Software and Game Heuristic Development

Nielsen [7] developed a list of heuristics that were aimed for use in productivity software.  Software productivity studies by Desurvire, et al. [1,2] demonstrated the effectiveness of these heuristics when combined with user studies.  In game development, there is a need to develop a corresponding set of heuristics.  Thus far, game heuristics have been developed by several individuals, groups, and professionals in the game industry, and researchers in the HCI community.  In 1982, Malone constructed a list of heuristics for instructional games [6].  In 2002, Federoff [4,5] compiled a list of game heuristics from a case study at a game development company and compared them with current game industry guidelines and J. Nielsen's heuristics from 1994 [7].  Since 2001, game designers Falstein and Barwood have been writing the first of 400 rules of game design "that can be used by designers to make better games" [3].

## Verification of Game Heuristics

Despite progress in game heuristic development, prior to this project, game playability heuristics have yet to be compiled into a comprehensive list or verified in any way. The intent of this particular project was to begin developing a list of verified heuristics.  A comprehensive list of game heuristics (see Table 1) was developed and comparatively tested against traditional user study methodologies during the critique of a new game design.  The results were examined to evaluate the face validity of the individual heuristics, as well as identify the strengths, weak-

nesses, and qualitative differences of the separate methodologies.

## PROCEDURE
### The Game
A new game at the beginning of the development design cycle was utilized as the mechanism to study the heuristics' efficacy. The game design concept was prototyped using Macromedia Flash. The displays consisted of screen shots that allowed users to navigate throughout the shell of the game but did now allow any game play.

### Heuristic Evaluation for Playability (HEP)
The HEP heuristics (see Table 1) were based on the current literature and reviewed by several playability experts and game designers. The playability evaluator performed the Heuristic Evaluation for Playability (HEP) while focusing on how each heuristic was supported or violated and then defined the playability issue. Alternative solutions for resolving the playability issues were generated by both the evaluator and the game designer.

### User Studies
After the HEP study was completed, four prospective users engaged in two-hour playability sessions. Each session was organized as a one-on-one evaluation session, in an environment similar to the one where they would actually play the game. Participants were given instructions to begin the game, asked to think aloud, and asked several probing questions while using the prototype. The player was then thanked, debriefed and asked to fill out a satisfaction questionnaire. The evaluator recorded a log of the player's actions, comments, failures, missteps, and coded each of these as a positive player experience or a negative player experience. A positive experience was defined as anything that increased their pleasure, immersion, and the challenge of the game. A negative experience was defined as any situation where the player was bored, frustrated, or wanted to quit the game. Probes and the players' comments were used to verify any assumptions made by the evaluator. After the sessions were complete, playability design problems were identified and alternative design solutions were generated. Each issue was assigned a severity score based on its consequence and the user's ability to continue with game play.

## ANALYSIS
The results from user studies are considered the benchmark of game evaluation tools. The HEP results were compared to the user study results, highlighting the contribution each made to the next iteration of design.

## RESULTS
The results, listed below, validated the usefulness of the heuristics and provided insight into the role of each methodology in the design cycle.

### Validating the Heuristics
The HEP heuristics proved effective in uncovering playability issues, especially in the Game Story and Game Usability categories. Six out of 8 Game Story heuristics and 11 out of 12 Game Usability heuristics were useful in uncovering playability issues. This demonstrates that for the initial game design, HEP is extremely useful in the categories of Game Story and Game Usability. Approximately half of the HEP issues were found in the Game Play (7 of 16) and Game Mechanics (4 of 7) categories. Because the game was at the beginning of the design cycle, the game play and game mechanics were not fully developed. Future research efforts could potentially find these additional HEP heuristics to also prove helpful in later phases of the design cycle.

### Comparing HEP Issues with User Studies Issues
The total number of issues identified from HEP was greater than the number of issues found from the user study (see Figure 1), but the nature of the user study issues were more specific to the game. The user study issues were very specific to the interface, such as the terminology, characters, and verbiage.
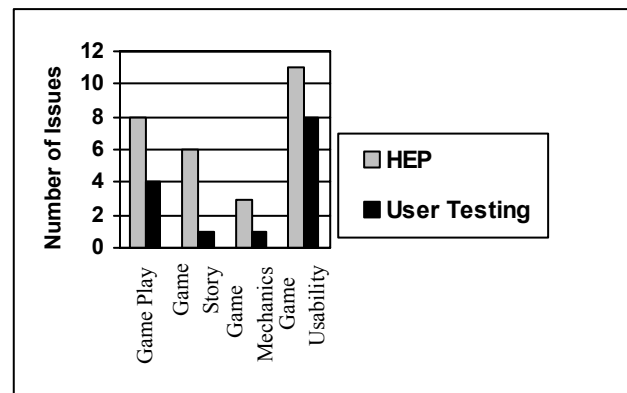


**Figure 1. Comparison of Number of Issues**

There was much overlap in the issues found by HEP and user studies. In the Game Play category there were two overlapping issues, which were mainly focused on player goals and motivations. HEP identified four additional issues, which were Game Play Heuristics 4, 6, 7 and 13 (see Table 1 for specific heuristics). For example, when using HEP, Game Play Heuristic 13: The first player action is painfully obvious and should result in immediate positive feedback, identified guidelines regarding the pace of game play that user studies did not directly indicate. Alternatively, user studies identified problems with Game Play Heuristics 1 and 5 which would only be found from observing the actual players playing.

In the category of Game Story there was one overlapping, high-severity issue found by both HEP and user studies. HEP identified five additional issues such as, Game Story Heuristic 3: Player spends time thinking about possible

**Table 1.  Heuristics for Evaluating Playability (HEP)**

| | **Heuristic and Description** |
|---|---|
| **Game Play** | |
| 1 | Player's fatigue is minimized by varying activities and pacing during game play. |
| 2 | Provide consistency between the game elements and the overarching setting and story to suspend disbelief. |
| 3 | Provide clear goals, present overriding goal early as well as short-term goals throughout play. |
| 4 | There is an interesting and absorbing tutorial that mimics game play. |
| 5 | The game is enjoyable to replay. |
| 6 | Game play should be balanced with multiple ways to win. |
| 7 | Player is taught skills early that you expect the players to use later, or right before the new skill is needed. |
| 8 | Players discover the story as part of game play. |
| 9 | Even if the game cannot be modeless, it should be perceived as modeless. |
| 10 | The game is fun for the Player first, the designer second and the computer third.  That is, if the non-expert player's experience isn't put first, excellent game mechanics and graphics programming triumphs are meaningless. |
| 11 | Player should not experience being penalized repetitively for the same failure. |
| 12 | Player's should perceive a sense of control and impact onto the game world.  The game world reacts to the player and remembers their passage through it.  Changes the player makes in the game world are persistent and noticeable if they back-track to where they've been before. |
| 13 | The first player action is painfully obvious and should result in immediate positive feedback. |
| 14 | The game should give rewards that immerse the player more deeply in the game by increasing their capabilities (power-up), and expanding their ability to customize. |
| 15 | Pace the game to apply pressure but not frustrate the player.  Vary the difficulty level so that the player has greater challenge as they develop mastery. Easy to learn, hard to master. |
| 16 | Challenges are positive game experiences, rather than a negative experience (results in their wanting to play more, rather than quitting). |
| **Game Story** | |
| 1 | Player understands the story line as a single consistent vision. |
| 2 | Player is interested in the story line.  The story experience relates to their real life and grabs their interest. |
| 3 | The Player spends time thinking about possible story outcomes. |
| 4 | The Player feels as though the world is going on whether their character is there or not. |
| 5 | The Player has a sense of control over their character and is able to use tactics and strategies. |
| 6 | Player experiences fairness of outcomes. |
| 7 | The game transports the player into a level of personal involvement emotionally (e.g., scare, threat, thrill, reward, punishment) and viscerally (e.g., sounds of environment). |
| 8 | Player is interested in the characters because (1) they are like me; (2) they are interesting to me, (3) the characters develop as action occurs. |
| **Mechanics** 1 | Game should react in a consistent, challenging, and exciting way to the player's actions (e.g., appropriate music with the action). |
| 2 | Make effects of the Artificial Intelligence (AI) clearly visible to the player by ensuring they are consistent with the player's reasonable expectations of the AI actor. |
| 3 | A player should always be able to identify their score/status and goal in the game. |
| 4 | Mechanics/controller actions have consistently mapped and learnable responses. |
| 5 | Shorten the learning curve by following the trends set by the gaming industry to meet user's expectations. |
| 6 | Controls should be intuitive, and mapped in a natural way; they should be customizable and default to industry standard settings. |
| 7 | Player should be given controls that are basic enough to learn quickly yet expandable for advanced options. |
| **Usability** 1 | Provide immediate feedback for user actions. |
| 2 | The Player can easily turn the game off and on, and be able to save games in different states. |
| 3 | The Player experiences the user interface as consistent (in control, color, typography, and dialog design) but the game play is varied. |
| 4 | The Player should experience the menu as a part of the game. |
| 5 | Upon initially turning the game on the Player has enough information to get started to play. |
| 6 | Players should be given context sensitive help while playing so that they do not get stuck or have to rely on a manual. |
| 7 | Sounds from the game provide meaningful feedback or stir a particular emotion. |
| 8 | Players do not need to use a manual to play game. |
| 9 | The interface should be as non-intrusive to the Player as possible. |
| 10 | Make the menu layers well-organized and minimalist to the extent the menu options are intuitive. |
| 11 | Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels. |
| 12 | Art should be recognizable to player, and speak to its function. |

story outcomes. At the time of testing the game story was at a cursory stage. These results illustrate how HEP can allow insight into the user's point of view and identify possible issues and remedies, even before actual user interaction is possible. These are the types of issues user studies would not find until they were already a problem designed into the game.

In the category of Game Mechanics one overlapping issues was uncovered. This was Game Mechanics Heuristic 7, which deals with a directly observable issue. In addition to this, HEP found one high-severity issue and one low-severity issue which would be more pertinent in the future when the game would be more developed. These issues were as follows, Mechanics Heuristic 1: Maximize potential so the Player will have a high expressive potential experience, and Mechanics Heuristic 3: A player should always be able to identify their score/status and goal in the game.

In the Game Usability category, there were eight overlapping issues and three issues unique to HEP, such as Game Usability Heuristic 11: Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels. In this category, although there was much overlap between the issues uncovered, the two methods produced qualitatively different information. The HEP results identified general interface design issues, whereas the user study results identified specific problems with the interface. The specific nature of the problems identified with user studies allowed for the creation of specific design solutions.

### The Advantage of User Studies
Overall, user studies uncovered many of the same issues found by HEP, but also identified specific behaviors and problems that could only be found by observing user play. The user studies findings highlighted issues specific to the game; boredom, challenge and pace level, as well as terminology. These issues were not found through HEP, whose benefit was in ensuring general game principles.

### CONCLUSION
User testing is the benchmark of any playability evaluation, since a designer can never completely predict user behavior. HEP appears to be very useful for creating highly usable and playable game design, particularly in the preliminary design phase prior to expensive prototypes. Future research may uncover how HEP could be utilized during a more advanced stage of design. Another area for future research would be to examine the differ-

ence in HEP results when performed by designers versus playability experts, as well as the reliability between multiple evaluators.

In conclusion, HEP is helpful in early game design and user studies are best suited to find specific problems once they already exist. HEP facilitates thinking about the design from the user's point of view. This process allows HEP to be proactive in avoiding expensive design problems. Still, we must rely on user testing since no matter how much we think we understand game players and humans, their behavior is still unpredictable. Once we observe the player's behavior, we have the specific knowledge necessary to resolve the design problems.

### REFERENCES
1. Desurvire, H. (1994). Faster, Cheaper: Are Usability Inspection Methods as Effective as Empirical Testing? In Nielsen, J. and Molich, R.L. (Eds.) *'Usability Inspection Methods'*, in "Usability Inspection Methods", New York: John Wiley & Sons, 173-202.

2. Desurvire, H. Kondziela, J., and Atwood, M. (1992a). What is Gained and Lost When Using Evaluation Methods Other Than Empirical Testing. In the proceedings of the ACM conference, CHI1992, collection of abstracts (1992), 125-126.

3. Falstein, N. and Barwood, H. The 400 Project. Available at http://theinspiracy.com/400_project.htm.

4. Federoff, M. User Testing for Games: Getting Better Data Earlier. *Game Developer Magazine* (June 2003).

5. Federoff, M. Heuristics and Usability Guidelines for the Creation and Evaluation of FUN in Video Games. Thesis at the University Graduate School of Indiana University, Dec. 2002 (melissafederoff@yahoo.com).

6. Malone, T.W. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In John C. Thomas and M. L. Schneider (Editors), *Human Factors in Computing Systems*, Norwood, NJ: Ablex Publishing Corporation, 1982.

7. Nielsen, J. Heuristic evaluation. In Nielsen, J. and Molich, R.L (eds.) *'Usability Inspection Methods'*, New York: John Wiley & Sons, 1994.