First Approach to Purposeful Sampling for Determining Key Factors on Outcome Bias

Dr. Mariana Tafur-Arciniegas, Universidad de los Andes

Mariana Tafur-Arciniegas is an assistant professor in the School of Education at University of Los Andes, Bogota-Colombia. She is a Ph.D. in Engineering Education from Purdue University, a M.S. in Education and a B.S. in Electrical Engineering from University of Los Andes. She is a 2010 Fulbright Fellow. Her research interests include engineering skills development, STEM for non-engineers adults, motivation in STEM to close the technology literacy gap, STEM formative assessment, and Mixed-Methods design.

Mr. Andres Felipe Lara Contreras, Universidad de los Andes

Andres Lara is an undergraduate student in Chemical Engineering at Universidad de Los Andes, Bogotá, Colombia, Class of 2017-20. His research interests include engineering learning and approach to technological challenges, nanostructured materials, and self-sustainable energy sources.
First Approach to Purpose Sampling For Determining Key Factors on the Outcome Bias

Abstract

It is common, in educational research studies, to select purposeful population samples. Independent of the selected method, there is an associated bias to the sampling process because of the subjectivity of the person or group of people who accomplish the task. This paper shows the pilot of a protocol development for observing how people approach purposeful sampling using a real research problem. Engineers and non-engineers were observed completing the task and were monitored for collecting the differences between both points of view, and for identifying possible bias in the purposeful sample.

Five stages were conducted for piloting the protocol for gathering information in how purposeful sampling is traditionally performed. Two participants with knowledge in educational research performed the activity for the first iteration of the pilot. In addition, four participants (two engineer students and two non-engineer students) performed a second pilot using an improved protocol. Finally, a panel of experts was asked to review the process for a final protocol. For the activity designed, participants were asked to review 118 profiles of people with diverse academic and social backgrounds. The goal was to choose 3 profiles for each of four categories, aiming to identify those participants who were the best representatives of each of the categories: 1. Engineers with a low level of Lifelong Learning (ELL), 2. Non-Engineer with a low level of Lifelong Learning (nELL); 3. Engineers with a high level of Lifelong Learning (EHL), and 4. Non-Engineer with a high level of Lifelong Learning (nEHL). The time for the assignment was limited to 90 minutes, and a think-aloud protocol was followed for data collection.

The iterative process design for the protocol allowed to improve the resources, time management, and activity logistics for the Purposeful Sampling Activity. The time for the whole activity was calibrated and defined as 90 minutes. The time alert was changed from 10 to 15 minutes, and 10 minutes were defined to be given to those individuals who asked for more time. After participants’ feedback, new resources such as pens, sticky notes, eraser, water, and digital database with the information were included. The activity allowed to identify bias in sample selection due to lack of usage of the complete data set. Likewise, each expert defined different criteria for selection, setting diverse start points. This bias was induced by variables such as age, undergraduate background, expertise in a specific field of study and degree of development of specific skills through the professional life.
In conclusion, this study showed the design of a protocol for collecting information about how non-intended bias was present in a purposeful sampling. This analysis may guide the process of purposeful selection of samples for qualitative research and provides a tool for measuring the bias reduction between traditional and statistical purposeful selection of information-rich cases.

Introduction

In different fields of research, the study of human characteristics is required to understand the way people interact with others and with the environment. Specifically, in human sciences, it is very common to use purposeful population sampling to select cases which may be of particular interest of researchers according to a certain observable phenomenon [1]. Particularly in education, individuals’ learning profile can be useful for understanding how people learn and how lifelong learning may facilitate the constant update for technological literacy. Cheryl Livneh [2] has presented a study which allowed to understand the lifelong learning characteristics of individuals associated with human service professions. This instrument, applied to 253 individuals in human service professions, allowed to determine representative factors of lifelong learning related to time spent on learning activities [2]. The instrument takes into account seven different factors: professional growth through learning, self-motivated achievement, educability, readiness for change, causation for learning participation, familiar educational background, and future orientation. Although this instrument facilitates the process of identifying individuals’ learning profile, it is important to analyze which factor is more relevant or how the relationship between these factors may provide information-rich cases for purposeful sampling, necessary for qualitative studies [3].

Tafur Arciniegas [4, 5] has implemented the instrument as part of a study related to how STEM and non-STEM related adults approach to technological challenges. As result of this research, data has been collected related to 118 profiles which follow different lifelong learning characteristics. In addition, qualitative data has been collected related to the familiar, professional, technological and educational background of each individual, compiling data for analyzing the relationship between levels of lifelong learning and STEM background. In the study, Tafur Arciniegas gathered and organized the information in tables and individual profiles in order to offer a consolidated tool for providing the information needed for selecting a purposeful sample based on these 118 profiles [4].

During the selection of purposeful sampling, bias is associated with qualitative data processing, and the magnitude of this bias depends on the used method [1]. For instance, bias in technology literacy research can be associated with factors as age due to stereotypes related to age-based ideologies [6]. Other bias may be associated with preconceptions of the researcher who does the selection, even if it follows or not a previously defined method. For example, selection or decisions made by an investigator could be strongly affected by anxiety or stressful events, as
well as other kinds of stimuli from the environment, the memories, emotions or disorders [7].

Another source of bias can be caused by omission, which is related to the impact of omitting information and the consequence obtained in comparison to when that information is not omitted [8].

In order to reduce the bias of the selection of a purposeful sampling, Tafur Arciniegas [4] designed a statistical procedure for selecting an information-rich sample, aiming to reduce the bias associated with purposeful sampling. In the study, the selection led to a qualitative observation of technological literacy, its variation between individuals with different levels of lifelong learning and STEM backgrounds, and their changes in technological literacy over time [9, 10]. This selection allowed to observe different levels of technological literacy and find relations between these levels and individuals’ backgrounds [5].

Although probabilistic sampling is commonly used to minimizing the potential bias associated to the selection [1], Tafur Arciniegas’s study did not measure the bias reduction due to the statistical procedure [4, 5]. The objective of this study is to create and validate a protocol for systematically gather data on how purposeful sampling is performed traditionally. This will allow measuring the bias difference between a purposefully selected sample using a probabilistic approach [4] and a traditional approach (manual selection performed by expert researchers). For the protocol design, the iterative method was used allowing the co-evolution of problem-solution. That means that the results could be evaluated periodically and be reproduced for confirming the validity of the methodology [11]. In addition, a ‘think-aloud’ methodology was used, due to its potential for collecting cognitive-processing data. This methodology is currently used in areas such as psychology, helping to evaluate cognitive competence in active tasks [12]. This study presents the final protocol for gathering data on how purposeful sampling is traditionally performed. This is part of a technology literacy research project lead by Tafur Arciniegas.

**Methods**

**Instrument**

The collected data during the doctoral research of Tafur Arciniegas was provided as this study is the extension of her work. It consists in 118 profiles built from interviews and surveys which provided information related to the following sections: population characteristics, familial, educational and professional background, technical and technological development of the individual, according to the factors established by Livneh [2] and explained in the introductory section of this document.

The data was organized using a Microsoft Excel® worksheet using the same described sections. The data was codified and anonymized to avoid any possible relationship between the
information and a specific person. Also, the information was displayed in 118 profile cards using two different formats:

- Table: All the information was contained in one table, organized in the described sections. The distribution was organized to give no priority to any specific section according to the purpose of the exercise.

- Text: All the information was displayed in a text template filled with the corresponding information in each space. The template was designed to show all the information in a progressive narrative, starting with the personal and familiar context and going through the educational and professional background to finally show the affinity of the individual to specific learning practices.

Study design

This study includes quantitative data resulting from participants’ selection of 12 profiles among the 118 total cards, and qualitative information resulting from the think-aloud activity and the open-ended interview conducted after the Purposeful Sampling Activity. Due to the parallel nature mixed data collection, this is an embedded mixed method study with qualitative and quantitative data gathered at the same time, using the same sample [13, 14]. The information was triangulated by integrating the purposeful selection and the interviews and think-aloud information for discussion and interpretation of reports [15]. For qualitative analysis, a case method analysis was performed, and for quantitative analysis, descriptive statistics were used [14].

Participants

This study comprised five stages, and data were collected during three of them: During the preliminary exploration, two undergraduate students, who had a minor in education, were invited to perform the Purposeful Sampling Activity. For the purpose of this study, they are called Simon and Sarah. During the execution stage, four senior undergraduate students with no required background in education were invited. Two of those students were studying engineering, two were studying other majors. For the purpose of this work they are called Samuel and Raphael (School of Engineering), and Daniel and Jenny (non-School of Engineering). Finally, during the external revision, eight individuals were invited to form a panel of experts. This group was comprised of undergraduate, master, and Ph.D. students, and professors. All of them had a background in education and were involved with STEM (science, technology, engineering, and math) education to a certain degree.
Design

An iterative process was conducted for designing the protocol for gathering data on how purposeful sampling is traditionally performed. Five stages were conducted in order to address protocol issues and adapt the initial version according to participants’ feedback. Each stage was conducted as described below:

1. Preliminary exploration

During this stage, Simon and Sarah were given the task to analyze the 118 profiles in the table or text formats depending on their preferences. The final objective was to obtain 12 different profiles to be classified according to the following categories (3 profiles in each category): 1. Engineer with high level of lifelong learning (EHL). 2. Engineer with a low level of lifelong learning (ELL). 3. Non-engineer with a high level of lifelong learning (nEHL). 4. Non-engineer with a low level of lifelong learning (nELL). The participants had 60 minutes to complete the whole task and had unlimited paper sheets, pencil and color pens. They were told each 10 minutes the remaining time to allow them to keep track of it. The study was recorded in audio and video, with previous authorization of the participants. During the whole activity, the participant had to perform a think-aloud exercise, in which they had to say everything that came to their mind.

2. Internal revision:

With the results obtained during the first phase of the study, the researchers analyzed the results and suggestions of the participants to improve the quality of materials and resources available to complete the task. The main assessed characteristics were stress, a disposition to complete the task, availability of resources, comfortability of the participant, time availability, and data format accuracy.

3. Execution:

Samuel, Raphael, Daniel, and Jenny were called to participate in the study and, with previous authorization, they were asked to complete the same task as described in the first stage with recorded audio and video. They had 90 minutes to complete the task, with the availability of the same materials plus sticky notes, markers, sharpie, and a bottle of water. No questions were allowed after the Purposeful Sampling Activity had started. A form was used for them to record the selection within a table. There was emphasized that the activity had to be completed by filling a table with the 12 asked profiles. Also, the think-aloud exercise was required. Although the activity had to be finished for individuals to be part of the study, participation was voluntary, and
they were able to drop-out the study as desired; if they wanted to remain in the study, they had to finish the task within the 90 minutes.

4. External revision:

After completed and processed the results of the previous stage, the panel of experts in education was called to be part of this study. During a presentation, all the obtained results during the previous stages were displayed with procedure and considerations included. The panel was asked to give feedback about the process, results, expectations and appropriate assessment during the study.

5. Final adjustment:

Once finished the external revision, researchers met and determined the changes to be implemented in order to finish the pilot stage of the study and determine the exact protocol to be implemented in order to assess how purposeful sampling is traditionally performed.

Results and Discussion

During the first phase of the study, the objective was to determine the accuracy of the proposed method in order to give the participant enough time to achieve the task. None of the participants could complete the table arguing lack of time to complete the review of all the presented data. According to Sarah and Simon, the stress caused by the remaining time did not allow them to think clearly. Also, a lack of materials was noticed by Sarah, who asked for an eraser, sticky notes, and color markers. This could be caused by lack of practice in summarizing information and doing quick notes taking into account the big amount of data which was required to be processed. As reported by Mogg, Bradley & Hallowell [7], the results were altered by external and/or internal stimuli taking into account the time, materials and amount of information. Also, as both of them had an educational background, they could be omitting information when processing data to “increase” the speed and achieve the task [8].

Simon, in contrast, reported misunderstanding in the task to be completed as the instructions were not completely clear according to him. He was asking for additional information, for example, a specific criterion for selecting and completing the task; however, that was the main objective of the activity. Special or experienced guidance could alter the way people understand specific dynamics [16], inducing undesirably to bias. They should determine and specify the criteria used to classify and select the 12 asked profiles. Also, he had a serious problem to adapt to the think-aloud exercise as he said this reduces his ability to process data faster because he was aware of the camera and it called his attention.
Sarah argued problems with the distribution of the table in the summary cards for each profile. She found missing information and some spelling mistakes which distracted her taking time of the study. Also, she asked for water once the exercise was finished taking into account she was speaking (in her case fluently and naturally) for about 60 minutes. Specifically, Sarah and Simon’s profile selections were not considered to be presented as part of the quantitative data, because the form was not fully filled and the explanation about the selection was not given to be analyzed.

This gave the authors an idea about the accuracy of the method to process and classify big amounts of information. Even when there were only two replicates during the first stage of the study, both of the participants reported similar data and valid arguments, giving validity to the experiment itself according to Ericsson and Simon’s criteria [12] to verbal report validity. In general, during the second stage of the process, the stability of the protocol was considered successful because, even if their final results were not complete, the participants could select a general algorithm to extract specific information of the summary cards. In this stage of the process, the authors decided to increase the time of the task up to 90 minutes and a strict protocol was built to reduce any misunderstanding or lack of information. There was built an explanation speech to be given to the participant before signing the informed consent. After that, further details were given in order to explain the purpose of sample selection, and any additional resource they could have available during the whole exercise.

To reduce the stress of the participants, the reminding of time was decided to be every 15 minutes. The camera location was discussed and after space consideration, it was decided to be placed discretely in the available space to reduce the impact it could have during the task. All the possible resources were acquired to be provided above the table so the participants could decide since the beginning of the study which would they use. It was decided that no cell phones or other distractors were allowed inside the space so they could focus on the task. It was expected that with the mentioned modifications, factors as stress, anxiety, time pressure and external phenomena, which could induce bias, would be reduced.

The third phase of the experiment (Execution) was completed successfully with no bigger problems than a couple of markers with no ink after two of the activities. Table 1 illustrates the results of the participants of the third stage of the study. All of them could fill the form with the 12 asked profiles. All four of the participants of this stage reported stress during the last 30 minutes of the activity, as they had processed about half of the available data. All of them picked the table format of the summary cards as they assumed from the beginning that the information was clearer. None of them found big issues during the use of the table format. None of them asked for additional resources and all used the available water bottle for them.
Table 1. Results obtained during the third phase of the study (Execution) by each participant.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Samuel</th>
<th>Raphael</th>
<th>Daniel</th>
<th>Jenny</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHL</td>
<td>105</td>
<td>63</td>
<td>87</td>
<td>8</td>
</tr>
<tr>
<td>EHL</td>
<td>51</td>
<td>75</td>
<td>103</td>
<td>100</td>
</tr>
<tr>
<td>EHL</td>
<td>117</td>
<td>95</td>
<td>118</td>
<td>103</td>
</tr>
<tr>
<td>ELL</td>
<td>60</td>
<td>8</td>
<td>69</td>
<td>63</td>
</tr>
<tr>
<td>ELL</td>
<td>75</td>
<td>60</td>
<td>115</td>
<td>69</td>
</tr>
<tr>
<td>ELL</td>
<td>104</td>
<td>109</td>
<td>67</td>
<td>105</td>
</tr>
<tr>
<td>nEHL</td>
<td>11</td>
<td>84</td>
<td>80</td>
<td>87</td>
</tr>
<tr>
<td>nEHL</td>
<td>42</td>
<td>97</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>nEHL</td>
<td>84</td>
<td>117</td>
<td>70</td>
<td>85</td>
</tr>
<tr>
<td>nELL</td>
<td>4</td>
<td>58</td>
<td>92</td>
<td>2</td>
</tr>
<tr>
<td>nELL</td>
<td>73</td>
<td>78</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>nELL</td>
<td>115</td>
<td>90</td>
<td>2</td>
<td>90</td>
</tr>
</tbody>
</table>

Samuel had trouble at the beginning with the data processing as he could not define a clear strategy to process big amounts of information. The think-aloud exercise was easy for him as he explained that it is a common habit in problem-solving situations among the School of Engineering. Regardless of that discomfort, he was asked to keep doing the exercise no more than 4 times during the 90 minutes of the study. Samuel also argued lack of time to fill the form when an explanation for his selection of the 12 profiles was asked during the interview.

Raphael had no trouble with the think-aloud exercise also arguing a common practice among engineering students. His technique gave him almost no time at the end to fill the form as he did his selection in paper sheets. 5 minutes extra were given to confirm if this time would be enough to complete the form filling. Even if the form was filled within the 5 extra minutes, when processing the obtained results, he picked an engineer as a nEHL even when his criteria to divide the group was undergraduate in engineering or not. After completing the task, he said he was confused because had many paper sheets with information on every single studied profile and at the end, it was not easy to distinguish between engineers and non-engineers. To reduce this distraction, he asked to tell to future participants not to make notes about every single profile.

Daniel by his side completed the task in about 70 minutes and the last 20 were allocated for filling the form and confirming his selection. It shows no extra time was required to complete the task using his technique which, according to him, just took into account some of the information presented for each profile. Same as Raphael, he picked one of the non-engineers as an engineer even when he said that his criteria, once again, was to divide the group according to the
undergraduate field of study. He explained that the remaining time constraint tended to stress him and, by the end, the information he used to select or classify the studied profiles was reduced to half, giving relevance to the obtained scores and the professional background. He asked for more color pens and no markers as the smell could be distractive.

Jenny completed the task in about 80 minutes with 10 minutes to fill the form and confirm her selection. She made a few notes of the relevant profiles and processed all of them in a quick but, according to her, meaningful technique. She started using sticky notes with the first 15 profiles but stopped using them as it could be confusing, distractive, and not very useful tool.

A reduction of stress in the participants was observed, all of them completed the task within the expected time. Omitting data while processing information and misleading individual differences of the profiles allowed confusion in the participants, inducing bias as profiles were misplaced showing a contradiction between the argument and the selection.

After processing the data obtained in the third stage, a meeting was set with people involved in different STEM activities at the School of Education, constituting a panel of experts. The established protocol and the results obtained were presented, including a brief background to show the relevance of this and further studies in this specific field. Once the audience heard carefully about the process we followed as researchers, there had a space for questions. All of the questions were related to details of the process, none of them about misunderstanding the procedure or criteria to select the participants of the previous stages. A specific question was focused on the available resources to complete the task. The person asked for a virtual database with the same information available at the summary cards in order to allow people with high technical skills to process the information in their own way. Also, another person commented that it would be helpful to give the informed consent 24 hours in advance to the study so that the participant could have time to complete a full and detailed reading of the consent that maybe, because of the pressure of time could not be accomplished just before the study. No further comments different from these, and positive feedback of the study were received during this session.

Peer evaluation was a successful method in the protocol design as it showed repeated and supported comments with the presented results and the new suggestions. With this methodology implemented in the design process, the validity of the protocol could be reinforced. Also, associated bias related to the audience’s viewpoint was triangulated, reducing the possibility of misunderstanding the objective of the study when receiving the suggestions.

On the fifth stage of the study, the researchers took into consideration all the relevant observations made by the participants of the previous stages to complete and define the final protocol of the Purposeful Sampling Activity. In addition to previous changes to the protocol,
some new adaptations were made: The participant would not be given a specific method to complete the task. No restrictions or comments about taking notes or any other technique they may or may not use were given. There was determined that, in case of being requested, 10 minutes extra would be given to fill the form. Also, a database was built in Microsoft Excel® with all the information summarized in the cards. This was kept in a USB formatted for both MacOS and Microsoft Windows. Finally, the informed consent was formatted in PDF to be distributed in advance, through secure webmail, directly to the participant of the study.

Conclusions

The protocol was obtained using a co-evaluated method in five stages consisting of three external and two internal revisions. The protocol was piloted in order to reduce stress and anxiety of the participant selecting the purposeful sampling; however, it seeks to measure how participants deal with both and solve the challenge regardless of time constraints, size of the population to be sampled, and the complexity of the task. Resources were adapted and improved for the protocol, as the stages were advancing, to reduce the omission of information during the data processing. A diversity of resources included in the protocol and modifications at each stage were validated by different methods in order to improve the quality of the study and for assuring the success of the protocol design for future applications. Multiple replicates using verbal reports were made to dive validity of each of the external revision stages.

This protocol is part of a larger study that analyzes the differences in perceptions and associated bias to misconceptions, emotions, and experience in engineering and non-engineering experts. Applications for this protocol may include an understanding of procedures and bias of purposeful sampling in research, human resources management, and personal selection. The impact of the bias reduction in the purpose sampling could lead to objectivity obtained by probabilistic sampling subject to future studies.

References


