Motivation, values, and work design as drivers of participation in the R open source project for statistical computing

Patrick Mair[a,1], Eva Hofmann[b], Kathrin Gruber[c], Reinhold Hatzinger[d,2], Achim Zeileis[e], and Kurt Hornik[f]

[a]Department of Psychology, Harvard University, Cambridge, MA 02138; [b]Department of Applied Psychology: Work, Education and Economy, University of Vienna, 1010 Vienna, Austria; [c]Service Marketing and Tourism, Wirtschatsuniversität Wien, 1020 Vienna, Austria; [d]Wirtschaftsuniversität Wien Vienna University of Economics and Business, 1020 Vienna, Austria; [e]Department of Statistics, University of Innsbruck, 6020 Innsbruck, Austria; and [f]Institute for Statistics and Mathematics, Wirtschatsuniversität Wien, 1020 Vienna, Austria

Edited by Robert J. Tibshirani, Stanford University, Stanford, CA, and approved October 1, 2015 (received for review March 26, 2015)

One of the cornerstones of the R system for statistical computing is the multitude of packages contributed by numerous package authors. This amount of packages makes an extremely broad range of statistical techniques and other quantitative methods freely available. Thus far, no empirical study has investigated psychological factors that drive authors to participate in the R project. This article presents a study of R package authors, collecting data on different types of participation (number of packages, participation in mailing lists, participation in conferences), three psychological scales (types of motivation, psychological values, and work design characteristics), and various socio-demographic factors. The data are analyzed using item response models and subsequent generalized linear models, showing that the most important determinants for participation are a hybrid form of motivation and the social characteristics of the work design. Other factors are found to have less impact or influence only specific aspects of participation.

R Project for Statistical Computing | Schwartz values | motivation | work design | item response theory

The story of the R environment for statistical computing (1) has been one of tremendous success. Since it was first conceived (2), R has been attracting more and more users and contributors from different fields where data analysis plays a major role. Fox (3) conducted a series of interviews with members of the R Core Team to explore the social organization of R and to identify factors crucial to its success.

The study presented here aims to examine why package authors participate in the R project. We use scales on work design characteristics, personal values, and types of motivation—based on theories from a general open-source software (OSS) perspective—to learn about factors and incentives that drive authors to develop R packages, as well as participate in R conferences and mailing lists.

The overwhelming majority of R packages are released under open-source licenses, thereby placing no restrictions on users and guaranteeing that these packages can become public goods (4). Although from a traditional economic point of view, it appears to make no sense to give away one's skills and efforts for free, thousands of highly skilled developers have organized into communities like the Comprehensive R Archive Network (CRAN; CRAN.R-project.org/), Bioconductor (5) (www.Bioconductor.org/), R-Forge (6) (R-Forge.R-project.org/), and GitHub (https://github.com/) to contribute code and documentation to open-source R packages distributed by these communities.

Studying software developer's motivations and determinants for participating in OSS projects is not a straightforward task. There are many internal and external factors that might potentially play a role and, hence, have to be taken into account when one wishes to explain OSS participation. Empirical findings in this research area are rather limited and partially ambiguous (7). In this study, we apply models from item response theory (IRT) and generalized linear models (GLMs) to data collected in a survey, conveyed on the popular platforms CRAN, R-Forge, and Bioconductor.

Psychological Findings on Participation in OSS Projects

In terms of internal factors that influence participation in OSS projects, psychological literature suggests to consider motivational theory, work design theory, and value theory. Motivational theory distinguishes between intrinsic and extrinsic motivation. Intrinsic motivation is the most pervasive motive for contributions to OSS (8–11). It represents the enjoyment of an activity itself and is strongly linked to an individual's perception of autonomy and competence (12). Extrinsic motivation refers to any scenario in which a person is motivated by external control. Some of the most salient extrinsic motives are monetary rewards and peer pressure. In addition, it has been found that satisfying a personal need (scratching a personal itch) (9, 13), further improvements by others (13, 14), enhancing personal reputation (7, 10, 15, 16), reciprocity and general exchange (9, 17), and social norms (8) are other extrinsic motives to be considered in OSS developments. Most researchers agree that a simple model of purely intrinsic and extrinsic motives is insufficient to capture the motivational patterns in OSS (7, 8). Instead, motivation is to be more accurately understood as a complex continuum of intrinsic, extrinsic, and internalized extrinsic motives. Motives evolve over time, as task characteristics are shifting from need-driven problem solving to mundane maintenance tasks within the community.

Significance

Over the last years, the open-source environment R has become the most popular environment for statistical computing and data analysis across many fields of research. The developer community is highly active: Thousands of packages are available in the official Comprehensive R Archive Network repository and more on developer platforms like GitHub or R-Forge. One question that has not been studied yet is as follows: why do people contribute to the R environment? What are the key motives that drive package authors? Do these developers have specific personal value structures? Are some work environments more conducive to productivity than others? This study is the first empirical study, to our knowledge, performed within the R package author community that finds answers to these questions.

Author contributions: E.H. and R.H. designed research; P.M., E.H., and K.H. performed research; P.M., K.G., R.H., A.Z., and K.H. analyzed data; and P.M. and E.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: mair@fas.harvard.edu.

[2]Deceased July 17, 2012.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1506047112/-/DCSupplemental.

The second potential influential factor for OSS contribution are work design characteristics (18, 19). Corresponding underlying traits refer to task complexity, significance of work, autonomy mastering the task, feedback from the task, etc. (20, 21). The model for work design allows organizations to assess the current state of specific task-related characteristics and, consequently, to change their design in a way that tasks become more motivating.

Third, personal values can be important for understanding contributions to OSS projects. The classic value theory in ref. 22 distinguishes 10 different values: benevolence, conformity, tradition, security, power, achievement, hedonism, stimulation, self-direction, and universalism. Oreg and Nov (23) determined the following three values to be relevant for OSS developments: self-direction, power, and universalism (24). Self-direction type values (e.g., creativity, choosing own goals, curiosity) are driven by independent thought and action. Thus, they are closely related to forms of intrinsic motivation. Power type values (e.g., social power, social recognition, authority) reflect abstract outcomes on an individual's achievements. These values do not refer to the direct outcomes of any particular action, but to the status in social structure an individual is able to derive from actions. Hence, they relate directly to forms of internalized extrinsic motivation. Universalism type values (e.g., equality, wisdom, social justice) refer to action for the welfare of all people and are derived from people's awareness of the scarcity of resources. They imply that individuals will consciously protect their own survival needs through the acceptance and just treatment of anyone outside their group (22).

Survey Design and Research Questions

Our population consists of package authors who contributed to R packages on CRAN, Bioconductor, or R-forge. This population includes package maintainers and people who received credit for contributing code and, therefore, appear in the package author list. We need to distinguish package authors clearly from users, that is, people who are just using packages or providing code snippets without being "officially" involved in a package development. Our study does not aim to generalize the results to the whole R community.

The online questionnaire for the package authors, provided as *SI Results*, included standard socio-demographic variables, as well as more specific dichotomous work-related variables such as whether respondents have a PhD degree, an education in statistics, are employed full time, work in academia, and work as statisticians.

Based on the research results described above, three lines of possible psychometric incentives are pursued: (*i*) hybrid forms of motivation, (*ii*) work design characteristics, and (*iii*) values. We investigate to which extent these factors determine the degree of the authors' participation in the R project. The following subsections describe these variables and constructs included in our study. Fig. 1 summarizes the latent structure of the psychometric scales we use and their relation to the measures for participation.

Degree of Participation. Participation in OSS projects will primarily manifest itself in the form of code contributions. As previous studies have shown, however, this is just one part of an underlying learning and information process (17). A prominent example of other forms of contribution is the active engagement in social media platforms such as mailing lists or blogs (9).

In the context of the R project, contributed code is typically conveniently organized in packages and distributed via repositories such as CRAN or Bioconductor. This fact makes packages the primary vehicle for communicating conceptual and computational tools related to R. Hence, the number of R packages (co)developed by an individual author (cf. Fig. S1) can easily be interpreted as the first, main variable of the extent of participation in the R project. As a second indicator, we use active participation in R project mailing lists (R-help, R-devel, special interest groups, ...) as an indicator for engagement in social media. Finally, as third participation indicator, we consider attending R conferences such as the annual useR! or the Directions in Statistical Computing (DSC) meetings.

Psychometric Constructs. As elaborated above, the classic distinction between intrinsic and extrinsic motivation is seen as too rigid within our context. Reinholt (25) presents a concept that distinguishes between extreme intrinsic motivation, well-internalized extrinsic motivation/moderated intrinsic motivation, and extreme extrinsic motivation. Well-internalized extrinsic motivation and moderated intrinsic motivation comprise hybrid types of intrinsic and extrinsic motivation. The corresponding scales are based on this concept of motivation because it provides a nuanced and

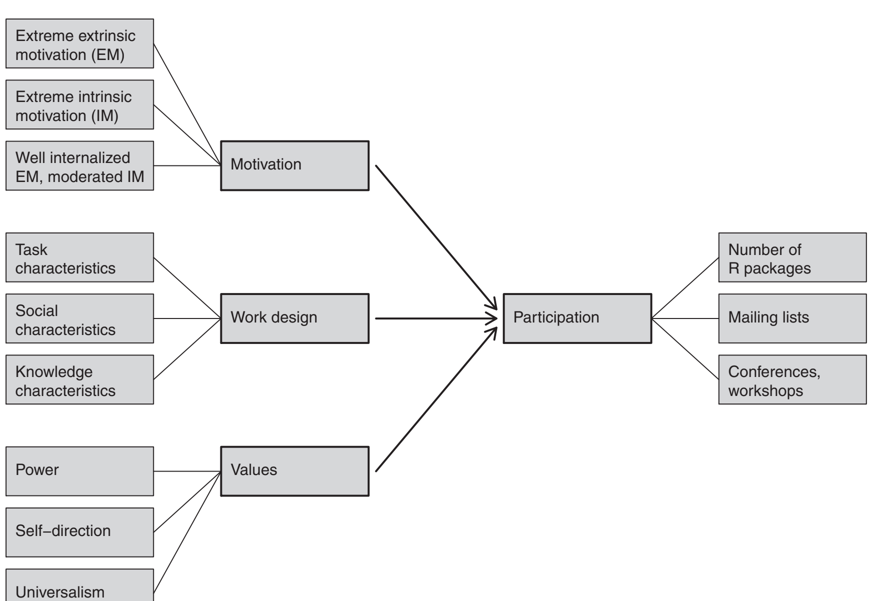

Fig. 1. Psychometric constructs. Hybrid forms of motivation (25), work design characteristics (21), and values (22) determining participation in the R project.

coherent understanding of motivational types along a continuum of motivation. This framework also accounts for potential interaction effects between intrinsic and extrinsic types of motivation. For the intrinsic and extrinsic motivation subscales, 36 items are included in our questionnaire. Each subscale (i.e., enjoyment based intrinsic motivation, self-reinforcement, obligation-based motivation, integrated regulation, identification, introjection-based regulation, external regulation) consists of four to eight items.

As suggested by previous studies (9, 10), the Work Design Questionnaire (WDQ) (21) is a prominent tool to investigate work design characteristics. This work design model captures, among others, the following three subscales: the effects of task characteristics (autonomy, task variety, task significance, task identity, feedback from job), social characteristics (received and initiated interdependence, feedback from others), and knowledge characteristics (job complexity, information processing, problem solving, skill variety, specialization). In its original form, the WDQ is composed of 77 items. Using the three subscales above reduces the questionnaire to 48 items. Note that WDQ items referring to work tasks in general were adapted to the work on R packages.

Regarding personal values, we consider 3 of the 10 values of the Schwartz value scale (self-direction, power, and universalism). All 19 items pertaining to these value subscales are included in the questionnaire.

Research Questions. Based on the theoretical extension of the concept of intrinsic and extrinsic motivation (25), we hypothesize that extreme extrinsic motivation (comprising external regulation and introjection-based regulation), extreme intrinsic motivation (stemming solely from enjoyment-based intrinsic motivation), and well-internalized extrinsic motivation/moderated intrinsic motivation (identification, obligation-based intrinsic motivation, self-reinforcement, and integrated regulation) are positively related to the participation in the R project.

Regarding work design, it is expected that task characteristics (comprising autonomy, task variety, task significance, task identity, and feedback from the job), knowledge characteristics (including job complexity, information processing, problem solving, skill variety, and specialization), and social characteristics (consisting of received and initiated interdependence and feedback from others) are positively related to participation. The more positive these characteristics are perceived, the more a package author should participate in R activities.

Finally, in line with earlier studies, it is hypothesized that the values self-direction and universalism relate positively to participation, whereas power is expected to relate negatively.

Statistical Analysis and Results

Statistical Analysis Work Flow. Our sample consists of 1,087 package authors. The statistical analysis work flow is the following: we scale each psychometric construct using a two-parameter logistic (2-PL) IRT model (26). Unidimensionality is checked using categorical principal component analysis (27), and itemfit is tested using the Q1 fit statistic (28). For the set of fitting items, the latent trait (person) parameters are estimated, which then act as predictors, in addition to demographic variables, in the subsequent GLMs. For the first participation response "number of packages," we fit a negative binomial regression, and for "participation in mailing lists" and "attending conferences," we fit logistic regressions. For each of these regression models, first a full model is considered using all potential determinants: the nine psychometric scores and all socio-demographic factors. Subsequently, a stepwise backward selection of the predictor variables in the GLM is carried out based on the Akaike information criterion (AIC) to highlight which determinants are most relevant. For the full model and the final model from stepwise selection, to account for the measurement error of latent trait scores as predictors in the GLM, we apply the simulation-extrapolation (SIMEX) approach (29). Methodological details about each statistical analysis step is given in the *SI Text*, as well as the outputs in terms of regression tables and effects plots.

Results

First, we look at the negative binomial regression with the number of packages an author has (co)authored as the response variable (Table S1). The effect plots for the final model are given in Fig. S2.

The number of packages are positively influenced by hybrid and extrinsic motivation. Work design is also an important determinant of the number of packages, with social characteristics being positively associated and task characteristics being negatively associated. Thus, the higher the intiated/received interdependence of an author and the more feeback he/she gets from the community, the more packages he/she is involved in.

Conversely, the higher a package author scores on the task dimension, the lower the number of packages (co)authored. In terms of the value scales, only power is found to be significantly associated with the number of packages showing a negative effect. On the socio-demographic side, the fact that a package author works full time and his/her field of work is statistics have a significant effect.

The results for the logistic regression model of participation in mailing lists are given in Table S2 and the effect plots are shown in Fig. S3.

Again, hybrid motivation significantly increases the probability of participation. However, extrinsic motivation has a similar absolute effect (both in terms of coefficient estimate and SE), but the effect is negative. Regarding the WDQ, social characteristics have a large positive impact and task characteristics a somewhat smaller negative impact. None of the value scale variables has a significant effect on the participation in mailing lists. For the socio-demographic predictor part, the fact that a package author works in the field of statistics leads to a significantly lower participation probability.

Finally, Table S3 presents the results of the logistic regression model for the binary response, indicating participation of package authors in R conferences and workshops. The corresponding effect plots are given in Fig. S4.

Regarding the motivational dimension, hybrid motivation is again found to be the most important determinant. Its influence is again positive. In terms of work design, social characteristics are significant with a positive impact on participation. Regarding values, universalism is significant at 5% after stepwise selection. The only significant socio-demographic variable is the occupational status: a full-time employment of a package author is a strong determinant to participate in R conferences. None of other socio-demographic variables (except, to a certain degree, statistics as the field of work that has a minor influence) has any impact on the model.

To summarize, the broad picture is very similar across all three participation responses (and corresponding models), even if the details vary to a certain degree: hybrid motivation and social characteristics are the most important determinants for higher levels of participation in the R project. The picture for extrinsic and intrinsic motivation is less clear and varies over the particular type of participation. Authors that score highly on the task characteristics scale generally participate less, whereas knowledge characteristics do not play an important role. Similarly, values are not found to be important drivers of participation as they rarely show up in the selected models. The influence of the socio-demographic variables varies across the models: full-time employment generally increases participation, whereas a job in academia somewhat lowers it. Working in statistics has a positive effect on the number of packages and participation in the conferences but a negative on participation in mailing lists. The remaining two variables (having a PhD and an education in statistics, respectively) cannot be shown to have an impact on participation in any of the models.

Discussion

This study has asked why R package authors participate in the R project for statistical computing. A survey was conducted and the data were analyzed using IRT models and, subsequently, GLMs (with SIMEX correction). In what follows, our findings are discussed in more detail and related to the literature on participation in OSS projects.

Hybrid Forms of Motivation. In line with the literature (7–9), hybrid motivation a crucial determinant for participation, whereas purely intrinsic and purely extrinsic forms of motivation are less important. These findings are reflected by our regression results and conform well with the academic life cycle. Various factors, including reputation, reciprocity, or social norms, can contribute to an internalization of extrinsic motives. On the one hand, many academics "do what they have to do." On the other hand, they select tasks they enjoy doing that can also encompass activities such as "fun coding" (8).

The influence of purely extrinsic motivation, which, in particular, includes monetary rewards (8), varies across the participation variables. In part, this may be due to a strong rooting of the R project in various academic communities. Although packages and conferences are by now regarded as scientific contributions, mailing list contributions have no (direct) impact on academic performance measures. This fact is somewhat substantiated by the positive (but not significant) influence of intrinsic motivation on contribution to mailing lists. We note that Bianchi et al. (16) found that contributions to "electronic networks of practice" are increased if the contributors perceive that this enhances their reputation (i.e., a typical extrinsic motive). Thus, participation in R mailing lists is apparently not perceived to do so. This situation might be different in the more recently established question and answer websites such as Stack Exchange, which work differently from classical mailing lists and explicitly try to capture the reputation of its contributors.

Work Design Characteristics. Social work design characteristics reflect the fact that work is performed within a broader social environment (21) where single individuals highly depend on each other. Our results show that OSS projects provide high degrees of social dependency and feedback as theoretically hypothesized in ref. 18. That social characteristics are such an important factor in our models is not too surprising, given that we are interacting in a social media dominated environment and social coding platforms are widely used (30). Psychological explanations for our results are the following: first, interaction with persons perceived as important leads to reputation (self-esteem, future job opportunities, etc.). Second, interaction with alike minded persons (i.e., interested in solving statistical problems) might be a possibility to express oneself and enjoy social inclusion.

From a perspective that goes beyond work design characteristics, social aspects include social recognition and identification. The R community seems to offer the opportunity for R developers to identify with this highly valued group and feel a sense of belonging. It can be assumed that they receive parts of their self-esteem by belonging to such a valued group (31) and are especially motivated to contribute to this group. It would be interesting to study such general social aspects of reputation gaining in a follow-up study.

Task characteristics are found to have a negative influence on participation which can be explained as follows: if the work is organized around the development of an R package as the central task (from development of code, via writing of manuals and vignettes to maintenance and bug fixing), R authors appear to do that but are less involved in the development of further packages or discussions on mailing lists. Or conversely, those authors who participate more and develop several packages, do not appear to be driven be the task of R package development as such but by the underlying knowledge characteristics involved.

Values. Our results indicate that in the context of R packages there appears only little additional direct effect of the values—other than potential indirect effects through the types of motivation. There are two notable exceptions: power is shown to have a clear negative effect on the number of packages and universalism has a clear negative effect on conference participation.

The former reflects that package authors, for whom social power, wealth, social recognition, and authority are important, produce fewer packages than their trait counterparts. The way the field of applied and computational statistics has developed over the last years, R package implementations have increased in scientific value. Thus, for a researcher, a corresponding implementation has become an academic status symbol to the effect that they refer to themselves as "R package author" even when involved in a single package only.

The latter shows that the higher a package author scores on the universalism dimension, the less likely he or she is to attend meetings. A closer look at what is meant by universalism provides an interesting interpretation of this result. According to Schwartz (22), attributes associated with universalism include the following: a world of beauty, unity with nature, protecting the environment, and inner harmony. These attributes are derived from an awareness of the scarcity of resources. Thus, universalism implies a strong environmental attitude that may be incompatible with carbon-intensive long distance travels to conferences.

Socio-Demographic Variables. Full-time employment always has a positive impact on participation; significantly for the number of packages and conference participation. This fact suggests that many contributions to the R project are made as part of the job. For mailing lists the influence is weaker but, as already argued above, such participation is typically not part of the job description. Additionally, there may also be direct effects of full-time employment on conference participation (e.g., through reimbursement of expenses).

Working in the field of statistics also has positive impact on the number of packages and conference participations but clearly negative impact on mailing list participation. Although the former is not surprising given that the R system is dedicated to statistics, the latter may not be obvious. However, statisticians will typically have other ways of asking questions related to R (e.g., colleagues within their department) and other ways of providing feedback about the corresponding statistical methods (e.g., in forms of papers, books, or lectures). However, for R authors and users coming from other domains (say, ecology, finance, or epidemiology), the R mailing lists may be a more crucial means of obtaining information related to R. This supposition overlaps with the findings in ref. 32, which showed that answers on the R mailing lists are mainly given by a few central players feeling responsible for certain topics.

Interestingly, an academic background (i.e., having a PhD or a job in academia) does not lead to more participation as hypothesized in ref. 14. In fact, it has almost no impact on any of the three response variables.

Conclusions. Our results show that growth of R-related projects is positively influenced by hybrid motivation, whereas purely intrinsic or extrinsic motives are less important. Hence, this suggests that extrinsic motives (such as monetary rewards or building reputation) can be important drivers but need to be balanced by possibilities of internalizing them. However, given the ongoing commercialization of the R ecosystem this aspect deserves reinvestigation in the future.

In conclusion, our results are important for institutions and individuals that want to stimulate growth of R developments: they must provide a work environment and corresponding incentives that foster a high amount of interdependence and feedback from others. Such collaborative research strategies also include the encouragement to work on projects with researchers outside the institution and the engagement in social coding platforms.

Materials and Methods

Sample. In total, we had 4,274 email addresses of R package authors. They were asked to fill out an online questionnaire within the following 3 wk. The survey was conducted in May 2010 using the online survey software Unipark. The platforms we used for the acquisition of the email addresses were CRAN, R-Forge, and Bioconductor. In total we sent out 4,274 emails, of which ∼200 could not successfully be delivered ("bounced"). Note that if packages had multiple authors, emails were sent out to those who provided an email address in the package description file. In addition, in the email list we used, some package authors had multiple email addresses. Therefore, the response rate below reflects a lower bound.

A total of 1,448 authors considered the questionnaire: 310 respondents quit immediately and 51 respondents scrolled through without answering. Altogether, a sample of 1,087 persons remained, which leads to a response rate of at least 27%. This response rate is in line with related OSS studies such as in refs. 10, 15, and 33. A total of 764 package authors completed the whole questionnaire without skipping any of the items. From a statistical power point of view, this sample size is sufficiently large to carry out all of our statistical analyses. The issue of possible nonresponse bias is addressed and analyzed in detail in *SI Text*. Our results are representative for R package authors who contributed to more than one package (see Fig. S5).

Reproducibility Materials. The following materials were submitted to fully reproduce the analysis in the article. The raw data are stored in Dataset S1, along with the variable descriptions (Dataset S2). The code file in Dataset S3 contains the R code for data preparation, IRT analysis, and all the GLM computations (including regression tables and effect plots) presented in *SI Text* and *SI Results*. In addition, it provides code to examine possible nonresponse bias.

ACKNOWLEDGMENTS. The authors would like to commemorate with sorrow the death of Reinhold Hatzinger. His contributions to this paper were invaluable and he will be fondly remembered by his friends and colleagues. The authors would like to thank the editor Robert Tibshirani and two anonymous referees for their careful and detailed comments.

1. R Core Team (2015) *R: A Language and Environment for Statistical Computing* (R Core Team, Vienna).
2. Ihaka R, Gentleman RC (1996) R: A language for data analysis and graphics. *J Comput Graph Stat* 5(3):299–314.
3. Fox J (2009) Aspects of the social organization and trajectory of the R project. *R J* 1(2): 5–13.
4. von Hippel E, von Krogh G (2003) Open source software and the private-collective innovation model: Issues for organization science. *Organ Sci* 14(2):209–223.
5. Gentleman RC, et al. (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol* 5(10):R80.
6. Theussl S, Zeileis A (2009) Collaborative software development using R-Forge. *R J* 1(1): 9–14.
7. Roberts JA, Il-Horn H, Sandra AS (2006) Understanding the motivations, participations and performance of open source software developers: A longitudinal study of the Apache projects. *Manage Sci* 52(7):984–999.
8. Lakhani KR, Wolf RG (2005) Why hackers do what they do: Understanding motivation and effort in free/open source software projects. *Perspectives on Free and Open Source Software*, eds Feller J, Fitzgerald B, Hissam S, Lakhani KR (MIT Press, Cambridge, MA), pp 3–22.
9. Shah SK (2006) Motivation, governance, and the viability of hybrid forms in open source software development. *Manage Sci* 52(7):1000–1014.
10. Hertel G, Niedner S, Hermann S (2003) Motivation of software developers in open source projects: An internet-based survey of contributors to the Linux kernel. *Res Policy* 32(7):1159–1177.
11. Li Y, Tan CH, Teo HH (2012) Leadership characteristics and developers' motivation in open source software development. *Inf Manage* 49(5):257–267.
12. Deci EL, Koestner R, Ryan RM (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* 125(6): 627–668, discussion 692–700.
13. Raymond E (1999) The cathedral and the bazaar. *Knowl Tech Policy* 12(3):23–49.
14. Henkel J (2006) Selective revealing in open innovation processes: The case of embedded Linux. *Res Policy* 35(7):953–969.
15. Wasko M, Faraj S (2005) Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *Manage Inf Syst Q* 29(1):35–56.
16. Bianchi AJ, Kang SM, Stewart D (2014) The organizational selection of status characteristics: Status evaluations in a open source community. *Organ Sci* 23(2):341–354.
17. Lakhani KR, von Hippel E (2003) How open source software works: Free user-to-user assistance. *Res Policy* 32(6):923–943.
18. Hertel G (2007) Motivating job design as a factor in open source governance. *J Manag Governance* 11(2):129–137.
19. Hemetsberger A, Reinhardt C (2009) Collective development in open-source communities: An activity theoretical perspective on successful online collaboration. *Organ Stud* 30(9):987–1008.
20. Hackman JR, Oldham GR (1976) Motivation through the design of work: Test of a theory. *Organ Behav Hum Perform* 16(2):250–279.
21. Morgeson FP, Humphrey SE (2006) The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *J Appl Psychol* 91(6):1321–1339.
22. Schwartz SH (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Adv Exp Soc Psychol* 25(1):1–65.
23. Oreg S, Nov O (2008) Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values. *Comput Human Behav* 24(5):2055–2073.
24. Engelhardt S, Freytag A (2013) Institutions, culture, and open source. *J Econ Behav Organ* 95:90–110.
25. Reinholt M (2006) *No More Polarization, Please! Towards a More Nuanced Perspective on Motivation in Organizations. Technical Report* (Center for Strategic Management, Copenhagen Business School, Copenhagen).
26. Birnbaum A (1968) Some latent trait models and their use in inferring an examinee's ability. *Statistical Theories of Mental Test Scores*, eds Lord FM, Novick MR (Addison-Wesley, Reading, MA), pp 395–479.
27. De Leeuw J, Mair P (2009) Gifi methods for optimal scaling in R: The package homals. *J Stat Softw* 31(4):1–21.
28. Yen W (1981) Using simulation results to choose a latent trait model. *Appl Psychol Meas* 5(2):245–262.
29. Cook JR, Stefanski LA (1994) Simulation-extrapolation estimation in parametric measurement error models. *J Am Stat Assoc* 89(428):1314–1328.
30. Dabbish L, Stuart C, Tsay J, Herbsleb J (2012) Social coding in GitHub: Tansparency and collaboration in an open software repository. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Burnham Inc Pub, Chicago), pp. 1277–1288.
31. Tajfel H, Turner JC (1986) *The social identity theory of intergroup behavior.* Psychology of Intergroup Relations, eds Worchel S, Austian WG (Nelson-Hall, Chicago), pp 7–24.
32. Bohn A, Feinerer I, Hornik K, Mair P (2011) Content-based social network analysis of mailing lists. *The R Journal* 3(1):11–18.
33. Wu CG, Gerlach JH, Young CE (2007) An empirical analysis of open source software developer' motivations and continuance intentions. *Inf Manage* 44(3):253–262.
34. Kish L (1965) Survey sampling (John Wiley & Sons, Inc., New York).
35. Armstrong JS, Overton T (1977) Estimating nonresponse bias in mail surveys. *J Mark Res* 14:396–402.
36. Rizopoulos D (2006) ltm: An R package for latent variable modeling and item response theory analyses. *J Stat Softw* 17(5):1–25.
37. Stefanski LA, Cook JR (1995) Simulation-extrapolation: The measurement error jackknife. *J Am Stat Assoc* 90(432):1247–1256.
38. Lederer W, Küchenhoff H (2006) A short introduction to the SIMEX and MCSIMEX. *R News* 6(4):26–31.