

Abdul Alzaid
Professor, Dondi
Interaction Design
09/26/2013

USABILITY REPORT

This report focused on measuring four out of five metrics from 8 different users. The metrics used are four, but only three are applied for each user. The first two are learnability, and efficiency which are measured in time, and only one at a time was measured on each user. The other two are, error which is measured in how many errors the user committed, and the second one is, satisfaction which is measured on a scale from 1 to 10. The users focused on using three different systems during the experiment, iTunes, Spotify, and Google Music to perform three different tasks.

The first task the users are asked to perform was creating three different playlists on each application, each of which should contain 10 different songs. The second task was, to search for five different songs on the online store. The third task the users were asked to do was to import 10 different songs from a local folder on the desktop that contained a list of song.

The users' learnability and efficiency have been measured depending on their User Proficiency Profiles. For example, some of the users are novice, and have neither used the three different applications and nor performed any of the three tasks before. Whereas, some of the users have used one or two of the applications, and performed one, or all of the different tasks before, which, depending on their level of interface, and domain, or task knowledge, that can be referred to as either first-time user or knowledgeable intermittent user, or expert frequent user.

Tasks Overview:

Task 1: The first task focused on testing the participants on creating three different playlists each playlist will have five songs from two different artists, and the participants are asked to name each playlist with the name of the two artists. So, the total number of songs will be 30 songs added to the playlists. The experiment will be timed from the beginning till the end, to measure the learnability, or the efficiency. The number of errors the participants were making was also measured. Finally, the overall satisfaction rate was measured out of 10, with 10 being the highest.

Task 2: The second the task is a little bit simpler than the first task, and it asked the users to search for 5 different songs by using the three different applications' online stores. The experiment will be timed from the beginning till the end, to measure their learnability or efficiency. The number of errors the participants were making was also taken under consideration, and their rating of the system they used to perform the task.

Task 3: The third task asked the participants to import 10 different songs from an existed local folder that contained a list of songs, to the application library. This task has so many different results with big gaps between them such that the first Four users tried to perform the task without any training. While, the last Four users got some training on how to perform the task.

Usability Metrics Results:

First User:

Background: The first participant has used iTunes before, but he is not familiar with using both Spotify, and Google Music.

Task # 1:

| FIRST USER | iTunes | Spotify | Google Music |
|---------------------|-------------------|--------------|------------------------------|
| Learnability (Time) | / | 3 Min | 5 Min |
| Efficiency (Time) | 5.30 Min.Sec | / | / |
| Error | Adding Song Twice | / | Adding to the wrong Playlist |
| Satisfaction | 8 out of 10 | 10 out of 10 | 6 out of 10 |

Task # 2:

| FIRST USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|-----------------------------|
| Learnability (Time) | / | 1.36 Min.Sec | 5 Min |
| Efficiency (Time) | 2.50 Min.Sec | / | / |
| Error | / | / | Searching the local library |
| Satisfaction | 9 out of 10 | 10 out of 10 | 7 out of 10 |

Task # 3:

| FIRST USER | iTunes | Spotify | Google Music |
|---------------------|--------------|-------------|------------------------------------|
| Learnability (Time) | / | 7 Min | 20 Min |
| Efficiency (Time) | 3 Min | / | / |
| Error | / | / | 15 Wrong Clicks 3 Drag and Drop |
| Satisfaction | 10 out of 10 | 6 out of 10 | 2 out of 10 |

Second User:

Background: The second participant is not familiar with any of the applications, as she only uses Pandora to listen to music on her phone.

Task # 1:

| SECOND USER | iTunes | Spotify | Google Music |
|---------------------|--------------------------------|--------------|----------------------------------|
| Learnability (Time) | 8 Min | 4.45 Min.Sec | 8 Min |
| Efficiency (Time) | / | / | / |
| Error | Wrong Click Naming Playlist | / | Wrong Click Adding wrong song |
| Satisfaction | 7 out of 10 | 9 out of 10 | 4 out of 10 |

Task # 2:

| SECOND USER | iTunes | Spotify | Google Music |
|---------------------|-------------------------------------|--------------|--------------------------------|
| Learnability (Time) | 6 Min | 4 Min | 5 Min |
| Efficiency (Time) | / | / | / |
| Error | Wrong Click Search local library | / | Searching the local library |
| Satisfaction | 6 out of 10 | 10 out of 10 | 5 out of 10 |

Task # 3:

| SECOND USER | iTunes | Spotify | Google Music |
|---------------------|--------------|-------------|------------------------------------|
| Learnability (Time) | 1 Min | 8 Min | 14 Min |
| Efficiency (Time) | / | / | / |
| Error | / | / | 10 Wrong Clicks 5 Drag and Drop |
| Satisfaction | 10 out of 10 | 6 out of 10 | 3 out of 10 |

Third User:

Background: The third person is not familiar with using any of the three applications, and here are the results of the experiment.

Task # 1:

| THIRD USER | iTunes | Spotify | Google Music |
|---------------------|-------------|---------------|------------------------------|
| Learnability (Time) | 13 Min | 10.49 Min.Sec | 14.12 Min.Sec |
| Efficiency (Time) | / | / | / |
| Error | Wrong Click | / | Wrong Click Drag and Drop |
| Satisfaction | 8 out of 10 | 7 out of 10 | 4 out of 10 |

Task # 2:

| THIRD USER | iTunes | Spotify | Google Music |
|---------------------|----------------------|--------------|-----------------------------------------------|
| Learnability (Time) | 5 Min | 3.11 Min.Sec | 14 Min |
| Efficiency (Time) | / | / | / |
| Error | Search local library | / | 16 Wrong Clicks Searching local library |
| Satisfaction | 9 out of 10 | 10 out of 10 | 3 out of 10 |

Task # 3:

| THIRD USER | iTunes | Spotify | Google Music |
|---------------------|--------------|----------------------------------------|----------------------------------------------|
| Learnability (Time) | 3 Min | 7.12 Min.Sec | 26 Min |
| Efficiency (Time) | / | / | / |
| Error | / | 3 Wrong Clicks Added iTunes Library | Multiple Wrong Clicks 13 Drag and Drop |
| Satisfaction | 10 out of 10 | 6 out of 10 | 1 out of 10 |

Fourth User:

Background: The fourth participant, is constantly using iTunes, and Spotify most of the time. But, he has never used Google Music.

Task # 1:

| FOURTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|--------------|
| Learnability (Time) | / | / | 3 Min |
| Efficiency (Time) | 3 Min | 2.30 Min.Sec | / |
| Error | / | / | / |
| Satisfaction | 10 out of 10 | 9 out of 10 | 7 out of 10 |

Task # 2:

| FOURTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|--------------|
| Learnability (Time) | / | / | 4 Min |
| Efficiency (Time) | 2 Min | 1.51 Min.Sec | / |
| Error | / | / | / |
| Satisfaction | 10 out of 10 | 10 out of 10 | 8 out of 10 |

Task # 3:

| FOURTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|------------------------------|
| Learnability (Time) | / | 3.32 Min.Sec | 10 Min |
| Efficiency (Time) | 3 Min | / | / |
| Error | / | / | Wrong Click Drag and Drop |
| Satisfaction | 10 out of 10 | 10 out of 10 | 6 out of 10 |

Fifth User:

Background: The fifth participant has used iTunes before and below is his outcomes.

Task # 1:

| FIFTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|-------------------------------|--------------|
| Learnability (Time) | / | 10 Min | 10 Min |
| Efficiency (Time) | 8 Min | / | / |
| Error | Wrong Click | Wrong Click Add Wrong Song | Wrong Click |
| Satisfaction | 9 out of 10 | 7 out of 10 | 8 out of 10 |

Task # 2:

| FIFTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|--------------|--------------|
| Learnability (Time) | / | 2.40 Min.Sec | 2.30 Min.Sec |
| Efficiency (Time) | 3 Min | / | / |
| Error | / | / | / |
| Satisfaction | 9 out of 10 | 7 out of 10 | 8 out of 10 |

Task # 3:

| FIFTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|-------------|--------------|
| Learnability (Time) | / | / | / |
| Efficiency (Time) | .8 Sec | .6 Sec | .33 Sec |
| Error | / | / | / |
| Satisfaction | 9 out of 10 | 7 out of 10 | 8 out of 10 |

Sixth User:

Background: The sixth participant had knowledge of all applications, and also used shortcuts.

Task # 1:

| SIXTH USER | iTunes | Spotify | Google Music |
|---------------------|------------------------|------------------------|--------------|
| Learnability (Time) | / | / | / |
| Efficiency (Time) | 5 Min | 4.47 Min.Sec | 3.35 Min.Sec |
| Error | Clicked Too Many Songs | Clicked Too Many Songs | / |
| Satisfaction | 6 out of 10 | 7 out of 10 | 10 out of 10 |

Task # 2:

| SIXTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|--------------|
| Learnability (Time) | / | / | / |
| Efficiency (Time) | 2.23 Min.Sec | 3.03 Min.Sec | 2.43 Min.Sec |
| Error | / | / | / |
| Satisfaction | 6 out of 10 | 7 out of 10 | 10 out of 10 |

Task # 3:

| SIXTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|-------------|--------------|
| Learnability (Time) | / | / | / |
| Efficiency (Time) | .6 Sec | .5 Sec | .30 Sec |
| Error | / | / | / |
| Satisfaction | 6 out of 10 | 7 out of 10 | 10 out of 10 |

Seventh User:

Background: The seventh participant has used iTunes before.

Task # 1:

| SEVENTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|--------------------------------------|--------------|
| Learnability (Time) | / | 9.30 Min.Sec | 11 Min |
| Efficiency (Time) | 8 Min | / | / |
| Error | Wrong Click | Wrong Click/Search Add Wrong Song | Wrong Click |
| Satisfaction | 9 out of 10 | 8 out of 10 | 7 out of 10 |

Task # 2:

| SEVENTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------|--------------|
| Learnability (Time) | / | 2.18 Min.Sec | 2.20 Min.Sec |
| Efficiency (Time) | 2.55 Min.Sec | / | / |
| Error | / | / | / |
| Satisfaction | 9 out of 10 | 8 out of 10 | 7 out of 10 |

Task # 3:

| SEVENTH USER | iTunes | Spotify | Google Music |
|---------------------|----------|----------|--------------|
| Learnability (Time) | / | .6 Sec | .33 Min.Sec |
| Efficiency (Time) | .8 Sec | / | / |
| Error | / | / | / |
| Satisfaction | 9 out 10 | 8 out 10 | 7 out 10 |

Eighth User:

Background: The seventh participant used iTunes before.

Task # 1:

| EIGHTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|--------------------------------------|---------------|
| Learnability (Time) | / | 10.33 Min.Sec | 11.37 Min.Sec |
| Efficiency (Time) | 9.05 Min.Sec | / | / |
| Error | Wrong Click | Wrong Click/Search Add Wrong Song | Wrong Click |
| Satisfaction | 9 out of 10 | 7 out of 10 | 7 out of 10 |

Task # 2:

| EIGHTH USER | iTunes | Spotify | Google Music |
|---------------------|--------------|-------------|--------------|
| Learnability (Time) | / | 3 Min | 2.14 Min.Sec |
| Efficiency (Time) | 3.12 Min.Sec | / | / |
| Error | / | / | / |
| Satisfaction | 9 out of 10 | 7 out of 10 | 7 out of 10 |

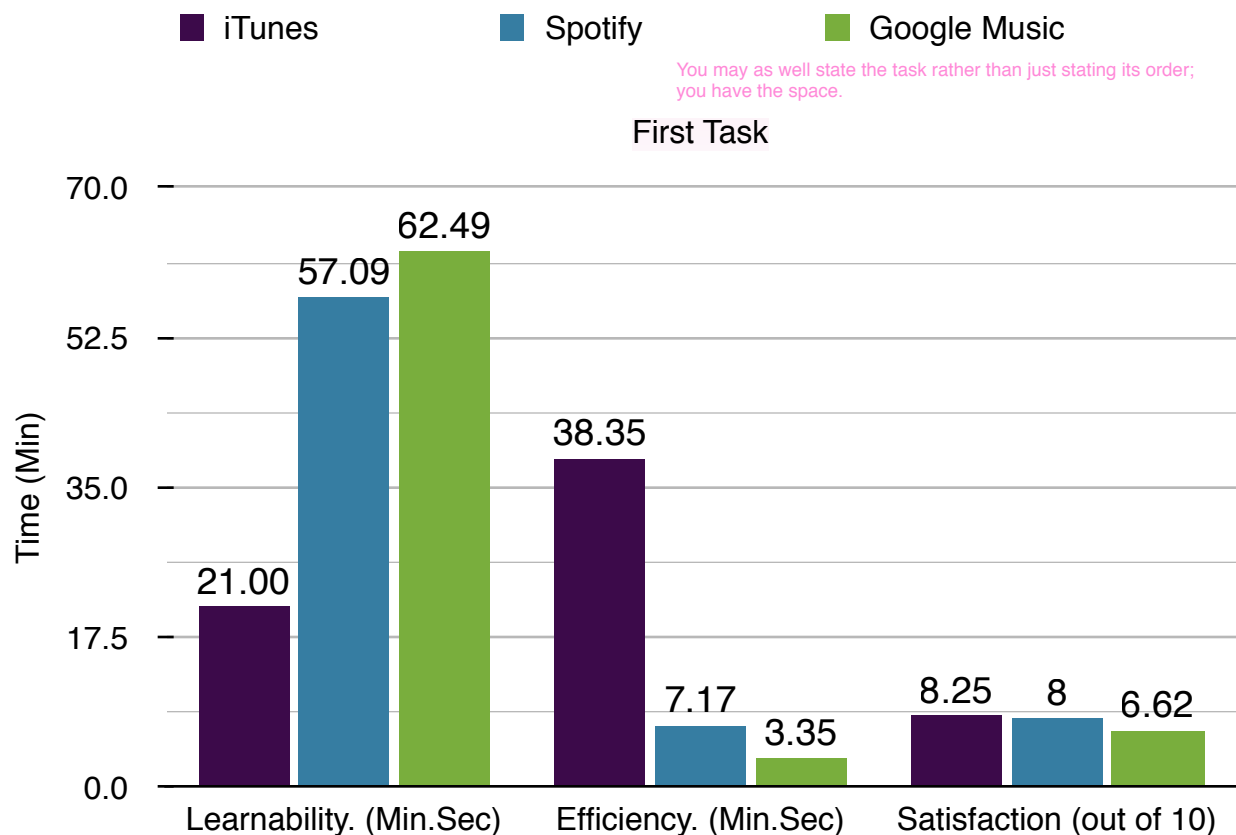
Task # 3:

| EIGHTH USER | iTunes | Spotify | Google Music |
|---------------------|-------------|-------------|--------------|
| Learnability (Time) | / | / | / |
| Efficiency (Time) | .10 Sec | .9 Sec | .33 Min.Sec |
| Error | / | / | / |
| Satisfaction | 9 out of 10 | 7 out of 10 | 7 out of 10 |

Heuristic Evaluation:

For this section, the evaluation will focus on demonstrating why each system performed the way it did. Moreover, the demonstration will focus on the mental model of the three systems from its developers' and users' perspectives, as possible. Also, it will be based on the results we collected from the experiment.

For everyone to clearly see the usability results, and before we discuss the mental model of the three systems, here is one of three charts that shows the learnability, and efficiency times for all the systems, and the average satisfaction rate. The chart shows the results of the first task.



This is the "sum" of the numbers, 13 + 8, and not the average. It gives the wrong impression, which is that it takes 21 minutes to figure out how to make playlists in iTunes. In reality, both users took significantly less time.

For the first task, we see that iTunes has the best learnability rate of 21 minutes, and that is because most of the users are familiar with the system. Therefore, they haven't been evaluated on it, instead they have been evaluated on their efficiency to complete the task. Meaning that the users are tested priorly on efficiency. However, iTunes users took them to accomplish the first task 21 (mins) in terms of learnability, and took them 38.35 (min.sec), based on the time of efficiency. On the other hand, Spotify users were tested priorly on learnability, and accomplished the task in a total time of 47.57 (min.sec) in terms of learnability, and 17.09 (min.sec) in terms of efficiency. Whereas, Google users' learnability, which was prioritized, was accomplished in times add up to 63.29 (min.sec), and the efficiency times add up to 3.35 (min.sec). Also, the users satisfaction rate on each system was as follows: For the first system iTunes rating was 8.25 out of 10, for Spotify 8 out of 10, and Google Music 6.62 out of 10.

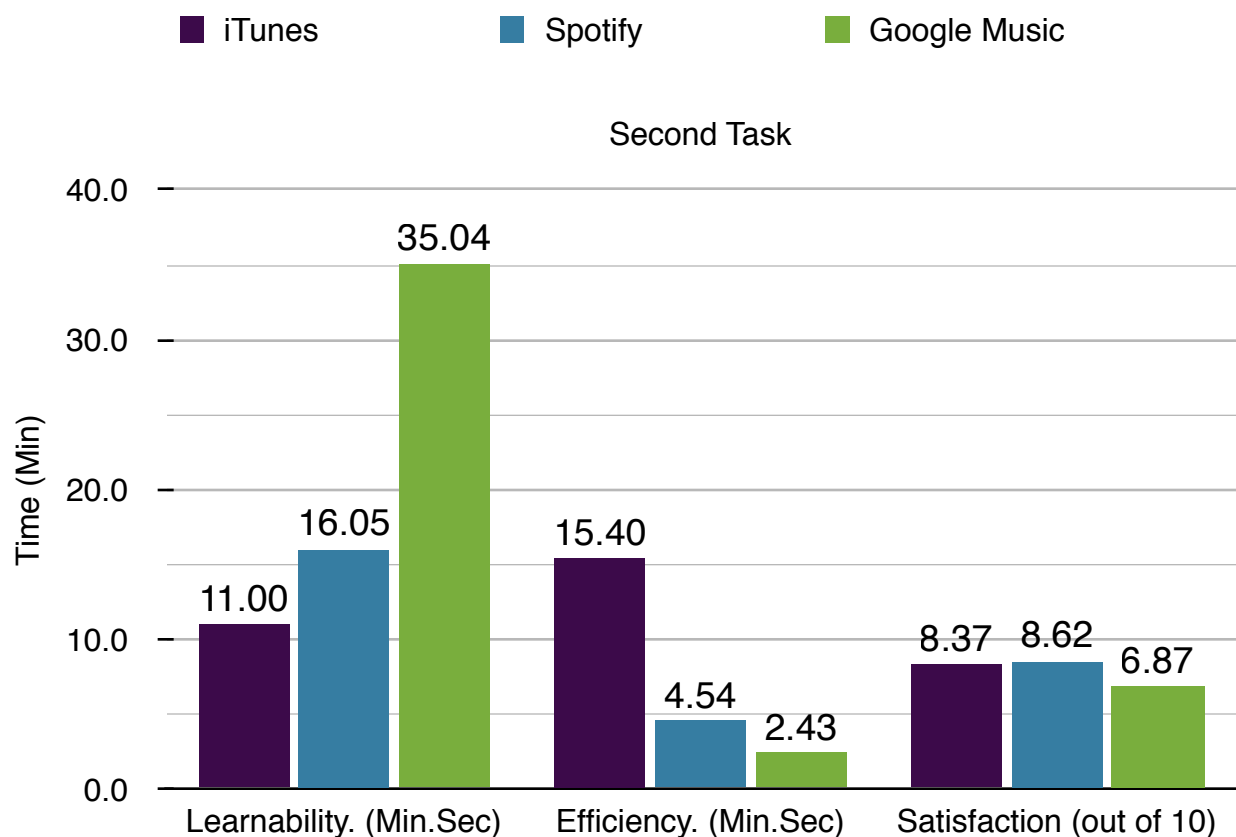
As the numbers show that iTunes was the best in accomplishing the task, than Spotify, and finally Google Music. One of the reasons was that iTunes, and Spotify developers, focused on putting the option "Add Playlist" in so many different places, unlike, Google Music developers which they made only one or two ways of doing it. For example, one of the ways that you can add a playlist using iTunes, and Spotify, is by clicking on the menu bar, and the first thing you see is "new playlist," which is most likely the first move to be performed by all kind of users. Also, if you just go and highlight on the songs that you need to add and right click on them, you will see from the options that will come out, "Add to a Playlist." Finally, if you just look at the sidebar where it contains all the playlists, you will find a Plus symbol that if you clicked on it, it will allow you to add a new playlist, which is the only possible way in Google Music.

For this first task, the users actually admired the way iTunes, and Spotify developers made it easy for them to add playlists. They said, the way iTunes, and Spotify included the option

to add a new playlist from the menu bar was very thoughtful. As for most of the people when trying to learn any new application, they will click on the menu bar first to discover the possible options.

Finally, we see that the developers' mental model for the design of iTunes, and Spotify was almost the same, as they put so many options, and ways for the same task, so the user can perform the task easily. Also, both of them thought that the user will first go to the menu bar and search there. Which, they took for consideration and made the first option there "Add a Playlist." Whereas, the Google Music developers missed that point which made the users commit errors when using it, and they made it time consuming to perform such a simple task.

The following is the chart for the second task, which shows the same elements as the first one.



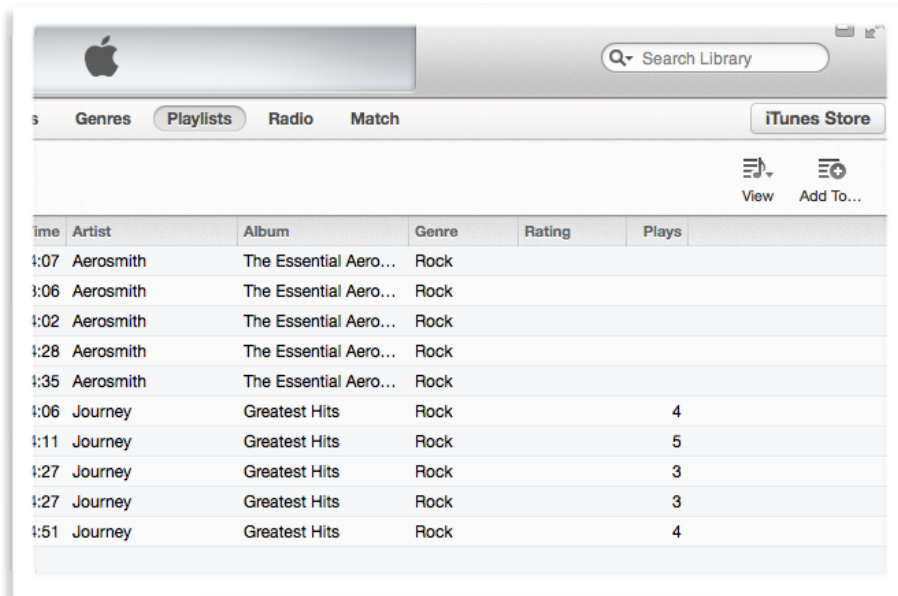
For the second task, we see that Spotify is on top this time when searching the online store for a song, followed by iTunes, and far behind is Google Music. The total time to accomplish the task when using Spotify was 16.05 (min.sec) in terms of learnability which was the prime element tested, and 11 minutes in terms of efficiency. When using iTunes, learnability was prioritized, due to the users familiarity with the system, and the total learnability time was 15.40 (min.sec). Whereas, the efficiencies was 11 (mins). Finally Google Music when tested by the users they accomplished the task in a learnability time of 35,04 (min.sec), and in efficiency time of 2.43 (min.sec). The users' satisfaction ratings when they performed the second task using the three application were, 8.62 for Spotify, 8.37 out of 10 for iTunes, and the least convenient system was Google Music with rating of 6.87 out of 10.

When performing the second task there were so many errors committed by the users when using iTunes, and Google Music. For example, most of the users at first searched the local library instead of the online store. Whereas, this problem wasn't there when they used Spotify. Because, when you use Spotify, the search space searches the online store as well as the local files. Furthermore, the reason behind these errors was the search space in iTunes, and Google Music. The user will automatically think that they will search the online store. But, it appeared that they only search the local, or the existed playlists. However, to search the online store using iTunes, you have to click on iTunes store button which is located just underneath the search space as the following image shows. However, it is a smart way from the developers to place the button there, as they obviously expected this kind of error to happen, even-though, they embedded a text saying, "Search Library."

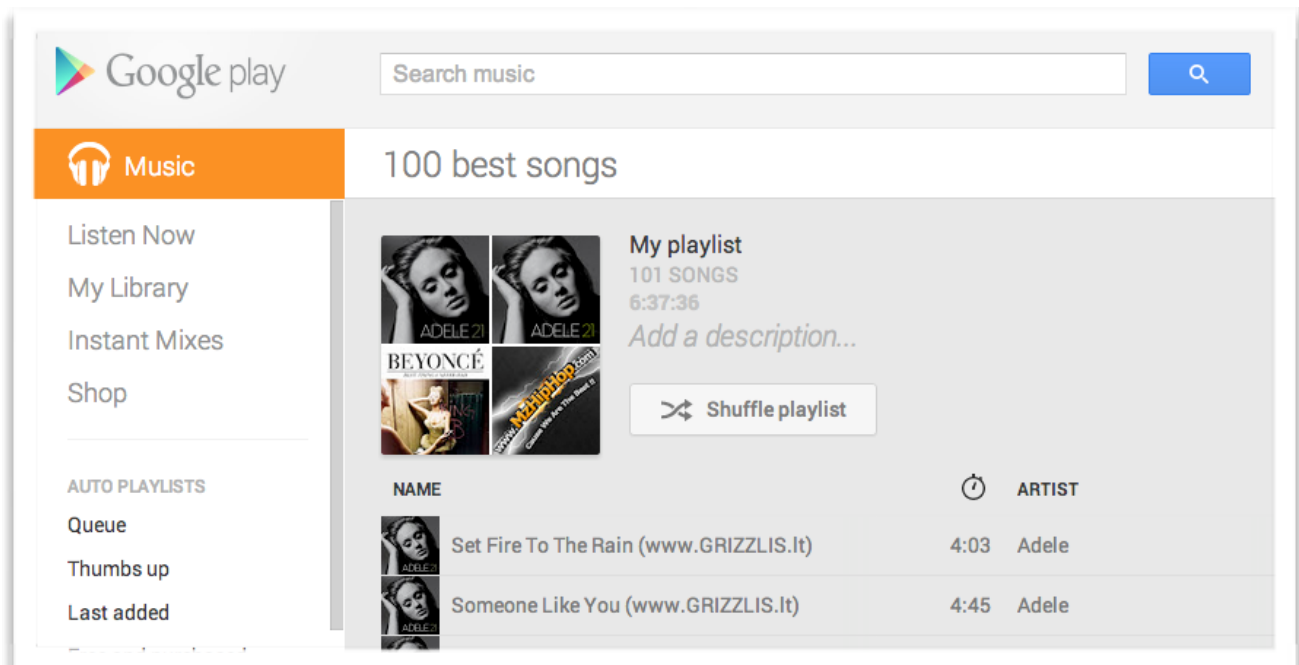
You have done this a few times now, saying that "for application x, some metric was prioritized." I don't understand that. Metric priorities don't change based on the specific application being tested.

I don't see anything in your chart with 11 minutes for efficiency.

OK, there is definitely confusion now. On your chart, 15.30 is for "efficiency" of iTunes, not learnability. And what is this with 11 minute "efficiencies" in the next sentence?

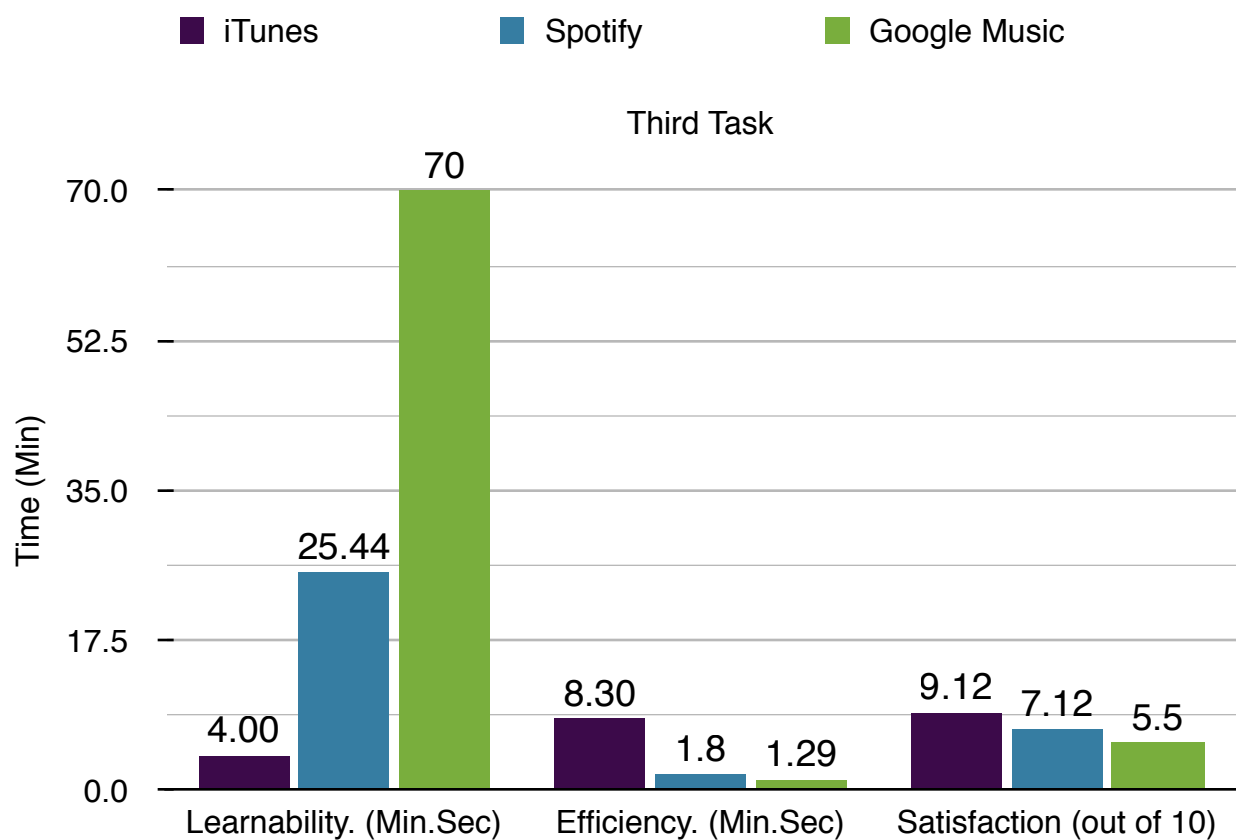


On the other hand, when you use Google Music you need to click on the option “Shop” from the sidebar, and not “Store” as in iTunes. So, this was a little bit confusing for the users. Also, what Google developers have missed, is that they didn’t embed anything within the search space to indicate, if you’re searching the local files, or the online store as in iTunes. Which, most of the users didn’t like, and found it difficult to do at first.



Finally, we see that Spotify developers' mental model differ from iTunes, and Google Music. As, they made their search space, searches for both the online store and the local files. Which, made it easier for the users to do, and look up their desired songs, with much less errors.

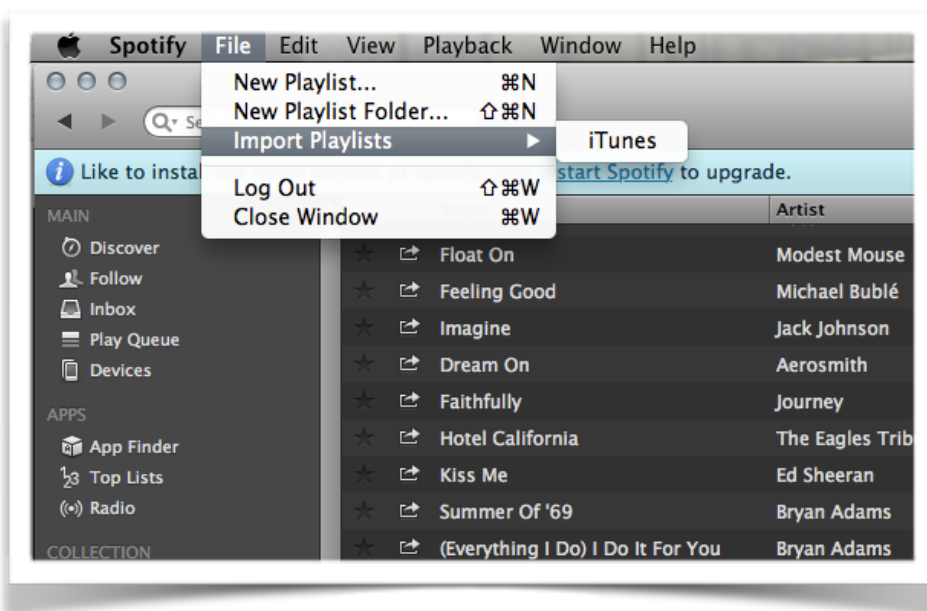
The following is the chart for the third task (Importing Music), which shows the same elements as the first, and the second one.

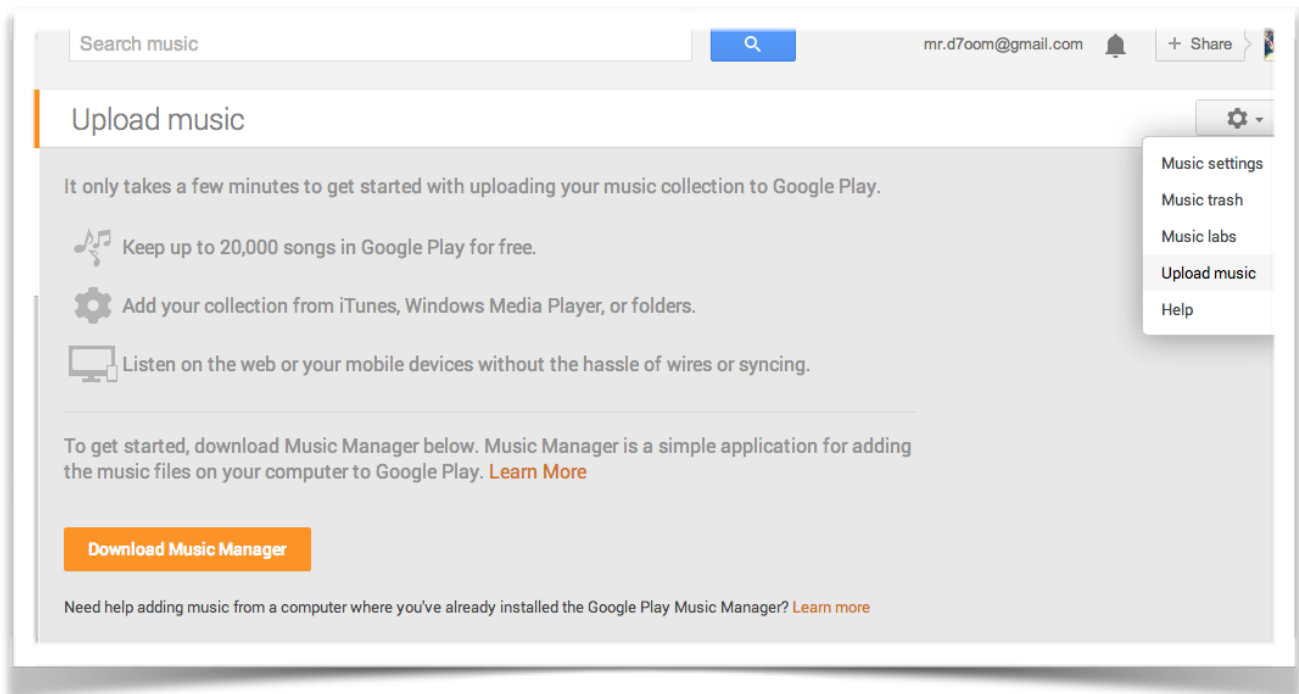
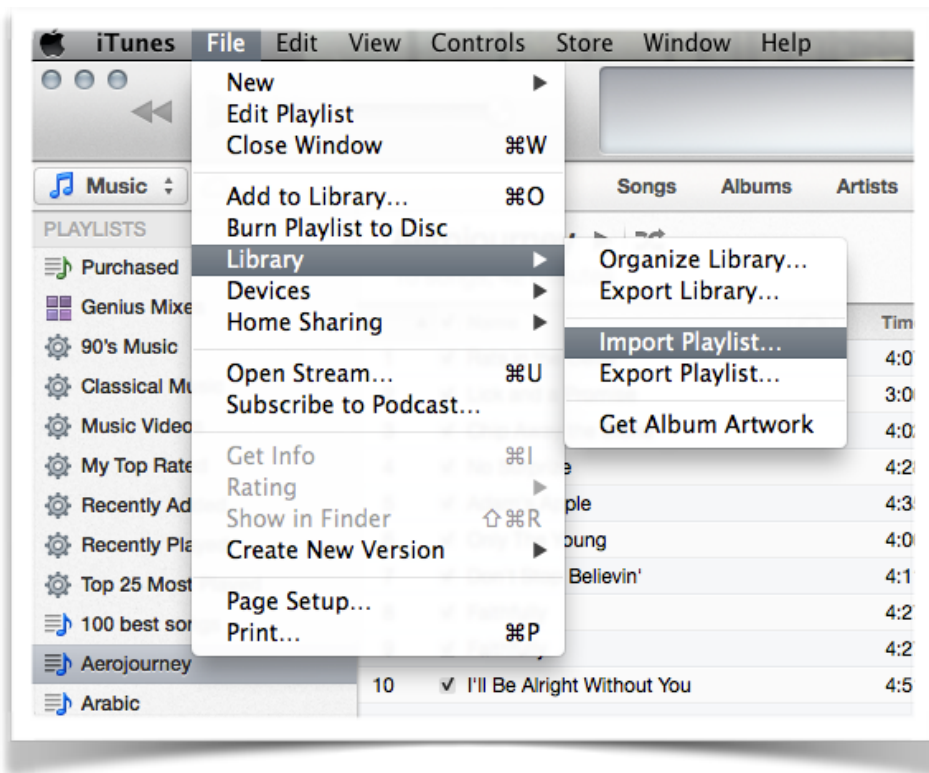


As this chart shows, we see that iTunes, is by far the best system to import music, in terms of learnability, or efficiency. Then, comes Spotify, and finally Google Music.

That's not what your numbers say. iTunes appears to be 4–8 times less efficient than Spotify or Google Music.

Again, iTunes developers mental model focused on putting most of the users' used options in the menu bar, for easy access and usage. Spotify used the same idea, but they only included import from the iTunes' library, but not from a local folder in their menu bar. Thus, In order to import songs to Spotify you need to drag and drop from the local folder to the application library. On the other hand, Google Music developers' mental model is totally different from iTunes, and Spotify in a way that, they use another application to import songs to the library, which first you need to download, then import the desired songs. This method was not expected by any of the users, which made the learnability taking a long time when performing a supposedly simple task. However, the next two images show how first Spotify and iTunes are similar, and the third one shows how Google asks you to download an application before you be able to do the task.





For the last part of the analysis we will focus on the relation between all these applications mental model designs, and the guidelines, principle, and theories of user interface design. The first guideline that all the application designers focused on implementing was Data Consistency. Whose guideline is this? And what does it mean? As it is obvious that all the designers made the top bar in showing data and the (drag down menus) that a user need to use consistent. Also, they kept consistent in labeling conventions. The second major implementation is that, all the applications designers used the principle of direct manipulation. Direct manipulation is an "interaction style", not a principle. For instance, all the applications used drag and drop to add songs to the playlists, which is a major form of direct manipulation.

However, let us turn now to some important user interface design principles, which they are the dialogue principles that represents the dynamic aspects of the interface, or let's say the feel. The following principles are the principles of "ISO 9241" which "is a multi-part standard from the International Organization for Standardization (ISO) covering ergonomics of human-computer interaction." You mention it, you cite it.

The first principle is the "suitability for the task," which as we noticed those tasks vary a lot, but the used interfaces were very suitable for the performed tasks. For example, the users' times measured when performing the three tasks were relatively short, and that reflects the effectiveness, and efficiency of those interfaces when performing the tasks.

Moreover, all the three interfaces used the principle of "suitability for individualization," which concerns with whether the interface is providing individual preferences or not. However, these three interfaces did. For example, they all allow the user to create his/her own playlist, and allow the user to have full control of them, such as, naming them, or adding certain music to them, and finally deleting them.

One final principle, which is “suitability for learning,” An interface that is said to be suitable for learning is when it shows and guides the user in a certain way in how to use and perform a certain task. For example, in the used interfaces, they all show how to add a song to the playlist by instructing the users by a written and obvious note saying “to add a song or songs use the mouse to drag and drop songs to the playlist.” Therefore, those interfaces were suitable for learning as they instruct and guide the user in every task needed to be performed.

In conclusion, as mentioned before, iTunes, and Spotify developers focused on making their mental model as easy as possible to perceive. Allowing the users to accomplish a certain task in different ways, and making their users find what they are looking for in one place most of the time. Also, the user’s feedback, and their mental model, and their perception on these two systems indicated the same as what the developers are trying to present, which made the learnability, efficiency, easier with fewer errors than in Google Music. On the other hand, Google Music developers’ mental model on what they are trying to achieve when they designed their system, did not meet with what the users are expecting it to be, therefore, it was the least satisfying system. Moreover, by implementing the interface design guidelines and principles, these interfaces managed to be efficient, effective, and satisfiable for most of the users, and therefore, they are the most suitable and famous user interfaces to perform those kind of tasks.

Mismatched mental models are not “solely” reflected in satisfaction.
“Effectiveness” is not a usability metric.
But making this blanket statement contradicts what you said about how Google Music was significantly mismatched in terms of communicating its developers’ mental model.