Motivation of Software Developers in Open Source Projects: An Internet-based Survey of Contributors to the Linux Kernel

Guido Hertel Sven Niedner Stefanie Herrmann

University of Kiel

In press: **Research Policy, special issue on Open Source Software Development.**

Running Head: MOTIVATION IN OPEN SOURCE SOFTWARE PROJECTS

Key Words: Open Source Software, Linux, Virtual Teams, Motivation, VIST Model

Author notes *

Correspondence should be addressed to Guido Hertel, University of Kiel, Institut fuer Psychologie, Olshausenstr. 40, D-24 098 Kiel, Germany. Phone: +49 431 880 2980. Fax: +49 431 880 1559. E-mail: hertel@psychologie.uni-kiel.de.

Abstract: The motives of 141 contributors to a large Open Source Software project (the Linux kernel) was explored with an internet-based questionnaire study. Measured factors were both derived from discussions within the Linux community as well as from models from social sciences. Participants' engagement was particularly determined by their identification as a Linux developer, by pragmatic motives to improve own software, and by their tolerance of time investments. Moreover, some of the software development was accomplished by teams. Activities in these teams were particularly determined by participants' evaluation of the team goals as well as by their perceived indispensability and self-efficacy.

Motivation of Software Developers in Open Source Projects:

An Internet-based Survey of Contributors to the Linux Kernel

The idea of Open Source Software (OSS) has gained more and more attention in the last years. OSS is usually developed by a loosely-knit community of programmers spread all over the world, who contribute to a software project via internet without necessarily being employed or paid by an institution. In most cases, such an institution does not even exist. The objective of the present study was to explore the motivation of such persons who spend considerable time and effort in OSS projects "for free". The data were collected within the Linux kernel project, one of the most active OSS projects in the world that has a relatively long history and a high popularity (see Moon & Sproull, 2002, for a recent documentation). In addition to scientific interests, a better understanding of motivational processes within the Linux kernel project might help to improve software development processes in other OSS projects as well as in the corporate world.

In the following sections, we will first describe briefly the main principles of OSS development and discus suggestions about motivational processes that originate from the OSS community itself. Then, we will present two theoretical models from social sciences that might provide a conceptual framework for the explanation of persons' involvement in OSS projects. The first has its roots in research on voluntary action for social movements and related community work (Klandermans, 1997; Simon et al., 1998) and specifies different classes of motives. If we understand the Linux community as a social movement with certain social and political interests, this model might be suitable to explain the motivation of its members. The second model specifies motivational processes of individuals when they work in small teams (Hertel, 2002; Karau & Williams, 2000; Karau, Markus, & Williams, 2000). Since some of the Linux kernel development might be conducted in project teams, team-based motives might also play an important role in order to understand developers' activities in OSS projects. A web-based questionnaire was developed based on these two conceptual models and completed by members of the Linux kernel community. After describing this questionnaire and its results, we will discus the implications of the results for our understanding of the motivational processes in OSS projects, as well as for the possible improvement of such projects.

Open Source Software

In Open Source Software (OSS) development, the source code (i.e., the human-readable commands) of a computer program is publicly available and usually shared via the internet. Every internet user with sufficient skills can join the project at any time, for instance by downloading the source code and implementing extensions or corrections. Usually, software developers contribute to OSS projects "for free", either as a hobby or during their regular working hours even when OSS development is not part of their usual work. However, some companies have recently started to sponsor OSS development, and pay developers so that they can work on OSS projects full time. Although this changes the voluntary nature of OSS development, it does not affect the general OSS principle that the source code is available to everybody.

The idea of OSS development is about as old as software development itself. In fact, when researchers in the 1960s began to use computers for their work, they often had to rely on open sharing of software code simply because commercial software solutions and support were not available at that time (Moon & Sproull, 2002). Later, when commercial software development increased and often only the software vendor had access to the source code of a program, OSS became an attractive alternative since it enabled the users to adapt and improve the software according to their personal needs. In the mid 1980s, R.M. Stallman, one of the first public advocates of the OSS idea, also introduced a political aspect to OSS, claiming that "information wants to be free", and computer programs should not be owned by companies but be a public good (Stallman, 1994 a, b). Following Stallman’s and others’ initiative, the Free Software Foundation created the "General Public License" (GPL), a legal paper assuring that once software was published under that license, the availability of the source code is guaranteed for all future enhancements. The Linux kernel is one example of software licensed under the GPL.

The label "Open Source Software", however, was coined to create a tamed variant of the

Free Software Foundation's ideological concept in order to enhance the acceptance of OSS by software companies. According to Raymond (1999), it suffices that the source code is publicly available and that it permits changes by users. As a consequence, a number of companies such as Netscape (including the Mozilla project) or Sun Microsystems (including the Solaris operating system and the StarOffice package) decided to publish the source code of parts of their software projects. Other examples of successful OSS products are the Apache webserver, the Free BSD operating system, and the Mozilla web browser. However, none of them has yet reached the popularity of the Linux operating system.

The Linux Kernel Development Process

Linux is a PC-based operating system that has been developed as Open Source Software along the structure of the UNIX operating system. The Linux system consists of a large number of programs or modules that are arranged around a kernel. The task of the kernel is to give the programs access to resources such as hard disk storage, random access memory, network band width, etc. The central role of the kernel makes it an essential part of the Linux operating system, currently containing about two million lines of source code (Bovet & Cesati, 2001).

The development of the Linux operating system was started in 1991 by the Finnish computer science student Linus Torvalds, who is still managing the project. In the very beginning, Torvalds did not plan to start a worldwide OSS project. In his first announcement of Linux he called the project a "pet project ", and his original motivation can be described as a mixture of dissatisfaction with existing software solutions and enjoyment to write software programs (Torvalds & Diamond, 2001). However, over the years he and the Linux project have attracted several thousand developers and other contributors from all over the world (see Moon & Sproull, 2002, for more details). The "credits" file of the Linux kernel alone contains already about 400 contributors, and about 3.500 people were subscribed to the Linux kernel mailing list at the time of this study (spring 2000). As a result of these activities, Linux products are today widely regarded as being high in quality and reliability, and compete successfully with software products that are developed by large commercial companies.

Although a formal organizational structure is lacking, there are some general characteristics of the Linux project that might be considered as structural conditions of successful OSS development (cf. Moon & Sproull, 2002). Among them are (a) a general culture in which authority comes from competence, (b) delegative and participative leadership principles combined with clear responsibilities, (c) a modular project structure that decreases unnecessary complexity, (d) a parallel release policy that simultaneously enables rapid development and a stable working system, (e) a motivating credit policy that not only acknowledges the contributions of developers but also, for instance, documentation work, (f) clear rules and norms of the community that are communicated online, and (g) simple but reliable communication tools that are available worldwide (e-mail, file transfer, Usenet discussion groups). Interestingly, most of these principles are similarly considered as structural pre-conditions for successful distributed work in business organizations (e.g., Duarte & Snyder, 1999; Lipnack & Stamps, 1997).

The activities of the Linux kernel programmers are mainly organized by the Linux kernel mailing list. This mailing list acts as a central place for discussions about the technical and organizational aspects of the kernel development. Since subscription to the mailing list is open to everyone, joining and leaving the project is easy. There is no formal criterion for being a Linux developer apart from the contributions (patches) one submits. However, in all questions, Linus Torvalds has still the final decision as the project maintainer, and he is often referred to as a "benevolent dictator". Besides the Linux kernel mailing list, there are a couple of mailing lists related to more specific technical aspects of the kernel development, such as the development of certain file system or hardware drivers. These fields are called "subsystems" of the kernel. We believe that most of the development takes place in these subsystems, and therefore we will focus on these subsystems particularly when we explore possible teamwork aspects of the Linux development process. However, it should be noted that these subsystems are technical and not organizational entities, and hence are not necessarily teams but can also contain only one developer.

We derived the possible motives of Linux kernel contributors from two different sources: First, we analyzed discussion papers that circulate within the OSS community itself, thereby building on the intuition and experience of the persons involved. Second, as a more systematic and structural approach, we consulted theoretical models developed from social scientists who try to explain and predict more generally when and why individuals voluntarily engage in community projects. Below, we describe these sources in more detail.

Suggestions from the OSS Community

While Stallmann’s publications (e.g., 1994 a, b) on OSS might be seen as rather ideological, first thoughts about motivational processes in OSS projects can be found in publications by Raymond (1999). According to Raymond, an OSS project is usually born by "scratching a developer's personal itch" (Raymond, 1999, p. 32). He compares the development process to a bazaar, where everyone can join and contribute, creating an inspiring, creative and democratic atmosphere. Raymond contrasted this bazaar style development with a rather hierarchical "cathedral style" which he observed in commercial software development. In the bazaar model, the democratic discussion of changes in the software is assumed to assure that only the best solutions are accepted for the source code. On the other hand, it is also assumed that public scrutiny avoids flaws (bugs) in the program more effectively since every user is also a potential developer. The quote "given enough eyeballs, all bugs [programming errors] are shallow" (Raymond, 1999, p. 41) is famous among OSS developers and captures this idea.

However, Raymond's postulation of how OSS development might work has also been criticized. For instance, Bezroukov (1999 a, b) claimed that OSS development is not as outstanding as Raymond stated, but resembles structures and processes of the scientific community. According to Bezroukov, members of both the scientific and the OSS community are not driven by monetary rewards but by competitive motives of status and reputation (Bezroukov 1999 a). Interestingly, Bezroukov also suggested that some of the motivation to work in OSS projects might stem from a perceived competition of OSS projects with commercial software companies. Bezroukov further argued that these competitive motives can also decrease the quality of decisions, and he described burnout syndromes and personal attacks ("flame wars") among developers as symptoms of such unproductive processes.

Linus Torvalds himself stated that he initially published the source code of the first Linux kernel version because it was just “natural within the community" (Torvalds 1998), thereby referring to a (virtual) community that already existed before the development of Linux was started, and that was based on internet mediated discussions on Usenet. One of the main personal motives for Torvalds was simply “fun to program” (Torvalds & Diamond, 2001), and he assumed that this also applies to a large number of his co-developers. However, Torvalds also admitted that the current success of the Linux project is connected with reputation that might provide career advantages of software developers, although this was not a goal at the beginning of the project.

In summary, the discussion within the Linux kernel community suggests two main motive classes of OSS developers: (a) intrinsic motivation ("fun to program") and personal challenges to improve existing software for own needs, and (b) social comparison motives such as competition with other developers (either within OSS projects or between OSS projects and commercial software projects) and/or the interest to build a reputation that might be helpful for their occupational career (see also Hars & Ou, 2002).

When the present study was conceptualized (January 2000), no empirical data were available that explored these speculations empirically beyond mere demographic information about OSS developers (e.g., Dempsey, Weiss, Jones, & Greenberg, 1999; Koch & Schneider, 2000). Thus, one main objective of the study was to collect first empirical data that explore the prevalence and impact of the mentioned motives within the Linux kernel community. In the meantime, a few other empirical studies on motivation in OSS projects have been presented (Hars & Ou, 2002, Lakhani & von Hippel, 2000) or are in progress (Lakhani, Wolf, Bates, & DiBona, 2002). Although these studies included developers from different OSS projects (Hars & Ou, 2002, Lakhani et al., 2002) or explored support systems instead of software development activities (Lakhani & von Hippel, 2000), the suggested motives of OSS developers are mainly consistent with our speculations presented above.

Relevant Models from Social Science

In addition to a rather intuitive approach to understand the motives of OSS developers, we wanted to base the present investigation also on more systematic approaches as provided by theoretical models from social sciences. Two models from social psychology are particularly relevant in this context.

Participation in Social Movements

A number of social researchers have explored motivational processes of persons who voluntarily engage in social movements such as the civil rights movement, the labor movement, the peace movement, or in the movements of specific social groups such as gay and lesbian people, older people, etc. (e.g., Klandermans, 1997; Omoto & Snyder, 1995; Simon et al., 1998). Social movements can be defined as "effort[s] by a large number of people to solve collectively a problem that they have in common" (Toch, 1965; see also Simon et al., 1998). Although the OSS community might not be a typical social movement, one can argue that some of the political and social goals of the OSS community -- and the Linux community in particular -- can be understood as collective effort to solve a common problem of those who participate (cf. Raymond, 1999). Examples of such goals are more autonomy in modifying software according to personal needs, or protecting the diversity of software solutions against too strong dominance of large economic enterprises. Moreover, the voluntary nature of OSS contributions is another important feature similar to social movements. Thus, even if one can question whether OSS communities such as the Linux community fulfill all criteria of a social movement, there are at least a number of similarities that let us assume that the underlying motives of those who contribute to OSS projects are similar to participants in social movements.

One well established model that explains social movement participation was developed by Klandermans (1997). According to his model, the motivation of participants in social movements depends upon various (expected) costs and benefits. Three basic types of expected costs and benefits are distinguished by Klandermans, comprising three different classes of motives for the participation in and contribution to a social movement. The first class is described as collective motives that are based on the evaluation of the movement goals weighted by the perceived likelihood that these goals are reached. Such expectancy x value constructs have a long tradition in social psychology (e.g., Ajzen & Fishbein, 1980; Atkinson, 1957; Vroom, 1964). According to this approach, the motivation to participate and contribute to a social movement will be higher the higher a person values its goals and the more likely the person perceives the attainment of these goals. Formally, this relation is conceptualized as a multiplicative function of the subjective evaluation of the goals and of the subjective expectation that these goals are reached.

The second class of motives is called social motives and contains expected reactions of significant others such as friends and family members. Motivation to contribute to a movement should be higher the more positive the expected reactions of significant others are, weighted by the perceived importance of these significant others. This relation is formally also expressed as a multiplicative function. It should be noted, however, that the term "social motives" can be misleading because there are other motives that can be considered as "social" as well and that are not included in this concept, such as making new friends or socializing with others. Thus, for better discrimination, we will call motives related to reactions of significant others "norm-oriented motives" throughout this paper.

Finally, the third class of motives includes reward motives that result from expected costs and benefits such as investments of time and money, making new friends, or risking one's health (see also Simon et al., 1998). Similar to the collective motives, these gains and losses are weighted by their expected likelihood and formally expressed as a multiplicative function. It is assumed that the motivation to contribute to a movement is higher the higher and the more likely the expected gains are perceived. The opposite holds for expected losses. Together, these three classes of motives are assumed by Klandermans to determine a person's willingness to actively participate in a social movement.

Recent research has extended this conceptual framework by the aspect of collective identification (Simon et al., 1998). Simon and his colleagues argued that persons not only weight costs and benefits when they decide whether they want be involved in a social movement, but that they also feel and define themselves as members of a specific group related to the social movement and behave according to the norms and standards of this group. Although one can argue that these identification processes are already included within Klandermans' model (e.g., as part of his collective motives class), the empirical data that Simon and his colleagues documented seem to suggest that collective identification processes are another independent factor for the explanation of social movement engagement. Interestingly, their data as well as the data of other studies have shown that identification with more specific subgroups are a better predictor of willingness to contribute to the social movement than identification with the movement as a whole. For example, identification as a feminist activist was a better predictor for the willingness to participate in women-related collective action than identification as a woman (Kelly & Breinlinger, 1995). It seems that the closer a person identifies with active subgroups of a social movement, the more s/he is willing to contribute personally. Similar processes can be assumed for projects of the OSS community.

Thus, when we refer to the "Extended Klandermans Model" (EKM) we refer to four motivational components to explain voluntary action in social movements: Collective motives, norm-oriented motives, reward motives, and identification processes. These components are assumed to contribute additively to the motivation of participants in a social movement or community project. When applied to OSS projects and the Linux kernel community in particular, we expect that the EKM components also explain and predict the motivation to contribute to the OSS project. Please note that most of the above mentioned motives derived from the discussion within the OSS community can be integrated into the EKM components. For instance, intrinsic motives (enjoying writing software) or expected advantages of Linux-related activities for furthering one's career are good examples of reward motives.

The VIST Model of Individuals' Motivation in Teams

The second model relevant for the understanding of developers' motivation in OSS projects stems from research on motivational processes in small work teams. Building on expectancy x value models that explain motivational processes in individual work (e.g., Atkinson, 1957; Vroom, 1964), recent research has extended these approaches to the more complex situation of teamwork in which (perceived) values and expectancies are contingent to a number of additional factors (e.g., Karau & Williams, 2000; Shepperd, 1993). One current model (Hertel, 2002; Hertel, Konradt, & Orlikowski, 2002) has conceptualized these motivational processes particularly for de-located or "virtual teams", i.e. teams in which team members work in different places and coordinate their work mainly via electronic media (cf. Lipnack & Stamps, 1997).

OSS development and Linux kernel development in particular is usually realized via the Internet without face-to-face contact of the contributors. Thus, the collaboration of the contributors can be considered as "virtual collaboration". However, due to the high number of participants and the ease of access, OSS projects are generally better understood as a community or collaborative network (Wellman, 1997) than as a team. Communities usually include a large number of people, and are open to anyone who wants to join as long as s/he obeys some general behavior rules. Collaborative networks are more restrictive in their access policy, relying on referral or reputation and develop a more specific community code including sanctions for violating this code. However, the boundaries of collaborative networks are still relatively flexible, allowing a rather frequent change of collaborators. A "team", in contrast, refers to a relatively small group of collaborating people (about 2 - 20) with clear and relatively stable team boundaries, functions, roles, and norms. Although OSS projects as the Linux kernel project are more a community or collaborative network, team-based approaches might still be relevant to understand developers' motivation as long as part of the development activities are realized in spontaneously formed project teams. Particularly in the Linux kernel structure, such teams might exist for the subsystems mentioned earlier. Thus, one objective of the present study was to explore whether teamwork exists among Linux developers. If such teamwork exists, models of individual's motivation in (virtual) teams might contribute to the understanding of motivational processes in OSS projects.

According to the VIST model (Hertel, 2002), individuals' motivation to work in a virtual team is determined by four main components: valence, instrumentality, self-efficacy, and trust, which are abbreviated by the label "VIST". Valence is defined as the subjective evaluation of the team goals similar to the valence part of collective motives in the EKM outline above. It is assumed that the motivation of a team member is directly proportional to her/his subjective evaluation of the team goals. However, while collective motives in the sense of the EKM refer to the more general goals of a social movement, the valence component refers to more specific goals of a smaller team that might be a subgroup of the social movement.

In addition to this valence component, the other three VIST components contain expectancy concepts of different kinds. Instrumentality is defined as the perceived importance or indispensability of one's own contributions for the group outcome. The higher the perceived instrumentality of one’s own contributions, the higher should be the motivation to exert effort for the team goals. On the other hand, if a person believes that her/his contribution does not matter or cannot be identified, performance motivation should decrease considerably even when this person values the team goals highly. The third component self-efficacy is adopted from work by Bandura (e.g., 1977) and includes team members' perceived capability of showing the required activities for the team tasks. Thus, self-efficacy describes the perceived contingency that one's own high effort leads to high performance. When a team member believes that s/he is unable to accomplish her/his part of the team task, her/his motivation should be low even if the team goals are highly valued and the contribution is perceived as necessary for the team's success.

Finally, trust as the fourth component is defined as the expectancy of team members that their efforts will be reciprocated and not exploited by other team members (interpersonal trust), and that the electronic support system works reliably (trust in the system). Trust is particularly relevant in virtual teams in which misunderstandings and fear of exploitation can escalate more quickly due to fewer face-to-face interactions (e.g., Jarvenpaa & Leidner, 1999). Together, each of the four components is assumed to contribute positively to a member's motivation in a virtual team. While related models of motivation in groups (e.g. Karau & Williams, 2000) assume a multiplicative function of at least some of these components, the VIST model refrains from such specifications since the empirical base is still too shallow. Moreover, a recent meta-analysis of similar concepts for individual work suggests that a multiplicative combination does not explain more variance than the single components alone (van Eerde & Thierry, 1996). Instead, the different components of the VIST model are assumed to contribute rather additively to team members' motivation similar to predictors in a regression approach.

The Present Study

A web-based questionnaire survey was conducted based on the two described motivational models in order to explore the motives of OSS contributors systematically. However, suggestions of possible motives from discussions within the Linux community were integrated in the operationalization of the model components.

The Extended Klandermans Model (EKM) was used to explore more general motives for participating in the Linux kernel community, either by contributing pieces of software or by following the discussion and contributing comments and ideas to the mailing lists. Collective motives were measured with questions about how much participants valued typical goals of the Linux kernel community. Norm-oriented motives were measured by exploring whether evaluations of friends and family members play a role in participants' involvement in the Linux development. Reward motives were measured by asking participants to rate the personal importance of possible gains and losses due to their involvement in the Linux kernel community.

Finally, the identification component was measured related to three different group concepts in order to explore whether identification with more specific subgroups might be more predictive for the participation in social movements than unspecific identification with the movement in general. Accordingly, we measured identification related to three different levels of group specificy, ranging from the general Linux user community over the Linux developer community as a moderately specific group to a certain Linux kernel subsystem as the smallest and most specific group category. This granularity is reflected in the structure of the Linux kernel as well as in the structure of the corresponding mailing lists. Together, the components of the EKM were expected to predict the level of engagement in the Linux development process as indicated by time spent on Linux development and by the willingness to engage in Linux development in the future.

On the other hand, the VIST Model was used to predict the motivation of those developers who actually contribute software to the Linux kernel based on teamwork. Such teamwork was expected to occur spontaneously in so-called "subsystems" of the Linux kernel. Thus, the items related to the VIST Model were only relevant for developers who contributed to such subsystems based on teamwork. Accordingly, the valence, instrumentality, self-efficacy, and trust components were measured with respect to the specific goals and processes within these subsystem teams. Given that such teamwork existed, we expected that the VIST components would also predict part of the average time spent on Linux development (hours per week) and the willingness to be involved in the subsystem work in the future. Moreover, the VIST components should predict more concrete performance criteria such as estimated contributions of lines of computer source code and number of “patches” for the Linux kernel submitted by a developer.[1] Finally, indicators of participants' satisfaction with the Linux kernel development process such as recognition and communication climate were measured for exploratory reasons.

Method

Recruitment and Data Collection Procedure

Data were collected between February 15th and April 12th 2000 via the Internet. About four weeks before the survey started, the study was announced on relevant mailing-lists of the Linux kernel community. Moreover, a web-page was installed on which the goals and the planed procedure of the questionnaire study were explained including information about the persons conducting the study (i.e., the authors of this paper). In order to increase the acceptability of the study as well as to increase the conceptual input, we made an attempt to integrate the Open Source philosophy into the development of the questionnaire by inviting the readers of the announcement to contribute and discus ideas for the questionnaire on a related mailing-list.[2]

About four weeks after this first announcement, the final version of the questionnaire was installed on the internet. The survey was then started with a second announcement on the same Linux kernel mailing lists. A lottery with small prizes (Linux-related products) was included as incentive to participate in the study. The e-mail addresses of participants who chose to participate in this lottery were stored separately from their questionnaire data in order to ensure anonymity. As another incentive, it was announced that a brief summary of the anonymized results together with the raw data of the study would be published on the internet as soon as the study was finished. Similar as the public discussion of the questionnaire items, this procedure integrated the Open Source philosophy in our research approach.

The questionnaire items were completed either by clicking answer options in drop-down boxes or by entering free answers. Since overlong questionnaires are often not completed on the internet (e.g., Batinic & Bosnjak, 2000), we restricted the length of the questionnaire to a completion time of about 15 minutes. After completion of the questionnaire, the participants clicked on a "send" button to transfer their answers to a server at the University of Kiel. At the end of the study, the lottery was conducted, and a first short review of the results was published on the internet along with the raw data as a download file.

<u>The Questionnaire</u>

The variables measured in the questionnaire fall into four main categories. The first category included demographic and context-related items. The second and third categories contained motivational measures based on the EKM and the VIST Model, respectively. Finally, the fourth category included measures of exerted effort and concrete performance as criteria to be predicted by the motivational measures. Moreover, additional questions addressed participants' satisfaction with the Linux developing process for exploratory reasons.

Demographic and Context-Related Items. Items of this category measured demographic data such as the sex, age, nationality, and employment status of the participants. Moreover, it was measured whether participants received payment for their Linux developing work, whether they were able to accomplish Linux-related work during their regular working hours, how much of their spare-time they devoted to Linux programming, and how long they had already been engaged in Linux-related activities. Finally, it was explored whether participants were active contributors to the Linux kernel or rather interested readers of the discussion list.

Items Measuring the Components of the EKM. These questions were developed in a similar way to the study by Simon et al. (1998).[3] The collective motive component was measured by participants' ratings of how much they value the central goals of the Linux kernel community, which were derived from the mentioned sources within the community: Improving the quality of the Linux kernel, improving the quality of the subsystem the participant was working on, and the more general goal that "software should be free". These items were answered on 5-point scales ranging from "very unimportant" (1) to "very important" (5). The norm-oriented motive component was measured by one item which asked participants what relevant others (family, friends, colleagues) think about their activities in the Linux kernel community. The answer scale ranged from "very negatively" (1) to "very positively" (5). The reward motive component was measured by eight items, which included both motives derived from the mentioned discussion within the Linux kernel community as well as motives derived from the literature on social movements (e.g., Klandermans, 1997). Participants rated the perceived importance of the following gains and losses on 5-point scales ranging from very unimportant (1) to very important

(5): Facilitating daily work due to better software, personal exchange with other software developers, gaining reputation as an experienced programmer inside the Linux community, having fun programming, improving one's own programming skills, career advantages, time loss, and lack of payment. Finally, the identification component was measured with a three item scale adapted from Doosje, Ellemers, & Spears (1995; e.g., "I identify with the Linux user community") that was re-worded for each of the three different group categories Identification with Linux users, Identification as Linux developer, and Identification with a certain subsystem, respectively.

Items Measuring the VIST Model Components. These items were adapted from earlier work (Hertel, 2002; Hertel et al., 2002) and referred only to the Linux kernel subsystems since these components were considered as most relevant for the motivational processes in (virtual) teams. Participants were instructed to answer these questions only for the subsystem project that was most important to them. The valence component was measured by two items (e.g., "The success of this subsystem is very important to me."). The instrumentality component was measured by three items (e.g., "I believe my personal contribution is crucial for the success of this subsystem project."). The self-efficacy component was measured by four items (e.g.: "I believe I have the necessary skills to accomplish my tasks in this subsystem project."). Finally, the trust component was measured by three items (e.g.: "I believe that the other developers in my project group invest high efforts in the development of our subsystem."). These items were answered on 5-point scales from "disagree strongly" (1) to "agree strongly" (5). As with the EKM, the number of items was a compromise between issues of reliability and acceptability.

Criterion Variables. To test whether the measured indicators were in fact valid predictors of participants' motivation to contribute to the Linux kernel development, several criterion measures were collected. First, as one main indicator of participants' engagement, participants estimated how many hours they currently spent on Linux development. Second, similar to Simon et al. (1998), participants were asked whether they looked forward to being involved in Linux development in the future, both in Linux more general and in their subsystem project (5-point answer scale from "strongly disagree" to "strongly agree"). Third, as more objective performance criteria, participants estimated how many lines of code and how many "patches" (cf. footnote 1) they had contributed to the Linux kernel. Finally, for exploratory reasons we measured participants' satisfaction with several processes within the development process such as communication, working atmosphere, recognition, and outcomes (5-point answer scales). Four items addressed the Linux kernel development in general and four items addressed processes within the subsystem, respectively.

The Sample

One hundred and fourty one participants (6 females, 135 males; mean age = 30 years, range: 16 – 54 years) from 28 different countries took part in the study. About 48 % of the participants lived in Northern America, 37 % in Europe, and 7 % in Australia. Less than 3 % lived in Asia, Southern America, and Africa, respectively. Most of the participants (67 %) were full-time employees, 5 % worked half-time, and another 5 % were unemployed. The remaining participants were students (23 %). The participants were on average involved in Linux kernel projects for about 17 months (range: 1 - 98 months). About half of the participants (69 persons; 49 %) were actively engaged in the Linux kernel developing process. These participants categorized themselves either as "active Linux kernel developer" or as "maintainer of a module", and were members of a kernel subsystem. We will call this group the "developer group". The other half of the participants (72 persons; 51 %) were readers of the Linux kernel mailing list. We will call this group the "interested readers". However, also the latter group spent some time on Linux kernel development (6.6 hours on average) according to participants' self-ratings.

Participants of the developer group worked on average 18.4 hours per week on Linux development (range: 1 - 70 hours). Twenty percent of the developers received a salary for their Linux programming work on a regular base, and another 23 % at least sometimes. The remaining 57 % indicated that they received no salary at all for their work on the Linux kernel. Thus, not all activities for the Linux kernel development were completely voluntary (see Hars & Ou, 2002, for similar results in other OSS projects). Not surprisingly, the more developers were paid for their Linux-related work the more time they spent, r (n=69) = 0.54, $p < .001$.

About 38 % of the developer group could carry out the Linux-related programming during their regular working hours, although this did not imply that this work was part of their official job. The remaining 62 % carried out Linux-related programming outside of their regular work. Paid developers were more likely to perform Linux-related programming within their working hours, r(n=69) = 0.69, $p < .001$. Investigating how much of participants' spare-time was devoted to Linux-related activities showed that about 40 % of the developer group spent less than 10 % of their spare-time on Linux programming, 48 % of the developer group spent between 10 % and 50 % of their spare-time on Linux programming, and 12 % of the developer group indicated that more than 50 % of their spare-time was devoted to Linux programming. No meaningful correlation occurred between this measure and the questions of payment, indicating that paid and non-paid developers spent more or less the same amount of spare-time on Linux development.

Importantly, 59 % (39 participants) of the developer group indicated that they worked for their subsystem in a group, thus providing evidence that at least some spontaneous teamwork exists within the Linux kernel community. These teams had an average of 12 developers (range: 2 - 50) who were on average 28 months (range: 3 - 98 months) involved in the Linux kernel development. The following team-related analyses based on the VIST model are restricted to the participants of this subgroup.

Results

The result section is divided into two parts. In the first part, we present the analyses related to the components of the EKM as potential motivational predictors of participants' more general engagement in Linux-related activities. Apart from writing software, these activities include also reading and writing comments on the Linux kernel mailing list. The EKM related analyses were conducted across both the developer group and the interested reader group. In the second part, we present the analyses related to the VIST model components as potential motivational predictors of the more specific activities in subsystem teams. These analyses were only conducted with members of the developer group who worked in such teams. In both parts of the result section, we first present tests whether our developed scales were reliable and valid measures of the motivational constructs. Then, we document the means and standard deviations of the measured constructs. Finally, regression analyses were conducted in order to explore whether the motivational constructs significantly predicted participants' Linux-related activities.

Preliminary Analyses of the EKM Items

First, we tested the reliability and construct validity of our newly developed EKM scales based on factor analyses. Please note that such analyses also provide information about the underlying motivational structure of the participating Linux community members. A principal component analysis (see e.g. Bryant & Yarnold, 1997, for an introduction) across all participants revealed seven factors with Eigenvalues > 1 (varimax rotation). The first factor explained 22.9 % variance and was mostly determined by items addressing identification as a Linux developer as well as identification with the subsystem (factor loadings between .38 and .85). This indicates that the participants did not discriminate much in their identification between the two group categories. Thus, we calculated for each participant one mean score for the specific identification as a developer and with the related subsystem (alpha = 0.85). The second factor explained an additional 11.2 % variance and was mostly determined by the three items measuring the more general identification as a Linux user (factor loadings between .82 and .85). These items were combined to a scale score of identification as a Linux user (alpha = 0.77). Together, these first two identification factors confirmed the assumed independence of identification processes related to more global and more specific group categories (cf. Simon et al., 1998).

The third factor explained an additional 8.5 % variance and was determined by the two collective motive items that addressed improving the quality of Linux (general and subsystem level), and by two reward motives that also addressed quality issues (facilitating daily work with

Linux software; career advantages due to experiences with Linux development; factor loadings between .31 and .73). This factor seems to measure pragmatic motives related to the improvement of the Linux kernel. A mean score of these items was calculated for each participant as another predictor for the following analyses (alpha = 0.71).

The fourth factor explained additional 6.5 % variance and was mainly determined by the item measuring norm-oriented motives (factor loading = .78). This result confirmed the assumed independence of norm-oriented motives from reward and collective motives, as well as from identification aspects. The fifth factor explained additional 5.8 % variance and was mostly determined by the reward motive "having fun programming" (factor loading = .74). As suggested in the introduction section, hedonistic motives or intrinsic motivation seem to be another independent factor relevant for Linux-related activities.

The sixth factor explained additional 5.6 % variance and was mostly determined by the four remaining reward motives (personal exchange with other software developers, gaining reputation as an experienced programmer inside the Linux community, improving one’s own programming skills, lack of payment [reversed]; factor loadings between .39 and .66). Moreover, the third collective motive item ("software should be free") showed its highest loading on this factor (.56). Thus, this sixth factor could be interpreted as social and political motives in the sense of supporting free software, and networking and socializing within the Linux community. Although the scale consistency of these items was low, alpha = 0.28, we still calculated a mean score for each participant based on these five items. As noted by Simon et al. (1998), a high internal consistency is not a necessary requirement for the reward motive to be included in the analyses since persons can be motivated by some outcomes but not by others, and this selectivity should vary across people. Finally, the seventh factor which explained additional 5.2 % variance was mostly determined by the [reversed] reward motive item measuring participants’ evaluation of time losses (factor loading = .87). Thus, a considerable tolerance of time investments on issues related to the Linux kernel seems to be another independent characteristic of the members in the

Linux kernel community.

Overall, the results suggest that the motives of persons to participate in the Linux kernel community can be categorized into seven major components that explain a total of 65.6 % variance. These factors largely overlap with components specified by the EKM. Apart from the distinction between more general and more specific group identification processes, norm-oriented motives could be distinguished as well. Collective and reward motives were not as clearly distinguishable in our sample but were instead split into four more specific factors measuring different aspects of possible outcomes: (a) a pragmatic component related to improving the Linux kernel for personal advantages, (b) a social/political component related to supporting free software and networking with other Linux developers, (c) a hedonistic component related to intrinsic motivation, and (d) a component mainly related to concerns of time losses.

While these analyses demonstrated which motivational components could be reliably measured and distinguished in our sample, in the following analyses we describe the relative importance of these motivational components for Linux-related activities.

<u>Means and Correlations of EKM Scores and Linux-related Activities</u>

As depicted in Table 1, the means of the motivational components were all rather high. Thus, participants identified highly both as a Linux user and as a Linux developer, indicated high pragmatic, norm-oriented, social/political, and hedonistic motives, and valued time losses due to Linux-related activities rather low.

< Insert Table 1 about here >

However, comparing the means of participants in the developer and in the interested readers group showed some interesting differences (see Table 2). First, participants in the developer group had higher means in the general engagement measures than participants in the interested reader group. The participants of the developer group spent significantly more hours per week on Linux development (M = 18.4 versus M = 6.6), t(136) = 4.82, p < .001, and showed significant higher willingness to be further involved in Linux development (M = 4.5 versus M = 3.9), t(125)

= 3.60, $p < .001$. More interestingly in terms of motivational differences, participants of the developer group compared to participants of the interested reader group showed marginally lower identification with the Linux community in general, M's = 3.9 vs. 4.2, $t(137) = 1.74$, $p < .09$, but significantly higher identification with more specific categories such as developers or a specific subsystem, M's = 4.0 vs. 3.3, $t(137) = 4.49$, $p < .001$. Moreover, participants of the developer group compared to participants of the interested reader group showed marginally higher means for the pragmatic motive, M's = 4.3 vs. 4.1, $t(139) = 1,9$, $p < .07$, and for the norm-oriented motive, M's = 3.9 vs. 3.6, $t(134) = 1.97$, $p < .06$. Thus, those participants who accomplished most of the programming work (the developer group) identified more strongly with more specific group categories in the Linux community (as a developer or with a subsystem), and indicated higher concerns for reactions of significant others as well as higher pragmatic interests in improving the quality of the Linux kernel.

< Insert Table 2 about here >

Next, in order to explore which of the measured motives were predictive for the engagement of participants, a number of correlational analyses were performed. As expected, most of the motivational variables correlated significantly with the number of hours participants invested in Linux development, as well as with participants' willingness to be further involved in Linux-related activities (cf. Table 1). Thus, the measured EKM indicators covered the underlying motivation of participants to engage in more general Linux-related activities quite well.

As a more integrative step, we then performed a number of regression analyses to explore which of the measured EKM-related indicators were the most relevant predictors for participants' more general engagement in the Linux kernel community. Compared to singular correlation analyses, a regression analysis considers not only one predictor but a certain set of predictor variables at the same time in the same analysis. The results provide insight into the relative importance of each predictor variable in this specific context (e.g., Licht, 1997, for an introduction to the logic of regression analyses).

In the first regression, all seven motivational components were included as predictor variables, and the amount of hours participants spent on Linux-related activities was entered as criterion variable. This analysis yielded a significant solution, $F(7,123) = 5.31$, $p < .001$ with $R^2 = .19$ (adjusted for number of predictors), and revealed significant effects for both general and specific identification components and for the time loss component. While specific identification as a Linux developer or with the subsystem showed a positive beta weight (beta = .28, $p < .01$), the beta weight of the more general identification as a Linux user showed a negative score (beta = -.33, $p < .001$). As already apparent in the correlation analyses (Table 1), participants spent more hours on Linux-related activities when they identified with the specific categories (Linux developer, subsystem) but not as a Linux user in a more general sense. In addition, the less important participants rated time losses due to Linux development, the more hours they spent, beta = .15, $p < .04$ (directional test). All other predictors showed no significant effects.

Performing this regression only for the developer group showed very similar results. However, performing this regression only for the interested reader group showed only a significant effect for the time loss component, while the identification measures were not significant this time. Thus, while a low concern for time losses seems to foster Linux-related activities for both groups of participants, the specific identification as a Linux developer or with a Linux subsystem is an important additional motive for the developer group.

Similar results were obtained when regression analyses were performed with the same motivational predictors and willingness to be involved in Linux development in the future as criterion variable. This time also the pragmatic motives component had a significant effect (beta = .30, $p < .01$). Thus, the higher participants perceived personal rewards for their Linux engagement, the more they were willing to be involved in Linux-related activities in the future. Separate analyses of participants who received payment and participants who received no payment did not lead to different results of the main regression analyses.

Together, these results suggest that while all components related to the EKM were significantly correlated with at least one of the criterion variables of Linux-related engagement, the most important predictors for participants' engagement were (a) a more specific identification as a Linux developer or with a subsystem, (b) a considerable tolerance in respect to time losses due to Linux development activities, and (c) a rather pragmatic interest in personal advantages due to improving the Linux kernel quality.

We now turn to the motivational processes of Linux developers who worked in a subsystem together with other developers as a team. This form of collaboration can be perceived as a "virtual team" which contains challenges particularly in respect to performance motivation due to low face-to-face contact (Hertel, Deter, & Konradt, in press). The following analyses explore motivational processes in Linux subsystem teams based on the VIST model.

Preliminary Analyses of Items Related to the VIST Model

Prior to the main analyses, we again explored the reliability and construct validity of our developed scales.[4] A principal component analysis with varimax rotation revealed four factors with Eigenvalues > 1 that explained 68.9 % variance. The empirical structure was equivalent to the theoretically expected structure of the VIST model, providing evidence for the assumed independence of its components. The first factor explained 27.5 % variance and was mostly determined by the instrumentality items (factor loadings between .70 and .90). The second factor explained additional 19.0 % variance and was mostly determined by the self-efficacy items (factor loadings between .64 and .72). The third factor explained an additional 14.0 % variance and was mostly determined by the trust items (factor loadings between .60 and .87). Finally, the fourth factor explained additional 8.5 % variance and was mostly determined by the valence items (factor loadings between .64 and .89). Based on these results, we calculated mean scores for the four scales for each participant, respectively (alpha of the trust scale =.66, all other alphas > .80). The following analyses of motivational processes within the Linux subsystems were based on participants who described themselves as developers and who indicated that they worked for their subsystem in a team with other developers (n = 39).

Means and Correlations of VIST Scores and Performance Variables

Means and standard deviations of the VIST components are documented in Table 3. These scores indicate a generally high motivation of developers in the subsystem teams.

< Insert Table 3 about here >

To explore whether the VIST components were also predictive for the engagement of the participants, we correlated the VIST components with several criteria variables. Two of them measured participants' motivation to work for the subsystem project (hours per week spent on Linux development, willingness to further engage in Linux development) while the other two criteria measured objective performance output (Number of patches, Lines of code). As can be seen in Table 3, a number of significant correlations occurred for the first three VIST components valence, instrumentality, and self-efficacy. Only the trust component seemed to be rather unimportant in the Linux subsystem teams. As a more integrative step, we again computed a number of regression analyses to explore the impact of the VIST components for each of the four performance criteria when all VIST components are taken into account simultaneously.

Hours Spent on Linux Development. Regressing the number of hours the developers spent on average per week for Linux development showed a significant effect for the instrumentality component, beta = .34, $p < .04$. Thus, developers spent more time on Linux development when they felt that their contribution was highly important for the progress of the subsystem. This result is consistent with other research on motivational processes in virtual teams (Hertel et al., 2002; in press). The other VIST components did not improve the level of explained variance ($R^2 = .09$), indicating that these components played only a minor role in the time invested by developers in this sample.

Willingness to Increase Participation in the Subsystem in the Future. Regressing this criterium on the VIST components showed a significant effect of the valence component (beta = .40, $p < .02$) and the instrumentality component, beta = .34, $p < .04$. Developers were willing to spend more time in the subsystem the higher they valued its goals and the higher they perceived their contribution as important for the project success. Including both components as predictors into a regression of this criterium yielded the best solution, $R^2 = .19$. No further improvement was found by including the other VIST components into this regression.

Number of Patches Accepted. The estimated number of patches a developer had contributed to the Linux kernel and which had been accepted is a performance output measure that is probably highly determined by the skills and experience of a person. However, since performance is always also a function of motivational factors, we still assume that the VIST components might contribute to at least some part of the variance of this criterion. Indeed, regression analyses showed significant effects of the instrumentality component (beta = .35, $p < .03$) and the self-efficacy components (beta = .36, $p < .03$). Higher numbers of accepted patches were accompanied by higher self-efficacy and stronger feelings that one's own contributions were crucial for the subsystem. Including both components as predictors into a regression of the number of patches yielded the best solution, $R^2 = .12$, although this time the beta weights of both the instrumentality and the self-efficacy component were only marginally significant (beta = .24 and .25, respectively; $p < .09$, directional tests).

Lines of Code. Regressing the (log.-transformed) estimations of lines of code added to the Linux kernel as another performance output indicator showed only a significant effect of perceived self-efficacy (beta = .32, $p < .04$, directional test), while the other VIST components did not improve the amount of explained variance, $R^2 = .10$. Thus, only perceived own ability was predictive for this performance indicator but none of the more group-related components of the VIST Model.

Together, these results showed that motivational processes in virtual teams as specified by the VIST model can also be observed in the Linux kernel development process. The instrumentality and valence components were particular predictive for the measured motivational criteria such as time investment and willingness to engage in the future. The output criteria were more strongly predicted by participants' perceived self-efficacy, presumably because objective performance is strongly determined by participants' programming skills and expertise. However, at least for one of the two output criteria we could also observe effects of perceived instrumentality of one's own contributions to the subsystem team. On the other hand, trust or fear of being exploited as the fourth VIST component seemed to play only a minor role in the Linux subsystem teams, at least in this study.

Satisfaction and Working Climate

For exploratory reasons, we also measured participants' satisfaction with the Linux development process and the working climate in the subsystems. Specifically, we asked the participants to value their satisfaction with the communication processes, with the working atmosphere, with the recognition of contributions, and with the results of the subsystem. These items were highly correlated and could be combined as a scale (alpha = 0.76). The average satisfaction score of developers who worked in subsystem teams was high, M = 4.4 (SD = 0.6, scale range between 1 and 5), indicating that these developers were generally quite happy with the working processes. Similar questions addressing satisfaction with the Linux kernel development more generally (alpha = 0.80) showed also high overall satisfaction, M = 4.2 (SD = 0.7; scale range between 1 and 5). While no significant differences occurred between participants of the developer group and interested reader group in the general satisfaction score, participants of the developer group more often indicated that they had experienced "burnout" during their Linux-related work. In the developer group, 17.4 % indicated that they had experienced burnout at least "once", and 46.4 % even "sometimes". In the interested reader group, only 4.2 % indicated that they had experienced burnout "once", and 18.1 % "sometimes".

Discussion

One of the most compelling aspects of Open Source Software projects is that they are predominantly based on voluntary contributions from software developers without organizational support in a traditional sense (Moon & Sproull, 2002). One central question is, what motivates these persons to contribute to OSS projects "for free", and what rewards do they expect. The objective of the present study was to explore these motives empirically within one of the most prominent OSS projects: The Linux kernel project. Together with other recent studies (Hars & Ou, 2002; Lakhani & von Hippel, 2000; Lakhani et al., 2002), the present investigation is one of the first studies that provide sound empirical data on motivational processes within OSS projects.

In doing so, we tried to integrate methodological and theoretical issues from the OSS community with research methods and models from social sciences. For instance, while the statistical analyses of the questionnaire data were based on standard methods from social sciences, the questionnaire items were discussed with the involved persons in a public mailing list and the questionnaire data were published on the Internet. In a similar way, the development of the questionnaire items was both based on rather intuitive considerations from involved members of the Linux kernel community (e.g. Moon & Sproull, 2002; Raymond, 1999; Torvalds & Diamond, 2001) as well as on systematic models from social psychology. The obtained results showed that these approaches complemented one another nicely. Central motives that were discussed within the Linux community could be integrated in the systematic social psychological models. However, these models also suggested additional motives that could be confirmed in our empirical study. Together, the results demonstrated that motivational processes within OSS projects, such as the Linux kernel development, do not differ completely from motivational processes in other social communities and teams and can be explained within existing social psychological theories.

With regard to the engagement for the Linux kernel more generally (i.e., including reading and contributing to the Linux kernel mailing list), seven factors have been identified as distinct motivational sources for the participants. These factors were largely consistent with the assumptions of the Extended Klandermans Model of voluntary action in social movements (Klandermans, 1997; Simon et al., 1998). Thus, engagement for the Linux kernel community seemed to be driven by similar motives as voluntary action within social movements such as the civil rights movement, the labor movement, or the peace movement. The main motivational factors were (a) a more general identification factor as Linux user, (b) a more specific identification factor as a Linux developer or with a Linux subsystem, (c) pragmatic motives related to the improvement of one's own software and career advantages, (d) norm-oriented motives related to reactions of relevant others (family, friends, colleagues), (e) social and political motives related to supporting independent software and networking within the Linux community, (f) hedonistic motives such as pure enjoyment of programming, (g) and motivational obstacles related to time losses due to Linux-related activities (cf. Hars and Ou, 2002, for similar results in a study including other OSS projects).

All seven factors showed rather high mean scores and correlated positively with the measured criteria of invested time for Linux-related activities and/or willingness to engage in Linux-related activities in the future. However, the more integrative regression analyses revealed that not all seven factors were equally predictive for participants' engagement. With regard to invested hours per week, the strongest predictors were the specific identification as a Linux developer or with a subsystem, and a considerable tolerance of time losses due to Linux-related activities. This is particularly true for those participants who could be categorized as active developers in a more narrow sense. For the group of interested readers, only concern for time losses showed a significant effect when all seven possible predictors were integrated simultaneously in a regression analysis. This suggests that lack of time is one of the biggest obstacles for participating in the Linux kernel project.

Finally, as a third important predictor in the regression analyses, the pragmatic motive factor showed significant effects in the regression of willingness to engage in Linux-related activities in the future. Thus, the interest to improve software for one's own use and to increase one's own career perspectives seemed to be particulary important motives for planned activities in this community. However, this motivation might diminish once a developer has actively joined a project and learned that these expectations are not always so easily fulfilled.

In respect to more specific programming work in subsystem teams, the four central factors derived from the VIST model (Hertel 2002) could also be confirmed as underlying motives of participants. This result is in accordance with current research on virtual teams in business organizations (Hertel et al., 2002), suggesting that similar processes can be found in virtual teams of OSS projects with rather low level of explicit organization. Similar to the EKM factors, the VIST factors also showed generally high mean scores. However, only the first three VIST factors correlated significantly with at least one of the motivation/performance criterion measures. Participants' activities were particularly determined by (a) their subjective evaluation of the subsystem goals, (b) the perceived importance of their own contributions for the subsystem, and (c) the perceived personal ability to accomplish the tasks.

On the other hand, perceived trust within the subsystem seemed to play only a minor role for motivation and performance in the Linux kernel subsystem teams. This is a surprising result since fear that one's own efforts might be exploited by other users (or by commercial companies) certainly is an issue in OSS development (e.g., Moon & Sproull, 2002; Stallman, 1994 a, b). For example, it has been vigorously discussed on the Linux kernel mailing list how to deal with contributions (so-called "modules") that are not available under the General Public License. Perhaps, distrust and fear of exploitation play only a minor role in Linux subsystems since the membership is determined by contributions per se, implying that a developer who neither posts constructive messages to the mailing list nor submits patches is not part of the subsystem.

The exploratory data on satisfaction with the communication processes and the working climate in the subsystems and within the Linux community more general showed quite positive results. Both participants of the developer group and the interested reader group seemed to enjoy their Linux-related activities (at least those who participated in our study). However, among the developers, there is also the danger of negative side effects due to high voluntary engagement. More than 50 % of participants in the developer group indicated that they had experienced feelings of burnout at least once. Although this certainly is not a diagnosis in a clinical sense, maintainer or organizer of OSS projects should be aware of possible burnout effects and should take preventive steps against it.

A number of other practical implications arise from the reported results that are not only relevant for the Linux kernel community and other OSS projects, but that might also be applied to software development within commercial enterprises if it follows a rather unstructured "bazaar"-like style. One important issue is that identification processes are important for such software development projects. In accordance with research in other areas of voluntary engagement (Kelly & Breinlinger, 1995; Simon et al., 1998), this identification should be related to smaller and more active units of a project rather than to the project in general. Concerns about time losses, as another important motivational factor, can be addressed by streamlining the organizational procedures in OSS projects as well as other software development processes so that ineffective steps and interactions are minimized. Helpful structural features that have been successfully realized in the Linux project are a module structure that prevents unneccessary coordination requirements and the simplicity in communication processes (Moon & Sproull, 2002). In addition, attempts can be made to increase the time-resources of participants whenever possible.

Pragmatic motives to improve one’s own software tools and to increase personal career chances were particularly relevant for participants’ willingness to engage in Linux activities in the future, but played only a minor role for the hours the participants actually have spent in the past. Thus, unrealistic expectations might be a danger for the motivation of participants in OSS projects, and it might be fruitful to inform new participants thoroughly about possible costs and gains of their activities in a project. Moreover, pragmatic motives should be mainly relevant for software projects for which users have basic programming skills. Further research might compare the demographic and motivational structure of OSS projects depending on the specific type of the software developed (operating system, business application, etc.).

Finally, the results revealed that in OSS projects some part of the development work is accomplished by (spontaneous) teams. The successful validation of the VIST model within the Linux subsystems provides a practical heuristic to address motivational challenges in such teams when members collaborate from different locations and different time zones. Similar to virtual teams in business organizations (Hertel et al., 2002), the perceived indispensability of one's own contributions for the team is an important motivational factor together with a high evaluation of the team goals and a high sense of personal self-efficacy.

Before closing, we want to mention two important limitations of the present study. First, although the sample size was sufficient to conduct first quantitative analyses, it certainly would be desirable to replicate the presented results in follow-up studies with larger samples, either within the Linux community or within other OSS projects (see Lakhani et al., 2002, for an example of such a study that is currently in progress). Apart from a higher reliability of the results and higher representativity of the sample, it would also be possible to prevent biases due to systematic differences between responding and non-responding OSS community members. In addition to self-report data, objective performance criteria and network analyses (e.g., Wellman, 1997) might be included in such follow-up research. Second, the present study is based on a cross sectional correlational data set. Thus, conclusions about causal processes are not clear. For instance, high correlation between subsystem identification and time investments can either occur because identification leads to increased time investments, because time investments lead to increased identification, or because both variables are affected by a third variable such as pragmatic motives. These questions have to be addressed in complementing longitudinal and experimental studies.

In conclusion, the present study has contributed to our understanding of motivational processes in OSS projects and has revealed various motivational forces that contribute to a person's willingness to engage in such projects both at the community level as well as at the team level. The motivational forces are in accordance with existing models of voluntary action and virtual teamwork. Moreover, this study also provides an example of how methodological principles of the OSS community might be integrated in social research. Together, we hope that this research helps to improve software development processes both in OSS projects as well as in other projects when the collaboration depends heavily on the voluntary engagement of persons in a rather low structured environment.

References

Ajzen, I. & Fishbein, M., 1980, Understanding attitudes and predicting social behavior. (Prentice-Hall, Englewood Cliffs).

Atkinson, J.W., 1957, Motivational determinants of risk-taking behavior, Psychological Review 64, 359-372.

Bandura, A., 1977, Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review 84, 191-215.

Batinic, B. & Bosnjak, M., 2000, Fragebogenuntersuchungen im Internet [Surveys on the Internet], in: B. Batinic (Editor), Internet for Psychologists (Hogrefe, Göttingen, Germany) pp. 278-318.

Bezroukov, N., 1999 a, Open source software as a special type of academic research (critique of vulgar Raymondianism). First Monday 4 (Retrieved from the World Wide Web, October 28, 2002: http://www.firstmonday.dk/issues/issue4_10/bezroukov/index.html).

Bezroukov, N., 1999 b, A second look at the Cathedral and the Bazaar. First Monday 4 (Retrieved from the World Wide Web, October 28, 2002: http://www.firstmonday.dk/issues/issue4_12/bezroukov/index.html).

Bovet, P.D. & Cesati, M., 2001, Understanding the Linux kernel (O'Reilly, Sebastopol).

Bryant, F.B. & Yarnold, P.R., 1997, Principal-component analysis and exploratory and confirmatory factor analysis, in: L.G. Grimm & P.R. Yarnold (Editors), Reading and understanding multivariate statistics (4th. edition) (American Psychological Association, Washington), pp. 99-136.

Dempsey, B.J., Weiss, B., Jones, D., & Greenberg, J., 1999, A quantitative profile of a community of Open Source Linux developers, University of North Carolina at Chapel Hill, School of Information and Library Science, Technical Report TR-1999-05.

Doosje, B., Ellemers, N., & Spears, R., 1995, Perceived intragroup variability as a function of group status and identification. Journal of Experimental Social Psychology 31, 410-436.

Duarte, D.L. & Snyder, N.T., 1999, Mastering virtual teams (Jossey-Bass, San Francisco).

Hars, A. & Ou, S., 2002, Working for free? Motivations for participating in Open-Source projects. International Journal of Electronic Commerce 6, 25-39.

Hertel, G., 2002, Management virtueller Teams auf der Basis sozialpsychologischer Modelle [Management of virtual teams based on social psychology models], in: E.H. Witte (Editor), Sozialpsychologie wirtschaftlicher Prozesse (Pabst Publishers, Lengerich, Germany), pp. 172-202.

Hertel, G., Deter, C., & Konradt, U., in press, Motivation gains in computer-supported teams. Journal of Applied Social Psychology.

Hertel, G., Konradt, U., & Orlikowski, B., 2002, Managing distance by interdependence: Goal setting, task interdependence, and team-based rewards in virtual teams. Manuscript under editorial review.

Jarvenpaa, S.L. & Leidner, D.E., 1999, Communication and trust in global virtual teams. Organization Science 10, 791-815.

Karau, S.J., Markus, M.J., & Williams, K.D., 2000, On the elusive search for motivation gains in groups: Insights from the collective effort model. Zeitschrift für Sozialpsychologie 31, 179-190.

Karau, S.J. & Williams, K. D., 2000, Understanding individual motivation in groups: The collective effort model, in: M.E. Turner (Editor), Groups at work: Advances in theory and research (Erlbaum, Mahwah) pp. 113-142.

Kelly, C. & Breinlinger, S., 1995, Identity and injustice: Exploring women's participation in collective action, Journal of Community & Applied Social Psychology 5, 41-57.

Klandermans, B., 1997, The social psychology of protest (Basil Blackwell, Oxford).

Koch S. & Schneider G., 2000, Results from software engineering research into open source development projects using public data, in: H.R. Hansen & W.H. Janko (Editors), Wirtschaftsuniversitaet Wien, Working Paper #22 “Diskussionspapiere zum Taetigkeitsfeld Informationsverarbeitung und Informationswirtschaft”.

Lakhani, K.R. & von Hippel, E., 2000, How Open Source Software Works: “Free” User-to-User Assistance, MIT Sloan School of Management Working Paper #4117 (May). Research Policy, forthcoming.

Lakhani, K.R., Wolf, B., Bates, J., & DiBona, C., 2002, The Boston Consulting Group Hacker Survey (Retrieved from the World Wide Web, September 11, 2002: http://www.bcg.com/opensource/BCGHackerSurveyOSCON24July02v073.pdf).

Licht, M.H., 1997, Multiple Regression and correlation, in: L.G. Grimm & P.R. Yarnold (Editors), Reading and understanding multivariate statistics (4th. edition) (American Psychological Association, Washington) pp. 19-64.

Lipnack, J. & Stamps, J., 1997, Virtual teams (John Wiley, New York).

Moon, J.Y. & Sproull, L., 2002, Essence of distributed work: The case of the Linux kernel, in: P. Hinds & S. Kiesler (Editors), Distributed work (MIT Press, Cambridge), pp. 381-404. (Also available on the World Wide Web: http://www.firstmonday.dk/issues/issue5_11/moon/index.html, retrieved October 28, 2002)

Omoto, A.M. & Snyder, M. 1995, Sustained helping without obligation: Motivation, longevity of service, and perceived attitude changeamong AIDS volunteers, Journal of Personality and Social Psychology 68, 671-686.

Raymond, E.S., 1999, The cathedral and the bazaar. Musings on Linux and Open Source by an accidental revolutionary (O'Reilly, Sebastapol, CA).

Shepperd, J.A., 1993, Productivity loss in performance groups: A motivation analysis. Psychological Bulletin 113, 67-81.

Simon, B., Loewy, M., Stürmer, S., Weber, U., Freytag, P., Habig, C., Kampmeier, C., & Spahlinger, P., 1998, Collective identification and social movement participation. Journal of Personality and Social Psychology 74, 646-658.

Stallman, R.M., 1994 a, The GNU manifesto (Retrieved from the World Wide Web, October 28, 2002: http://www.gnu.org/gnu/manifesto).

Stallman, R.M., 1994 b, Why software should not have owners (Retrieved from the World Wide Web, October 28, 2002: http://www.gnu.org/philosophy/why-free.html).

Toch, H., 1965, The social psychology of social movements (Bobbs-Merrill, Indianapolis).

Torvalds, L., 1998, Interview with Linus Torvalds: What motivates free software developers? First Monday 3 (Retrieved from the World Wide Web, December 14, 2001: http://www.firstmonday.dk/issues/3_3/torvalds).

Torvalds, L. & Diamond, D., 2001, Just for Fun: The Story of an accidental revolutionary (Harper Business, New York).

Van Eerde, W. & Thierry, H., 1996, Vroom's expectancy models and work-related criteria: A meta-analysis. Journal of Applied Psychology 81, 575-586.

Vroom, V.H., 1964, Work and motivation (John Wiley, New York).

Wellman, B. (1997). An electronic group is virtually a social network, in: S. Kiesler (Editor), Culture of the internet (Erlbaum, Mahwah), pp. 179-208.

Footnotes

* We thank SuSe Germany for sponsoring a number of lottery prizes for the participants in this study.

1 Computer source code is arranged in lines, where each line corresponds to one instruction. Thus, the number of these lines of code can be used as an objective estimation of the quantity of contributions by one developer. Several lines of code, in turn, form a "patch" which is an atomic contribution to a software project, submitted as one file and addressing one single issue. Consequently, the number of patches produced by one developer is an estimation of his or her number of issues addressed. For reasons of anonymity, it was not possible for us to measure the contributions of the participants objectively.

2 Unfortunately, the traffic on this mailing list was rather low. Only 15 persons subscribed to the list, and a real discussion never started.

3 Due to the mentioned space restrictions for internet-based questionnaires we refrained from including expectancy weights for all motives. The reported analyses were restricted to the collected importance ratings. Such a procedure is appropriate since importance ratings often already include expectancy considerations implicitly.

4 Although the later analyses are restricted to developers who worked in subsystem teams, the validation of the VIST scales was based on all participants of this study for power reasons.

Table 1

Participation in Linux Kernel Development (n=141): Means and Intercorrelations of Seven Motivational Components and Two Effort Criteria.

Variable	1	2	3	4	5	6	7	8	9
1. General Identification (Linux User)	----								
2. Specific Identification (Developer/ Subsystem)	.08	----							
3. Norm-oriented Motives	.09	.29 **	----						
4. Pragmatic Motives	.15	.43 **	.33 **	----					
5. Social / Political Motives	.26 **	.18 *	.21 *	.31 **	----				
6. Hedonistic Motives	.21 *	.28 **	.12	.10	.10	----			
7. Time Loss (Rev. Coding)	.06	.09	.03	.10	.12	-.07	----		
8. Hours/Week Spent on Linux Development	-.28 **	.32 **	.15 †	.16 †	.01	.01	.18 *	----	
9. Willingness to be Involved in the Future	-.02	.38 **	.12	.40 **	.22 *	.19 *	.20 *	.22 *	----
M [a]	4.0	3.6	3.7	4.2	4.1	4.6	3.4	12.5	4.2
SD	0.8	0.9	0.8	0.7	0.5	0.9	1.0	15.4	0.9

[a] Scale range of all scales except "hours/week spent on Linux development" varied between 1 and 5; see Method section for details. Higher means indicate higher identification, stronger norm-oriented, pragmatic, social/political, and hedonistic motives, and lower importance of time losses due to involvement in Linux development, and higher willingness to be involved in Linux development in the future.

† $p < .10$. * $p < .05$. ** $p < .01$ (two-tailed)

Table 2

Means of Motivational Components and Effort Criteria for the Developer Group and the Interested Readers Group.

	Developer Group (n = 69)		Interested Reader Group (n = 72)		Difference
	M	SD	M	SD	
1. General Identification (Linux User)	3.9	0.9	4.2	0.7	$p < .09$
2. Specific Identification (Developer/ Subsystem)	4.0	0.7	3.3	1.0	$p < .001$
3. Norm-oriented Motives	3.9	0.8	3.6	0.9	$p < .06$
4. Pragmatic Motives	4.3	0.5	4.1	0.8	$p < .07$
5. Social / Political Motives	4.1	0.4	4.1	0.5	n.s.
6. Hedonistic Motives	4.7	0.6	4.5	1.0	n.s.
7. Time Loss (Rev. Coding)	3.6	1.0	3.3	1.1	$p < .11$
8. Hours/Week Spent on Linux Development	18.4	18.0	6.6	9.2	$p < .001$
9. Willingness to be Involved in the Future	4.5	0.8	3.9	1.1	$p < .001$

[a] Scale range of all scales except "hours/week spent on Linux development" varied between 1 and 5; see Method section for details. Higher means indicate higher identification, stronger norm-oriented, pragmatic, social/political, and hedonistic motives, and lower importance of time losses due to involvement in Linux development, and higher willingness to be involved in Linux development in the future. The Difference column indicates the results of t-tests (two-tailed) between the means of the two groups for each variable.

Table 3

Developers' Motivation in the Kernel Subsystem Teams (n = 39): Intercorrelations between VIST Components and Effort and Performance Criteria.

Variable	1	2	3	4	5	6	7	8
1. Valence	----							
2. Instrumentality	.19	----						
3. Self-efficacy	.07	.46**	----					
4. Trust	.21	.03	-.03	----				
5. Hours/Week Spent on Linux Development	.19	.34*	.21	.19	----			
6. Increase Participation in the Subsystem (Rev. Coding)	.40**	.34*	.09	.06	-.02	----		
7. Number of Patches	.05	.35*	.36*	.06	.32*	.09	----	
8. Lines of Code (log.-transformed)	-.13	.08	.32*	-.06	.05	-.33*	.31*	----
M [b]	4.7	3.4	4.0	3.8	20.5	2.4	2.4	2.6
SD	0.5	1.0	0.6	0.8	19.4	1.0	1.6	0.9

[b] Scale range of all scales except "hours/week spent on Linux development", "number of patches", and "lines of code" varied between 1 and 5; see Method section for details. Higher means indicate higher valence, instrumentality, self-efficacy, trust, and willingness to increase participation in the subsystem.

* $p < .05$. ** $p < .01$ (two-tailed)