PhD student training: principles and implementation

L F Panchenko, H O Korzhov, T V Kolomiiets and M N Yenin
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, 37 Peremohy Ave., Kyiv, 03056, Ukraine

E-mail: lubov.felixovna@gmail.com

Abstract. The article deals with the problem of PhD student training. The Salzburg Principles about the improvement and quality assurance of doctoral programs at universities in every country from European University Association are analysed and materials of LERU on quality culture in Doctoral Education in Europe are used. The three focuses of the research component of doctoral program are proposed. The first focus relates to reproducible research principle. The second focus is related to the use of multivariate models of phenomena’s study and SEM methodology. The SEM methodology is mostly based on deductive logic, involves the preliminary construction of a structural model of the relationships between the variables in order to further check for consistency with the experimental data. The third focus combines qualitative and quantitative methods and the use of triangulation (data triangulation, investigation triangulation, theory triangulation etc.). The content of selected courses for doctorate students is proposed: Reproducible Research and Multivariate Methods in Scientific Research courses. The importance of courses related to the future career is demonstrated, in particular the career of academic researcher. Further development of our work is the creation of teaching and methodological support for selected course “Twitter for professional development of PhD students”.

1. Introduction

1.1. Problem setting and analysis of research
In recent years, we have seen the attention of scientists to the quality of PhD training programs. It is time to accredit such programs, and thus increase the requirements for their design, conceptualization and implementation.

The Salzburg Principles outline a set of ideals that are relevant to the improvement and quality assurance of doctoral programs at universities in every country (European University Association, 2005, 2010) [8]. The materials of LERU [15] present the quality culture in Doctoral Education in Europe.

In 12.02.2021 at the site of the Ministry of Education and Science of Ukraine [16] there were projects of standards in the following sciences: 122 Computer Science (MSc), 122 Computer Science (PhD), 142 Power Engineering (PhD), 151 Automation and computer-integrated technologies (PhD), 185 Oil and Gas Engineering and Technology (MSc), 185 Oil and Gas Engineering and Technology (PhD), 186 Publishing and Printing (PhD), 192 Construction and Civil Engineering (MSc). Thus, to date, there is neither an approved standard, nor a project of PhD standard on educational, pedagogical sciences and sociology.

The first Salzburg principle says “The core component of doctoral training is the advancement of knowledge through original research. However, doctoral training must also meet the demands of an
employment market wider than academia”. Principles also include institutional strategies and policies of Universities; the importance of diversity doctoral programs in Europe; understanding doctoral candidates as early stage researchers; crucial role of supervision and assessment; increasing mobility, and other.

The complex and successful training programme of PhD students to prepare them for research-focused careers both in academic and non-academic settings has to be based on three major foundations: research, teaching and leadership [6]. Each of these components as building blocks of the future successful career inextricably intertwines with the others creating an integrated complex. Research skills can be efficiently performed and perfected in the process of teaching which “helped students develop their professional identity as they broadened their subject knowledge and practised a disciplinary research discourse in the relatively safe environment of the classroom” [11]. Both research, especially in teams, and teaching are instrumental in the mastering and practicing soft skills, including leadership, while the latter contribute to the competent performance of the social roles of researcher and teacher.

The importance of harmonious development components of modern doctoral training (education, science, business) is emphasized by the Rector of the Igor Sikorsky Kyiv Polytechnic Institute Michael Z. Zgurovsky. Science Park “Kyiv Polytechnic” was the first experience of organizing innovation activities on the basis of combining the interests of all participants from education, science and business. The purpose of such innovative activity is commercialization of scientific research’s results and their implementation in the employment market of Ukraine and foreign countries [31]. Such harmonious development can be ensured by the creation of the “Sikorsky Challenge” business incubator and a business start-up school at Kyiv Polytechnic Institute.

In the present article, we deal with the research component of doctorate programs which should be based on the ideas of reproducible research supported by scientists of various specialties [2], [5], [4], [18], [23], [28]. In reproducible research, scientific statements are published along with the collected data and program code, so that other researchers can verify the data obtained and rely on them.

Among the possible advantages of such studies are greater confidence in the reliability of the results; improving the quality of data analysis; expanding the possibilities of further research based on existing ones; opportunities for training future researchers and their cooperation with experienced ones; introduction of standards for the submission of PhD research materials.

The relevance of this topic is evidenced by our experiment conducted in May 2019 to find sources of relevant topics on the amazon.com website and a number of mass online courses and initiatives [20].

The second focus of the research component should be aimed at training future researchers in the field of multidimensional statistics to build multidimensional models of pedagogical or sociological phenomena that they will study. Qualitative scientific research is impossible without a systematic description of the studied phenomena; multidimensionality of the investigated phenomena requires the use of multidimensional analysis methods that are capable to identify causal relationships, latent factors, etc. A promising area in the field of multidimensional applied analysis is the structural modeling or structural equation modeling, which is becoming an increasingly popular tool for researchers in the field of education, psychology and social sciences [1], [19].

As a third focus, we propose to combine qualitative and quantitative methods in research practice with the method of triangulation as a path for cross-checking the results of the study and their interpretation [7], [13], [14], [24], [25].

The research question of the article is: what is the content of training PhD students in the field of education on the subject of the research component and in which courses it is appropriate to do so? The research methodology included analysis and generalization of scientific articles on reproducible research, on research methods as well as analysis of the content of PhD programs; analysis of EUA-CDE survey data of [12] on doctoral education in European universities. Survey focuses on key aspects: doctoral candidates’ qualifications, funding, completion rate and time to completion; purposes; organization; application and admission; supervision; training and activities; quality assurance; career development; decision-making powers and strategic priorities in doctoral education and their implementation.
2. Results of study
From the results of the EUA-CDE survey [12] in 2017-2018 on doctoral education in European universities we can clearly see that universities in Europe focused on research competence training (table 1).

| How important are the following elements of doctoral training? ([12], p. 15) |
|---------------------------------------------------------------|
| Specific research competencies | Generic academic competencies | Knowledge valorization | Teaching competencies | Management and leadership competencies |
| Extremely important | 75 | 35 | 11 | 11 | 6 |
| Important | 22 | 47 | 36 | 34 | 31 |
| Somewhat important | 2 | 14 | 38 | 38 | 40 |
| Somewhat unimportant | 3 | 3 | 11 | 13 | 20 |
| Unimportant | 0 | 0 | 4 | 3 | 3 |

We calculated a weighted average and visualized this data (figure 1).

![Figure 1. Importance of skill training evaluation on doctoral training (weighted average).](image)

Doctoral education is dominated by training activities focused on specific research competencies: 97%, with 75% finding it “extremely important” and 22% “important” (weighted average is 1,71). They included advanced methods, up-to-date data knowledge, new techniques. Generic academic competencies – grant writing, publishing, ethics – were chosen by 82% universities (35% finding it “extremely important” and 47% “important”), the weighted average is 1,31. Knowledge valorization (intellectual property rights, entrepreneurship, product development) is a focus for 47% of universities, who find it either “extremely important” (11%) or “important” (36%), weighted average is 0,39; and management and leadership competencies were chosen by 37% of universities, finding it either “extremely important” (6%) or “important” (31%), weighted average is 0,17. 45% of universities note the importance of pedagogical competences, finding it either “extremely important” (11%) or “important” (34%), weighted average is 0,38.

Consider the authors' proposed three focuses of research component: Reproducible Research, Multivariate Modeling and Mixed Methods.
2.1. Reproducible research
During research we studied the reproducible research content in PhD program. OSF initiative [4] involves ungraduated and graduated courses on these topics for different fields of science [2], [18], [23], [28]. We selected courses for PhD programs (table 2) and presented their distribution by the fields (figure 2).

| Course name                                      | University, Institution                                      | Author          | Fields                        |
|--------------------------------------------------|--------------------------------------------------------------|-----------------|-------------------------------|
| Scientific Computing and Visualization          | Doctoral school Frontières du Vivant, Paris                  | K. Hinsen, 2015 | Interdisciplinary             |
| Meta-Analysis and Open Science                  | Summer schools. The Swedish Institute for Disability Research | H. Danielsson, 2017 | Interdisciplinary             |
| Ethics, Open Science, and Reproducible Research | PhD course Stockholm University                              | G. Nilsonne, 2018 | Interdisciplinary Social Science |
| Rigorous & Reproducible Research Practices      | Experimental Psychology, Ph.D. program, Tufts University     | H. L. Urry, 2018 | Psychology and Neuroscience   |
| Open Science and Reproducible Research          | Karolinska Institutet                                         | G. Nilsonne, 2018 | Medicine and biomedical sciences |

As we see in the figure 2, there are 2 interdisciplinary courses, 8 in the category “Psychology and Neuroscience”, 2 in “Social Science”, 2 in the category “Medicine and biomedical sciences”. There are no such courses in the field of pedagogy.

![Figure 2](image_url)  
*Figure 2. Reproducible research courses in PhD programs (according to [4]).*

We analysed also the content of mass online courses on relevant topics, including the course “Transparent and Open Social Science Research” offered on the Future Learn platform, courses “Reproducible research” and “Communicating Data Science Results” offered on the platform Coursera. The topics and subtopics of the courses, software and services used in them were analysed. We collected generalized data in the table 3 [20].

The Transparent and Open Social Science Research course [29] can be seen as an introduction to transparent and open research; it proposes a plan for preliminary analysis of research and re-registration data; the state and future of open science is clarified.
Table 3. MOOCs for Reproducible Research.

| Course name                        | Platform       | Part of specialization      | Software                                      |
|------------------------------------|----------------|----------------------------|-----------------------------------------------|
| Transparent and Open Social Science Research | Future Learn | –                          | OSR, Stata, R                                 |
| Reproducible research Communicating Data Science Results | Coursera       | Data Science               | R, R Studio                                   |
|                                    |                | Data Science at Scale      | Elastic Map Reduce, Pig, Amazon Cloud Service |

The Reproducible research course [22] is part of the Data Science specialization, which includes 10 courses and is offered by Johns Hopkins University on the Coursera platform. The course focuses on concepts and tools that allow you to reproduce modern data analysis. R-Markdown and knitr tools are discussed in detail as tools for developing reproducible research reports. The Checklist for evidence-based data analysis is provided.

The latest course in Communicating Data Science Results states that reproducibility is important not only for scientists: data analysts must be able to share data, explain it, and protect their methods and data from unauthorized access. The course uses Amazon's cloud services, Elastic Map Reduce, and Pig to perform big data analysis. Cloud computing is seen as a friendly environment for reproducible research with powerful capabilities.

Analysis of scientific sources and content of PhD programs and mass online courses on relevant topics showed the need to include the topic of reproducible research in PhD student training in Ukraine in the form of a special selective course. Such a special course may consist of the following three blocks: 1) introduction to reproducible research; 2) tools and services for reproducible research; 3) environment R as a basis for reproducible research.

The first block plans to introduce Ukrainian PhD students to key terms related to open science, including re-registration, registered reports, p-hacking, HARKing, open articles [27].

- Re-registration is the practice of digitally registering an in-depth data analysis plan before data collection.
- Filed reports are scientific articles whose potential for publication is assessed through peer review and editorial decisions.
- P-hacking is the use of data mining methods to identify patterns in data that are “statistically significant”, based on the use of p-values, but for which no previous causal hypotheses have been developed.
- HARKing is the practice of formulating hypotheses by scientists after obtaining scientific results (by the first letters of English words – Hypothesizing After Results are Known).
- Open articles (OA). There are two types of open articles. Green OAs refers to the practice of creating a copy of a published journal article that is freely available in the repository or on the researcher's personal website. Gold OA is the availability of journal articles on a journal's website, sometimes at additional processing costs.

Within the framework of the second block of the special course the software of information technologies for use in realization of transparent research is studied: Swirl, Dataverse, Git Version Control, Zotero. The third unit focuses on the use of the R environment and its markdown and knitr packages as means to develop reproducible research reports.

As areas of further research, we consider the development of methodological support for a special course on reproducible research on the technology of blended learning.

2.2. Multivariate methods

Recently, there is a growing need to train future researchers in the field of multidimensional statistics. Future specialists should study, understand and explain the phenomena and processes that occur,
investigate their relationship, the influence of certain factors, predict development, build multifactorial and multidimensional models of such phenomena.

In our opinion, the selected course “Methods of multidimensional analysis” offered to doctorate students of pedagogical specialties may play an important role in the formation of the research culture of future professionals. The study of master's programs and training programs for doctors of philosophy in the field of social and behavioral sciences of the world's leading universities has shown the mandatory presence of such a course for future researchers.

Consider the content and structure of such a course “Methods of Multidimensional Analysis”, which is offered to PhD students in the Social Sciences as selective course. The list of topics studied includes the following:

- Introduction to multidimensional statistics and its use in research in the social sciences;
- One-factor and multifactor one-dimensional analysis of variance;
- Multidimensional analysis of variance and its use in sociology research;
- Discriminant analysis and its varieties;
- Use of cluster analysis in sociology research;
- Factor analysis and features of its use in sociology;
- Use of multidimensional scaling and collaborative analysis in sociological and marketing research;
- Modeling by structural equations as a means of constructing and verifying multidimensional models of social reality.

We have developed educational and methodological support of the course, which consists of Power Point lectures on each topic, laboratory class on methods of multidimensional analysis, examples of data files used in the classroom and independent work. An important part of the course is the discussion of the relationship between the various methods used in it.

As software support for the course, we proposed the use of three environments, namely: spreadsheets, SPSS statistical environment (or PSPP), and free environment for statistical modeling R. Spreadsheets that implement only some methods of multidimensional statistics (one-factor and two-factor analysis of variance). It can also be used to demonstrate and perform individual calculations in the study of complex methods, such as calculations in the analysis of variance, comparing different approaches to calculating distances between objects, cluster analysis, construction of spatial maps by given coordinates in multidimensional scaling, etc.

Among the tools of statistical data analysis, the free environment R has recently received a lot of attention. It includes descriptive statistics functions, tools for intelligence data analysis, analysis of variance, methods of multidimensional analysis, neural networks, decision trees, and so on. R functionality can be extended through packages that are available on the CRAN website.

The R environment is evolving with the participation of the entire world community and is used in virtually all mass online data analysis courses, including many specializations of the Data Science Coursera projects and in academic research. Methodological aspects of the use of R were considered in the works of John Fox and Robert Andersen [9], and others. An important feature of the R environment is the ability to create dynamic reports based on the data being processed (R Markdown). Such features contribute to the conduct of reproductive research conducted by the scientific community around the world [10].

Our experience shows that the proposed content and structure of the course and its program support can be used to prepare graduates and future doctors of philosophy in other humanity fields, particularly in the field of pedagogy.

2.3. Combining qualitative and quantitative methods in research practice and use of triangulation
The third focus in PhD training research component is the combination of qualitative and quantitative methods in research practice and triangulation use [7], [13], [14], [24], [25]. Norman K. Denzin identifies four basic types of triangulation [7]: data triangulation, investigator triangulation, theory
triangulation, and methodological triangulation. Data triangulation includes time, space, and persons. Investigator triangulation involves multiple researchers in an investigation. Theory triangulation assumes using more than one theoretical scheme in the interpretation of the phenomenon. Methodological triangulation implies using several methods of gathering data (interviews, questionnaires, observations, experiment, expert analysis and documents analysis).

Note also the special issue “Mixed Methods and Triangulation in History Education Research” recently published by the “International Journal of Historical Learning, Teaching and Research” (https://doi.org/10.18546/HERJ.16.1.01).

In the table 4 four periods of history of doctorate education are presented in [17] (p. 7). As we see in the modern period the focus is on the research and research based training and academics and professional school’s fields.

| Period                          | Purposes of Doctoral Education | Subjects Studied                        | Type of Examination                                      |
|---------------------------------|--------------------------------|----------------------------------------|----------------------------------------------------------|
| Medieval Europe (1200-1535)     | To foster communities of scholars teaching and writing textbooks | Medicine, law, theology                  | Oral (teacher wrote thesis, student had to defend or oppose it) |
| Reformation Europe (1535-1750)  | To train priests and church administration | Theology, low                          | Examination by board of professors in all subjects (no specialization) |
| Enlightenment Europe (1750-1865)| To create new knowledge; to train future professors | Philosophy, humanities, natural sciences | Written (student wrote thesis, faculty took role of opponents) |
| Era of modern research university (1865-present) | To foster settings for research and research based training | The subject matter of most academic fields as well as that of professional schools | Written and oral (student writes and defends dissertation) |

In search of strategic priorities of doctorate education, we turned to the EUA-CDE survey data [12] (table 5). Universities rated the importance of the following parameters: funding of doctoral education; research ethics; attracting doctoral candidates from abroad; career development of doctoral candidates; gender equality; Open Access / Open Science; health /wellbeing of doctoral candidates; increasing the number of doctoral candidates; university-business cooperation within doctoral education; societal engagement of doctoral candidates.

We calculated a weighted average and visualized the data (figure 3).

Therefore, we can clearly see that universities in Europe mainly focused on strategy priorities: funding of doctoral education: weighted average 2,68, research ethics: weighted average 2,64, career development: weighted average 2,33, gender equality: weighted average 2,22, open access / open science: weighted average 2,21 (figure 3).

It is interesting to compare strategic priorities based on data from a survey of European universities [12] with the content of scientific reports of All Ukrainian Science and Practice Conference 2017 “Training of Doctors of Philosophy in the minds of reforming high education” [30]. We presented the results of content analysis in the table 6.
Table 5. Strategic priorities in doctoral education (according to [12], p. 30)).

|                               | Not at all (%) | To a small extent (%) | To some extent (%) | To a great extent (%) |
|-------------------------------|----------------|-----------------------|--------------------|-----------------------|
| Funding of doctoral education| 2              | 2                     | 22                 | 74                    |
| Research ethics               | 1              | 4                     | 25                 | 70                    |
| Attracting doctoral candidates from abroad | 3           | 10                    | 26                 | 61                    |
| Career development of doctoral candidates | 1           | 12                    | 40                 | 47                    |
| Gender equality               | 9              | 11                    | 29                 | 51                    |
| Open Access / Open Science    | 4              | 15                    | 37                 | 44                    |
| Health /wellbeing of doctoral candidates | 6          | 16                    | 37                 | 41                    |
| Increasing the number of doctoral candidates | 14         | 11                    | 32                 | 43                    |
| University-business cooperation within doctoral education | 5          | 19                    | 48                 | 28                    |
| Societal engagement of doctoral candidate | 6           | 28                    | 46                 | 20                    |

Figure 3. Strategic priorities in doctoral education.

We can see in the works of Ukrainian researchers, little attention is still paid to the issues of career, grants and funding, Open Science, the well-being of doctoral students, and societal engagement of doctoral candidates.

3. Conclusion
Research component is an important part of PhD training. Some focuses for PhD student training in Ukraine in the research field were proposed: reproducible research, multivariate methods and mixed methods, as well as content of relevant selected courses.

The characteristics of future Doctors of Philosophy according to the Salzburg Principles are: “Advanced knowledge of the discipline; Capacity for Research; Proficiency in English; Technical and Contextual Intelligence; Transferable Skills”. To help prepare PhD candidates for an academic career several courses on skills are offered in PhD programs, such as Art of Presentation; Advanced Academic Writing; Academic Career Building; Social Networks for Academics.
Table 6. Main aspects of the scientific reports in materials of the All Ukrainian scientific-practical conference 2017 “Training of Doctors of Philosophy (PhD) in the context of higher education reform”.

| Aspects                                                      | Number of reports | % of reports |
|--------------------------------------------------------------|-------------------|--------------|
| General issues of training and improvement                   | 12                | 25           |
| Academic integrity                                           | 8                 | 17           |
| Foreign experience, European experience                      | 5                 | 10           |
| Mobility                                                     | 4                 | 8            |
| Publications and scientific journals                         | 4                 | 8            |
| Doctoral training experience of a certain university          | 4                 | 8            |
| Competences                                                  | 3                 | 6            |
| Funding, projects, grants                                    | 2                 | 4            |
| Research training                                            | 2                 | 4            |
| Career                                                       | 1                 | 2            |
| Leadership                                                   | 1                 | 2            |
| Inclusion                                                    | 1                 | 2            |
| Communication                                                | 1                 | 2            |

A comparison of strategic priorities in a survey of doctoral programs in Europe [12] and in the materials of a conference on training PhD in Ukraine [30] shows a general interest in academic integrity, mobility and insufficient attention of Ukrainian scientists to finding funding, open science, leadership, the well-being of doctoral students, societal engagement of doctoral candidates, and scientific communication.

Further development of our work in this direction is the creation of teaching and methodological support for the selected course “Twitter for Professional Development of PhD students” [21] in the field of pedagogy and social sciences.

References

[1] Badri M, Alnuaimi A, Yang G, Rashidi A A and Sumaiti R A 2017 Structural Equation Model of Determinants of the Perceived Impact of Teachers’ Professional Development – The Abu Dhabi Application SAGE open 7 URL https://doi.org/10.1177/2158244017702198
[2] Ball R, Bowman S, Christensen G, Frank M C, Funder D, Hinsen K, Janz N, Medeiros N, Miguel E, Nosek B A, Poldrack R, Vazire S, Schönbrodt F, Wagenmakers E J, Campbell L, Uhlmann E L, Schweinsberg M, Bearden M, Zelenski J M, Simons D J, Nilsonne G, Kane M J, Hilgard J, Howell J, Nuijten M B, Sakaluk J K, Damian R I, Mackinnon S, Danielsson H, Roberts B, Urry H L, Corker K S, Machery E, Maximino C, Mellor D T, Grahe J, Parker Y H, Moore D A, Azevedo F, Hartwright C, Le B, Strand J, Salholz-Hillel M, Orben A, Syed M, Herrmann A P, Vatakis A, Rahal RM, Schwab S and Umpierre D 2021 Course syllabi for open and reproducible methods URL https://osf.io/vkhbt
[3] Berkeley Initiative for Transparency in the Social Sciences 2021 Berkeley Initiative for Transparency in the Social Sciences URL https://www.bitss.org/
[4] Center for Open Science 2021 OSF URL http://osf.io
[5] Christensen G S, Freese J and Miguel E 2019 Transparent and Reproducible Social Science Research: How to Do Open Science (Oakland: University of California Press) p 272
[6] Conn V S, Zerwic J, Rawl S, Wyman J F, Larson J L, Anderson C M, Fahrenwald N L, Benefield L E, Cohen M Z, Smith C E, Topp R and Markis N E 2014 Strategies for a successful PhD program: Words of wisdom from the WJNR editorial board Western Journal of Nursing Research 36 6–30 URL https://doi.org/10.1177%2F0193945913492144
gapminder/15-knitr-markdown/

[27] Toelch U and Ostwald D 2018 Digital open science – Teaching digital tools for reproducible and transparent research PLoS Biology 16 e2006022 URL https://doi.org/10.1371/journal.pbio.2006022

[28] Tufts University 2020 Graduate Program in Experimental Psychology URL http://web.archive.org/web/20201117064738/https://ase.tufts.edu/psychology/graduate/

[29] University of California, Berkeley and Center for Effective Global Action (CEGA) 2020 Transparent and Open Social Science Research FutureLearn URL https://www.futurelearn.com/courses/open-social-science-research

[30] Zaporizkyi nationalnyi universytet 2017 Pidhotovka doktoriv filosofii (PhD) v umovakh reformuvannya vyshchoi osvity: Materialy Vseukrainskoi naukovo-praktychnoi konferentsii 5-6 zhovtnia 2017 r. (Training of doctors in philosophy (PhD) in the context of higher education reform: Proc. All Ukr. Conf. (Zaporizhzhya: Zaporizhzhya National University) p 216 URL http://phd.znu.edu.ua/page/img/Tezy_PHD_ZNU.pdf

[31] Zghurovskyi M 2018 “Ostriv proryvu” nauky ta innovatsii (The island of breakthrough science and innovation) Holos Ukrainy 10 8 URL http://www.golos.com.ua/article/300332