Learning in the GNU/Linux Community

Don Davis
University of Texas at San Antonio
College of Education and Human Development
Department of Interdisciplinary Learning and Teaching
210-458-5969
Don.Davis@utsa.edu

Iffat Jabeen
University of Texas at San Antonio
College of Education and Human Development
Department of Interdisciplinary Learning and Teaching
210-458-5969
Iffat.Jabeen@utsa.edu

ABSTRACT
The GNU/Linux operating system is one of most familiar examples of Free and Open Source Software (FOSS). The Free and Open Source Software paradigm encourages a sense of community and participation, building upon shared knowledge. The same tenets of collaboration and shared knowledge are espoused by academia and the educational system [8]. However, there is little discussion of the GNU/Linux community and the FOSS framework in much of the K-16 education system.

If identified as a participatory learning community, the GNU/Linux community and its utilization of the FOSS framework would represent a stark contrast to the proprietary, black box software and software practices currently adopted in much of K-16 education. Therefore, researchers conducted an online survey to determine if the GNU/Linux community exhibited significant elements of the legitimate peripheral participation (LPP) learning process. This research does not focus on specific units of learning using GNU/Linux but rather is a more general investigation of a possible GNU/Linux community to determine whether more specific research evaluating the educational efficacy of FOSS and the FOSS framework is warranted.

This paper presents an analysis of survey data provided by 4603 respondents reflecting on learning within the context of GNU/Linux and other FOSS communities. The intent of the paper is to facilitate an evaluation of GNU/Linux and its properties as a participatory learning community by educational technologists, educators, education administrators, and other education stakeholders. In this paper, researchers analyze data to determine whether the GNU/Linux community exhibits significant characteristics of a community engaged in the process of legitimate peripheral participation. Researchers then identify some of the more significant correlations identified among participatory activities within the GNU/Linux community and learning. Last, researchers present community members' perceptions of learning within the GNU/Linux community and their beliefs regarding the possible role(s) of GNU/Linux and the FOSS framework within the K-16 educational paradigm.

Categories and Subject Descriptors
K.3.2 [Computers and Education]: Computer and Information Science Education; K.3.0 [Computers and Education]: General; K.3.1 [Computers and Education]: Collaborative Learning; K.3.m [Computers and Education]: FOSS, GNU/Linux

General Terms
Human Factors

Keywords
GNU/Linux, FOSS, Legitimate Peripheral Participation

1. INTRODUCTION

1.1. Evolution of GNU/Linux
The GNU/Linux operating system sprouted from the Linux kernel developed by Linus Torvalds and GNU software packages developed by Richard Stallman and others at the GNU project. GNU/Linux quickly blossomed and has become a successful OS with a community of more than five million users and developers. Since Linus Torvalds first announced his plans for an operating system, Linux has relied on a community of users and developers sharing and co-creating knowledge [1][4][7][8][13]. This presents a stark contrast to current software paradigms taught and perpetuated in K16 education environments.

The existing trend of using proprietary software hinders the promotion of flexibility in education. The rights to modify or copy code stay exclusively with software developers and publishers leading to licensing issues and often at exorbitant costs.

1.2. Shared Goals
The ethos and philosophy of the FOSS community (and in particular the GNU/Linux community) appear to be closely aligned to the professed beliefs and goals of the educational community; this seems especially true in regards to shared knowledge, collaboration, and a co-creation of knowledge [1][4][7][8].

1.3. Theoretical Framework
In this paper, researchers examine learning within the context of the GNU/Linux community through the lens of legitimate peripheral participation as initially described by Lave and Wenger [9]. Legitimate Peripheral Participation (LPP) can be understood as a process whereby novices entering a field learn and hone their skills with interaction and guidance from the experts and seasoned members in that area. As defined by Lave and Wenger [9], “a person's intentions to learn and the meaning of learning is configured through... becoming a full participant in a sociocultural practice” (p.29) and therefore learning occurs as a function of identification with the group. The LPP construct emphasizes that a full participant is “able to be involved in new activities, to perform new tasks, [and] to master new understandings” (p. 53) within the context of a social community.

As Fang and Neufeld[4] observe, “[a]s new members become increasingly established in the community, their access to resources and understanding escalates and this in turn stimulates them to participate even more actively in the future” (p. 14).

Given a cursory glance, practices within the GNU/Linux community seem to exemplify Lave and Wenger's description of LPP. It then behooves researchers to establish whether learning and participation within the GNU/Linux community fulfill the requirements of LPP. For that purpose, researchers evaluated the GNU/Linux community in terms of the constituent components of legitimate peripheral participation, although Lave and Wenger (p. 35) emphasize the composite nature of LPP.
A community exhibiting characteristics of LPP should possess characteristics of legitimacy, peripherality, and participation.

1.3.1. Legitimacy
In contrast to proprietary software environments, GNU/Linux community members appear to be legitimate participants. Users can, and frequently do, communicate directly with developers and package maintainers via email and public listservs. Bug reports and features requests are rapidly acknowledged and incorporated into software packages. Although some bugs and features take longer to incorporate, their status is publicly viewable [10]. Users often create and share patches (tiny changes in source code) with one another and developers. The developers use publicly readable open mailing lists and IRC channels to discuss development. The programming tools and source code are freely available to all. Learners are encouraged to tinker and learn as bricoleurs[11]. It is the responsiveness to users and acceptance of their input that legitimizes the participation of GNU/Linux community members. As Lave and Wenger [9] indicate, legitimacy is a “crucial condition for learning” (p. 35).

1.3.2. Peripherality and Participation
“Peripherality,” as described by Lave and Wenger, “suggests that there are multiple, varied, more- or less-engaged and -inclusive ways of being located in the fields of participation defined by a community” (pp. 35 – 36). The GNU/Linux community strongly exhibits ‘peripherality’ in its numerous roles in which users and developers can participate. Users are encouraged to ask questions, help others, participate in Linux User Groups (LUG) and forums, submit bug reports and feature requests, contribute code, and provide translations. Lave and Wenger expound that “[a]… one moves toward more-intensive participation, peripherality is an empowering position” (p. 36). Viewing GNU/Linux through the filter of legitimate peripheral participation, a GNU/Linux community member might be expected to use the computer for typical desktop computing and then progress in the community through increasing contributions such as filing bug reports, asking and answering questions, debugging, providing documentation, contributing patches and workarounds, and possibly becoming involved in the further development of GNU/Linux software and the GNU/Linux community [7][9].

1.3.3. LPP Traits of the GNU/Linux Community
To determine the extent that the GNU/Linux community exhibits LPP traits, researchers examined correlations among participation, identification, and learning within the GNU/Linux community. Lave and Wenger emphasize that full participants “do not exist in isolation” (p. 53); it then follows, that in terms of LPP, full participants are those who participate most frequently and should consequently experience the greatest amount of learning. Therefore, researchers determined whether there is a significant relationship between GNU/Linux community members’ participation and their learning.

The sociocultural aspect of LPP, where identification with a group is a prerequisite, consequently lead the researchers to investigate whether there are correlations between members’ identification with the group and learning, and members’ identification with the group and participation. In this study, researchers do not seek to establish whether the GNU/Linux community is best described as community of interest or a community of practice. Indeed, within the GNU/Linux community there may be substantial overlap. Therefore, in order to avoid semantic analysis and discussion which lies outside the scope of this paper, researchers utilize the term participatory learning community.

2. METHODOLOGY
This study was conducted using an online survey of 21 questions with multiple components. An online survey was utilized in order to facilitate greater participation from GNU/Linux users worldwide. The survey was designed to understand the GNU/Linux community; therefore, non GNU/Linux users were not directly targeted. Posts soliciting volunteers for participants in the online survey were posted to online message forums dedicated specifically to GNU/Linux such as Linuxquestions.org, Ubuntu, Red Hat, OpenSuse, and Debian user groups. Participants were also invited to this study through other means: LUG mailing lists, email, Facebook, and blog posts. The application MrInterview was utilized to create the survey; data was analyzed using SPSS and R.

The survey, primarily consisting of Likert scale type questions, was made available online for twelve weeks. It was completed by 4603 participants. The largest number (n > 3000) of respondents completed the survey after Ken Starks, well known for his Linux based charity work, posted the survey link on his blog [12]. Respondents were predominately male ranging in age from under 18 to over 45 years. Respondents consisted of 40% advanced, 24% developing, 34% beginning, and 2% unspecified levels of end-user ability.

3. RESULTS
Welch’s t-tests and Pearson's product moment correlation tests were conducted to compare responses in relation to members’ participation, learning, and identification with the GNU/Linux community. The data provided confidence levels of 95% and p < .001. Regression analysis was the primary tool used to determine the validity of relationships and identify significant regressors for further evaluation; correlation coefficients (r-values) and coefficients of determination (adjusted r² values) are provided.

3.1. General Findings
3.1.1. Participation and Learning
The data indicate a significant correlation (r = 0.473, r² = 0.224, p < .001) between overall participation, the sum of participation values for all activities, and overall learning, the sum of values for learning related to programming, teamwork, and general computer knowledge. The data display a significant correlation (r = 0.432, r² = 0.186) between software participation and overall learning. Additionally, a significant correlation (r = 0.426, r² = 0.181, p < .001) was identified between communication participation and learning.
Given the large sample size the authors have applied the guidelines for the social sciences presented by Cohen where for small effect size, $r = 0.1 - 0.23$; medium, $r = 0.24 - 0.36$; large, $r = 0.37$ or larger [2][3].

Figure 5. Correlations among identification, participation and learning

3.1.2. Identification, Learning, and Participation

The data exhibited a moderate correlation ($r = 0.373$, $r^2 = 0.139$, $p < .001$) between identification with the GNU/Linux community and respondents' learning. Additionally, in order to determine whether the GNU/Linux community exhibited prerequisite characteristics of a community engaged in the process of LPP whereby identification is a function of participation, data was analyzed to determine the relationship between participation (software participation, software contributions, communication participation) and identification with the GNU/Linux community. Analysis of the data revealed a significant relationship ($r = .324, r^2 = .105, p<.001$) between software participation (including frequency values for filing bug reports, compiling programs from source code, changing programs source code, and providing software contributions) and identification with the community. Additional analysis indicated a lesser (though significant) relationship ($r = 0.311, r^2=0.097, p<.001$) between software contributions (writing patches, work-arounds, original code, and maintaining packages) and identification with the community. Further, an evaluation of the data indicated a stronger, though still in the medium range for Pearson's $r$, correlation ($r = 0.412, r^2 = 0.169, p<.001$) between communication participation (asking and answering GNU/Linux related questions) and identification values.

3.2. Specific Relationships

Significant relationships were identified within the GNU/Linux community embodying traits of LPP which serve to bolster the perception of the GNU/Linux community as a participatory learning community. Having presented correlations establishing the GNU/Linux community as a participatory learning community, more specific correlations will now be presented among more specific participatory acts and learning within the GNU/Linux community.

3.2.1. Software Participation and Learning

Within the LPP framework, Lave and Wenger [9] considered transparency of technology to be an essential prerequisite for learning. Similarly, access to source code (transparency of technology) is considered essential within FOSS communities including the GNU/Linux community [4][7][13]. Therefore, correlations related to source code (transparency of technology) and learning were examined. In considering these relationships, it is important to note that although FOSS communities (and the GNU/Linux community in particular) consider access to source code essential, many community members have no interaction with source code. From 4603 respondents, 88% (n=4034) report having ever compiled programs from source, 78% (n=3604) report looking at source code, and 57% (n=2637) report ever changing source code.

Figure 6. Frequency of compiling packages from source

Figure 7. Frequency of viewing source code

Figure 8. Frequency of changing source code

Although few respondents frequently interact with source code, strong correlations were identified between interacting with source code and learning. Compiling from source code displayed a significant correlation to perceived levels of programming learning. An analysis of the data indicated one of the strongest correlations identified ($r=0.514, r^2=0.264, p<.001$) between looking at source code and programming related learning. Interestingly, there was a slightly lower correlation between changing source code and learning programming ($r = 0.459, r^2=0.210, p<.001$). (Bear in mind that 2637 respondents (57%)
indicated ever changing source code.) Additionally, analysis of the data revealed a medium correlation \((r = 0.403, r^2=0.162, p<.001)\) between software contributions and overall learning though very few respondents \((n=2569, 56\%)\) indicated making software contributions.

**Figure 9. Correlations among identification, learning, and participation**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Overall Learning</th>
<th>Learning Programmability</th>
<th>Corresponding correlation values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compile Source Code</td>
<td>0.34</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>View Source Code</td>
<td>0.40</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Change Source Code</td>
<td>0.34</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>File Bug Reports</td>
<td>0.33</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Ask Questions</td>
<td>0.33</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Provide Answers</td>
<td>0.37</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Overall Participation</td>
<td>0.47</td>
<td>0.48</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2. Communication and Learning

Analysis of the data indicated that the GNU/Linux community may constitute a participatory learning community engaged in the process of LPP and as such may exemplify a sociocultural practice. Within a sociocultural practice, communication plays an important role. Therefore, researchers analyzed data pertaining to communicative aspects of the GNU/Linux community and learning, such as the frequency and relationships among communication practices and learning. In the following section, data is provided detailing quantitative summaries of communicative practices within the GNU/Linux community. These data are followed by values of correlation.

The following graph (Figure 11) illustrates respondents’ reported frequency of electronically asking questions related to GNU/Linux or other FOSS projects. This is followed by a table which illustrates respondents’ indicated frequency of electronically providing instructionally beneficial answers.

**Figure 11. Density of learning by frequency of asking Linux/FOSS related questions**

A review of the data indicated a lesser correlation \((r = 0.325, r^2=0.106, p<.001)\) between asking questions related to the GNU/Linux community and overall learning. Further analysis indicated a stronger correlation \((r = 0.373, r^2=0.139, p<.001)\) between providing instructionally beneficial answers and overall learning. Data analysis also indicated a mild correlation \((r = 0.328, r^2=0.108, p<.001)\) between filing bug reports and overall learning.

4. IMPLICATIONS

4.1. The FOSS Framework in education

In this study, researchers sought to determine whether the GNU/Linux community exhibits significant characteristics of a community engaged in LPP. Upon establishing the validity of previous assertion, researchers examined the more specific characteristics of the GNU/Linux community in relation to learning. Now, researchers will present GNU/Linux community members’ perceptions of the possibilities for GNU/Linux and the FOSS in K-16 educational environments.

Respondents’ responses indicate the presence of characteristics within the GNU/Linux community which may facilitate the inclusion of new learners and promote a culture of learning. The majority of respondents (61%) indicated having very positive experiences with the GNU/Linux community. Respondents also
indicated that GNU/Linux community members frequently guide lesser experienced members and that community members prefer to guide rather than 'simply provide answers.' These attributes indicate that the GNU/Linux community environment may be exceptionally conducive to learning.

Figure 14. Frequency of guiding less knowledgeable members in GNU/Linux

Respondents perceived benefits of FOSS inclusion in schools. The majority of survey respondents felt that the use of FOSS is educationally beneficial and that the use of FOSS was better suited to education than the proprietary model.

Figure 16. Recommend use of FOSS in schools

The majority of respondents recommended teaching the FOSS framework to students in K-16. Additionally, it is important to note, not only do respondents recommend involving programming students in FOSS projects, the majority of respondents recommend involving non-programming students in FOSS projects. (Italics added for emphasis.)

Figure 17. Preference for FOSS to proprietary for education

Figure 18. Recommend teaching FOSS framework to K-12 students

Figure 19. Recommend teaching FOSS framework to college students
students engaged in the sociocultural practices of the GNU/Linux community. Further research should seek to identify how the learning of non-programming (non-CS) students as well as IT and non-IT students. Therefore, further research is warranted eliciting educators' perceptions regarding the educational efficacy and feasibility of using FOSS and implementing the FOSS framework in classrooms and the educators' perceptions of possible impact on students’ learning. It may also be of interest to recognize educators who are currently utilizing this framework and their experiences regarding its benefits or deficiencies.

Almost all survey participants belonged to the GNU/Linux community which might be perceived as a potential bias to data. However, in order to establish whether the GNU/Linux community is engaged in LPP it was important to purposefully sample compares to that of students familiar only with proprietary software and proprietary software practices. In doing so, learning should be separately evaluated for programming and non-programming students as well as IT and non-IT students.

Additionally, there was a significant gender discrepancy among survey respondents; Ninety-six percent (96%, n=4415) of respondents were male. Future research should seek to identify factors affecting female participation in the GNU/Linux community.

6. CONCLUSIONS
The GNU/Linux community exhibits significant characteristics of LPP; therefore, the GNU/Linux community may represent an asset for lifelong learners as participatory learning community. It may behooves education stakeholders, including IT educators and K-16 education stakeholders not involved with computer science education, to further investigate learning as a function of the sociocultural practices associated with the GNU/Linux community and other FOSS communities, including communities easily accessed by non-programmers such as the LibriVox community. Given the ever-present constraints of the digital divide and the fiscal limitations of proprietary software, it behooves educational stakeholders to closer examine the possibilities and impact of involvement with FOSS for all learners.

ACKNOWLEDGMENTS
The authors give a special thanks to Ken Starks for posting the survey link on his blog which generated a large number of respondents and to Raymond Holland for his assistance with statistical analyses.

REFERENCES